XVII JORNADAS DE INGENIERÍA TELEMÁTICA

CÁCERES
2025

JUNTA DE EXTREMADURA
Consejería de Economía, Empleo y Transformación Digital

Departamento de Ingeniería de Sistemas Informáticos y Telemáticos
Universidad de Extremadura

Escuela Politécnica

AYUNTAMIENTO cáceres

UNIVERSIDAD DE EXTREMADURA

SCITEL
Sociedad Científica de Ingeniería Telemática

Cátedra de Ciberseguridad
INCIBE - UEx - CUMe

CÁTEDRA INCIBE-UEX-EPCC

LIBRO DE ACTAS

UNIVERSIDAD DE EXTREMADURA

jitel2025.unex.es

XVII JORNADAS DE INGENIERÍA TELEMÁTICA

Actas de las XVII Jornadas de IngenieríaTelemática
(JITEL 2025)
Cáceres (España),12-14 de noviembre de 2025

UNIVERSIDAD DE EXTREMADURA

Caceres, 2025

Esta obra ha sido valorada, evaluada y aprobada su publicación por el Consejo Asesor del Servicio de Publicaciones de la Universidad de Extremadura.

1ª edición, 2025

Edita:

Universidad de Extremadura. Servicio de Publicaciones
Plaza de Caldereros, 2. 10003 Cáceres (España)
Tel. 927 257 041
publicac@unex.es
http://publicauex.unex.es/

E- ISBN.: 978-84-9127-339-4 (edición digital)

Libro electrónico en acceso abierto en el Repositorio Institucional de la Universidad de Extremadura

Presentación

Desde su creación hace 28 años, JITEL se han convertido en unas jornadas de referencia en el área de Ingeniería Telemática en España, abarcando los avances en las tecnologías de las redes del futuro, con énfasis en tecnologías habilitadoras, arquitecturas, servicios telemáticos e innovación en docencia en telemática.

En los últimos años, las redes han continuado su profunda transformación, evolucionando hacia infraestructuras altamente adaptativas que difuminan los límites entre comunicación, computación y control. La convergencia de los principios de diseño cloud-native, las redes basadas en intenciones y autónomas, y la orquestación impulsada por IA está permitiendo que las redes sean más resilientes, sostenibles y centradas en el usuario. Paradigmas emergentes como los Gemelos Digitales de redes, la automatización zero-touch, la integración multicamino y multidominio, así como las aplicaciones inmersivas conscientes de la red, están redefiniendo cómo la conectividad da soporte tanto a humanos como a máquinas. Al mismo tiempo, el impulso global de la investigación en 6G está acelerándose, con visiones de sistemas terrestres y no terrestres integrados, rendimiento adecuado para experiencias inmersivas y nuevos requisitos de confianza, privacidad y sostenibilidad que están dando forma al Internet de las siguientes décadas.

Esta décimo séptima edición de las Jornadas (JITEL 2025), celebrada en Cáceres durante los días 12, 13 y 14 de noviembre de 2025, ha sido organizada por el área de Ingeniería Telemática de la Universidad de Extremadura y la Sociedad Científica de Ingeniería Telemática (SCITEL), con el apoyo y patrocinio de la Junta de Extremadura.

En este libro de actas se recogen las 79 contribuciones aceptadas en las jornadas, cuyas sesiones se organizan en los siguientes ámbitos temáticos: Análisis de datos de redes, Aplicaciones y servicios, Docencia en telemática, Gestión y operación de redes y sistemas, Inteligencia Artificial para redes, Investigación en tecnología educative, Multimedia, salud y sociedades digitales, Redes dedicadas, Redes de nueva generación, seguridad en Comunicaciones, redes y sistemas; además de una sesión de pósters interactiva para fomentar el intercambio de líneas de investigación y exponer trabajos ya publicados.

La organización desea expresar su agradecimiento a todos los participantes y colaboradores que han hecho posible este evento. Confiamos en que las investigaciones y resultados presentados inspiren nuevas ideas y colaboraciones que moldeen el futuro de la Ingeniería Telemática.

Comité de Organización

Presidente del Comité de Organización

- Jaime Galán Jiménez (Universidad de Extremadura)

Miembros del Comité de Organización

- Javier Carmona Murillo (Universidad de Extremadura)
- David Cortés Polo (Universidad de Extremadura)
- Jesús Calle Cancho (Universidad de Extremadura)
- María Mar Ávila Vegas (Universidad de Extremadura)
- Lorenzo Martínez Bravo (Universidad de Extremadura)
- Francisco Javier Rodríguez Pérez (Universidad de Extremadura)
- José Antonio Gómez de la Hiz (Universidad de Extremadura)
- Juan Luis Herrera González (Universidad de Extremadura)
- Jesús Galeano Brajones (Universidad de Extremadura)
- Sergio Frejo Martín (Universidad de Extremadura)
- Juan Luis Arenas Sánchez (Universidad de Extremadura)

Comité de Programa

Presidente del Comité de Programa

- Javier Carmona Murillo (Universidad de Extremadura)

Miembros del Comité de Programa

- Iria Estevez Ayres (Universidad Carlos III de Madrid)
- Elisa Rojas (Universidad de Alcalá)
- Javier González Castaño (Universidad de Vigo)
- Jaume Ramis (Universitat de les Illes Balears)
- María Canales (Universidad de Zaragoza)
- Víctor Corcoba Magaña (Universidad de Oviedo)
- Xavier Larriva Novo (Universidad Politécnica de Madrid)
- Xavier Hesselbach (Universitat Politècnica de Catalunya)
- Ramón Sánchez Iborra (Universidad de Murcia)
- Francisco Luna Valero (Universidad de Málaga)
- Carlos Ramírez (Universidad de Las Palmas de Gran Canaria)
- Ramón Agüero (Universidad de Cantabria)
- Jorge Navarro (Universidad de Granada)
- Víctor Carneiro Díaz (Universidade da Coruña)
- Miguel Bote (Universidad de Valladolid)
- Mariví Higuero Aperribay (Universidad del País Vasco)
- Vincenzo Mancuso (IMDEA Networks)
- Guiomar Corral Torruella (La Salle – Universitat Ramon Llull)
- Antonio Estepa Alonso (Universidad de Sevilla)
- Pilar Manzanares López (Universidad Politécnica de Cartagena)
- Santiago Felici Castell (Universitat de València)
- Jaime Lloret (Universidad Politécncia de Valencia)

Revisores

- Javier Carmona Murillo
- Santiago Felici-Castell
- Cristina López-Bravo
- Jose Salazar
- Antonio Estepa
- Luis Cruz-Piris
- Vinie Lee Silva Alvarado
- Manuel Domínguez-Dorado
- Pilar Manzanares López
- Llorenç Huguet
- Rosa Pericàs Gornals
- Vincenzo Mancuso
- Jorge Mata
- Ramon Sanchez-Iborra
- David Franco
- Ignacio Martínez-Ruiz
- Jose David Fernández Rodríguez
- Andreu Pere Isern-Deyà
- Marisa Catalan
- Noelia Perez Palma
- Jorge Gallego-Madrid
- Jose Ruiz-Mas
- Álvaro Paricio
- David Melendi
- Jaime Galán-Jiménez
- Xabiel García Pañeda
- Mar Avila

- José Ángel Irastorza
- Juan Asensio-Pérez
- Jose Arco
- Sandra Sendra
- Guiomar Corral Torruella
- Antonio Bazco-Nogueras
- Xavier Hesselbach Serra
- Victor Carneiro
- Juan Manuel Lopez-Soler
- Ane Sanz
- Sara Matheu
- Jesús Martínez
- Felip Riera-Palou
- Luis Sanchez-Fernandez
- Jesús Díaz-Verdejo
- Francisco Luna
- M. Francisca Hinarejos
- Francisco-Javier Rodríguez-Pérez
- Victor Villagra
- Andres Munoz-Arcentales
- Jorge Navarro Ortiz
- Juan Antonio Martinez
- Manuel Rodríguez
- Jesús Galeano-Brajones
- Juan-Antonio Carral
- Mehdi kimiaee
- Roberto García Fernández
- Jesús Calle-Cancho

- Dan Garcia-Carrillo
- Eduardo Gómez Sánchez
- Elisa Rojas Sánchez
- Miguel Garcia-Pineda
- Guillermo Azuara
- Luis Diez
- Juan Ramos-Munoz
- Mario Sanz Rodrigo
- Cristian Martin
- Pablo García Zarza
- M. Magdalena Payeras-Capellà
- Josemaría Malgosa-Sanahuja
- Juan Burguillo-Rial
- Asier Atutxa
- Josep Pegueroles
- Celeste Campo
- Natalia Chinchilla-Romero
- Francisco Ortega Zamorano
- Diego Fernandez
- Carlos Garcia Rubio
- Enrique Costa-Montenegro
- Pablo Sotres
- Juan Marcos Ramirez
- Oscar Araque
- Francisco Jaime
- Federico Simmross-Wattenberg
- Iria Estevez Ayres
- iman valizadeh

- Roberto Magán-Carrión
- Florina Almenares
- Jorge Sasiain
- Oscar Esparza
- Fidel Cacheda
- José García Moros
- Marco Álvarez
- Luis Sanchez
- Macià Mut-Puigserver
- Antonio Soriano Asensi
- Francisco de Arriba
- David Cortés-Polo
- Antonio Ruiz Martínez
- Marivi Higuero Aperribay
- Luis de la Cruz
- Jaume Segura Garcia
- Diego López Pajares
- Carlos Alario-Hoyos
- Miguel Zaragoza
- Alberto Bazán
- Carlos Lentisco
- Lorena Mehavilla Martín
- Antonio Estepa Alonso
- Juan Pedro Muñoz-Gea
- Francisco González Castaño
- Jose Ramon Gallego
- Vicente Mayor
- Diego Rivera

- Alberto Eloy García Gutiérrez
- Félix Delgado-Ferro
- Jorge Lanza
- Delia Rico
- Ramón J. Durán Barroso
- Pablo Fondo-Ferreiro
- Samira Amiri
- Maider Huarte
- Ramón Agüero Calvo
- Jesús Fisteus
- Juan José López Escobar

Keynote

Jorge Crichigno (University of South Carolina, USC)

Dr. Jorge Crichigno is a Professor in the Molinaroli College of Engineering and Computing at the University of South Carolina (USC). He also serves as the Director of the Cyberinfrastructure Lab (CILab), where his research focuses on domain-specific accelerators (DSAs) for high-speed networks and cybersecurity. His work includes offloading real-time applications to packet-processing, dynamic programming, regular expression, and tensor-core/AI accelerators. Dr. Crichigno's research has been supported by private industry as well as U.S. federal agencies, including the National Science Foundation (NSF) and the Office of Naval Research (ONR). He has authored over 100 peer-reviewed publications and has received three best paper awards. Dr. Crichigno earned his Ph.D. in Computer Engineering from the University of New Mexico, USA, in 2009.

Domain-specific Accelerators and Their Applications in Computer Networks

For decades, Moore's Law—the observation that the number of transistors on a chip would double approximately every two years—drove consistent improvements in computing performance. This trend allowed developers to rely on advancements in general-purpose CPUs to run their applications more efficiently with each hardware generation. However, Moore's Law has recently stalled, leading to a slowdown in CPU performance gains just as computing-intensive applications, particularly those powered by artificial intelligence (AI), are rapidly expanding. As a result, simply waiting for the next generation of CPUs to deliver faster performance is no longer a viable strategy. To address this challenge, the computing industry has increasingly shifted from general-purpose processors to domain-specific architectures (DSAs) or accelerators, which are tailored to meet the performance demands of specialized workloads.

This presentation will discuss the use of DSAs in computer networks, including applications for cybersecurity and network optimization. DSAs are specialized hardware engines designed to accelerate operations within specific domains. The presentation will begin with an overview of increasingly popular DSAs and related technologies, including Packet-Processing Accelerator (PPA), Regular Expression Accelerator (RegEx), Dynamic Programming Accelerator (DPA), and P4, a domain-specific language for packet processing. Following the overview, the presentation will highlight several applications developed at the University of South Carolina: 1) Self-configuring network applications – These applications leverage the PPA's ability to generate fine-grained network measurements and perform closed-loop control, enabling real-time reconfiguration of buffers, queues, and security policies. 2) Anomaly detection applications – These applications use machine learning (ML) algorithms running on PPAs and RegEx accelerators to detect and classify cybersecurity threats, including malware and distributed denial-of-service (DDoS) attacks. 3) Network resilience applications – These applications use DPAs to implement optimal routing strategies that quickly

identify alternative paths in response to link failures, enhancing network fault tolerance. Each of these applications demonstrates both novel approaches and performance improvements made possible by DSAs.

Índice de trabajos

1 vFFR: A Very Fast Failure Recovery Strategy Implemented in Devices With Programmable Data Plane 16

2 HATHOR: Huerto hidropónico terapéutico 17

3 UEBA design for insider detection within IAM systems 21

4 Detección de intrusiones basadas en anomalías para alumbrado público inteligente (Resumen) 29

5 Codificación de acceso aleatorio aplicable a la trazabilidad de documentos 30

6 Aula invertida en docencia remota a nivel de máster: experiencia y resultados obtenidos 34

7 eBPF-Hub: gestor de funciones eBPF 41

8 Evaluación de Mecanismos de Operación Multicanal para Redes Vehiculares 49

9 Comparación de técnicas de calibración para el sensor O3 del módulo ZPHS01B de calidad del aire en Nodos IoT 57

10 Simulador web mulsemedia en 3D 65

11 Plataforma web mulsemedia 73

12 Platform for Gathering QoS-related Metrics for Training AI-based Media Synchronization Models 81

13 Analysing the Effectiveness of CRM Processes in Acceleration Services: The Business process of Research Infrastructure Catalogue in Unite! Agora platform 87

14 Evaluación y Mejora de los Sistemas de Detección de Intrusiones en IoMT: La Importancia de los Entornos Específicos 94

15 MuHoW: Distributed protocol for resource sharing in collaborative edge-computing 95

16 Beyond 5G Satellite Transmission and Indoor UAVs in Dams Monitoring: An Opportunistic Data Collection Strategy for Isolated Rural Areas 96

17 Proposing AI-based Network Forecasting in Proactive Sync for Shared Media 97

18 DE-RCM: Desanonimización Explicable de MACs Aleatorias en 802.11 WLANs 19 105

Control en bucle cerrado para mitigar alucinaciones en modelos de Inteligencia Artificial. Aplicación a pacientes virtuales. 109

20 Línea de investigación en "Análisis del sesgo de género en las herramientas de IA Generativa". Grupo Lovelace-Lena. 113

21 Orchestrated Service Provisioning in the GÉANT Global Platform for Labs 114

22 Paving the Way for Digital Twins in Network Operations and Management 118

23 Análisis de Algoritmos Post-Cuánticos de Cifrado en Plataformas Heterogéneas 119

24 Diseño e implementación de una arquitectura abierta basada en la nube para planificar rutas personalizadas en función de la calidad del aire 123

25 Gestión Segura de Datos Médicos en Entornos IoMT mediante Blockchain y Smart Contracts 131

26 Integración de Herramientas de IA Generativa en Tareas Académicas: un Proyecto de Innovación Docente en Ingeniería Telemática 132

27 Estafa a través de ataques BEC: Un análisis de casos reales 140

28 Hacia una Arquitectura para Apoyar el Aprendizaje Ubicuo en Múltiples Dominios Utilizando Datos Abiertos Enlazados 148

29 Mejora del posicionamiento Wi-Fi RTT mediante clustering y desviación estándar 152

30 Autoorganización de servicios inteligentes basada en Grafos 153

31 Virtualización de arquitecturas de voto electrónico 157

32 Bienestar digital: Detección y bloqueo de flujos IP en redes Wi-Fi 161

33 Pronóstico del rendimiento citrícola con IA híbrida, imágenes captadas por drones y variables climáticas 165

34 Metodología docente para el análisis de tráfico de red mediante sistemas inteligentes 173

35 Evaluación de sistemas de bajo coste para medir consumo energético de algoritmos criptográficos en microcontroladores 181

36 Una arquitectura de votaciones electrónicas: Un anillo para democratizarlos a todos 185

37 Deployment, Analysis and Optimization Simulator for 5G Wireless Networks 189

38 Propuesta de mejora de Eclipse Zenoh sobre protocolos multi-flujo 194

39 Aprendizaje learning by doing y evaluación sin estrés para la enseñanza de protocolos de enrutamiento 198

40 Rendimiento académico y asistencia: Un estudio con métodos estadísticos y LLM **206**

41 Transmisión de vídeo HD en Canal Físico Restringido usando Tiled Stream Time Division Multiplexing y evaluación de su calidad **212**

42 Proposal of Architecture for Extended Reality Network Functions in Digital Twin Networks **219**

43 Migración en vivo de microservicios con estado para alta disponibilidad en redes asistidas por drones **223**

44 Sistema de Reconocimiento Facial Federado en Conjuntos Abiertos basado en OpenMax **231**

45 Arquitectura para la gestión de offerings en un mercado de activos descentralizado **235**

46 Enfoque práctico simplificado de integraciones IT-OT seguras en entornos industriales **243**

47 A Reinforcement Learning-Based Telematic Routing Protocol for the Internet of Underwater Things **249**

48 INCA-LLM: Innovative Network Slicing in 5G/6G Core Auto-Configuration with LLM **257**

49 Vulnerabilidades en Redes Intra-Vehiculares Modernas: Un Análisis de Ataque de Denegación de Servicio sobre Automotive Ethernet **261**

50 Propuesta de implementación de catálogos federados para espacios de datos sobre DataHub **265**

51 Towards a quantum-safe 5G: Quantum Key Distribution in core networks **273**

52 Interpretación de vídeos mediante Inteligencia Artificial. Estudio preliminar para la detección de errores en cuidados médicos o de enfermería **274**

53 Fomentando el interés científico en las aulas mediante el Aprendizaje Basado en Juegos en áreas STEM: Matemáticas, Pensamiento Computacional y Ciberseguridad **278**

54 Secure and Integrated LoRaWAN-5G Platform for Unidirectional Communications **282**

55 Un Sistema de Detección de Intrusos para IoT basado en DNN y Preprocesamiento Híbrido **290**

56 Entendiendo Wazuh: Análisis mediante redes complejas **300**

57 Prueba de concepto de comunicaciones semánticas para aplicaciones inmersivas 304

58 Herramienta de Preprocesamiento de Tráfico IoMT para Generación Configurable de Datasets 308

59 Sobre la robustez de LLMs contra ataques de envenenamiento en entornos federados 316

60 La Ingeniería Telemática en la Economía Circular: Digitalización de una Planta de Reciclado de CO2 320

61 PALABRIA: Language Learning Enhancement through Linguistic Rule-Based Analytics and Artificial Intelligence 324

62 Redes para un futuro con bajas emisiones: Optimización del enrutamiento consciente del carbono en redes SDN 332

63 Un enfoque simplificado de trazabilidad de criptoactivos con Page Rank RWR, PCA y K-means 340

64 Vehículo Terrestre Autónomo para la Automatización de Auditorias de Seguridad Wi-Fi 346

65 Plataforma escalable de honeypots IoT en Azure: captura, análisis y respuesta automática frente a botnets Mirai y Gafgyt 354

66 Desarrollo de un sistema eficiente para asignar bloques de recursos en Slices 5G, utilizando Aprendizaje Por Refuerzo 359

67 Riesgos de seguridad en Open RAN: Vulnerabilidades, deficiencias y mitigación 367

68 Tráfico de sensado en la interfaz E2: integración de nodos radio 6G en O-RAN 375

69 Comparación de Algoritmos de Aprendizaje Automático para Predicción de Contaminación en la Ciudad de Cartagena 379

70 Políticas avanzadas de scheduling MAC para tráfico XR en redes 5G y 6G 387

71 Laboratorio Abierto Científico-Tecnológico de Investigación en 6G de la UPC 394

72 Despliegue de actualizaciones y comunicaciones seguras en redes IoT 402

73 Modelado e Implementación de un problema de Virtual Network Embedding en redes O-RAN 410

74 Arquitectura para el Slicing Dinámico en Redes Privadas 5G 414

75 5G-Enabled Industrial Mobile Robotics: Integration, Coordination, and Performance 422

76 Computación Autonómica para el Uso Eficiente de Energía en Redes de Sensores Destinadas a la Monitorización Ambiental 429

77 Impacto de la distancia de visualización en la percepción de calidad de representaciones humanas volumétricas comprimidas en entornos de realidad virtual436

78 Aproximación hacia arquitecturas modulares asistidas por computación en el borde para optimización de comunicaciones holográficas 444

79 Optimización del Despliegue de SDNs Mediante Algoritmos de Colonias de Hormigas 453

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

vFFR: A Very Fast Failure Recovery Strategy Implemented in Devices With Programmable Data Plane

David Franco, Marivi Higuero, Ane Sanz, Juanjo Unzilla, Maider Huarte, Eduardo Jacob
Departamento de Ingeniería de Comunicaciones,
Universidad del País Vasco / Euskal Herriko Unibertsitatea (UPV/EHU)
Plaza Ingeniero Torres Quevedo, 1, 48013.
{david.franco, marivi.higuero, ane.sanz, juanjo.unzilla, maider.huarte, eduardo.jacob}@ehu.eus

This paper [1] proposes a very Fast Failure Recovery (vFFR) strategy that can recover from link failures in sub-millisecond timescales by reacting directly from the data plane of the network devices while maintaining a synchronized state with the centralized controller. The design of the vFFR strategy combines three data plane recovery algorithms to reduce latency and packet loss under varying network conditions. The vFFR strategy has been modeled in P4 language, tested on an emulation platform, and validated on Intel Tofino devices.

***Palabras Clave*—Link failure recovery, programmable data plane, P4, SDN, 6G**

I. FAILURE RECOVERY STRATEGY

The operation of the vFFR strategy is divided into three main phases, as depicted in Fig. 1. As a first step, following a proactive failure recovery approach, the centralized controller **computes and installs** the most optimal primary and alternate routes. Second, in the **fast failure reaction** phase, a network device detects and notifies the failure to its own data plane at a minimum delay. The data plane is then responsible for starting the recovery phase. Third, in the **fast failure recovery** phase, the information about the failure is propagated to the adjacent nodes and the alternate paths are activated directly in the data plane, causing the lowest impact on the network performance. This phase is also responsible for the synchronization between the centralized controller and the data planes, thereby enabling them to share a common view of the network. Once the controller receives the failure notification, the computation and installation phase is initiated in order to recalculate alternate paths.

II. PERFORMANCE ANALYSIS

The vFFR strategy is modeled for a well-known emulation platform and compared to a state-of-the-art failure recovery strategy. This testbed is focused on evaluating the performance of the three data plane algorithms in a complex topology under different traffic conditions. The vFFR provides stable values of packet loss and latency

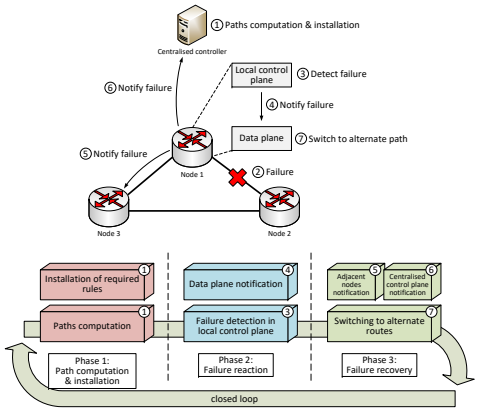


Fig. 1. The vFFR strategy in three phases: path computation & installation, failure reaction and failure recovery.

in the event of link failure and demonstrates superior performance compared to state-of-the-art failure recovery strategies under a variety of network conditions.

On the other hand, the vFFR strategy is modeled for a P4-programmable switch and tested in a simple topology in order to validate the main features of the vFFR strategy in a high-performance switching device. The vFFR strategy has been modeled in a P4-programmable hardware switch and tested using real traffic generators. In addition to latency and packet loss, this testbed estimates the overall failure recovery time of the vFFR strategy, achieving values lower than 500 μs in all cases, including those with a 35 Gbps load.

ACKNOWLEDGMENTS

This research was supported by the Spanish Ministry of Science and Innovation (Grant No PID2022-137329OB-C44) and by the SNS JU (Grant No 101096456).

REFERENCIAS

[1] David Franco, Marivi Higuero, Ane Sanz, Juanjo Unzilla, Maider Huarte, "vFFR: A Very Fast Failure Recovery Strategy Implemented in Devices With Programmable Data Plane", IEEE Open Journal of the Communications Society, 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

HATHOR: Huerto hidropónico terapéutico

Andrade-Garda J, Carneiro V
Centro de Innovación en Tecnologías de la Información y las Comunicaciones (CITIC). Dpto. Ciencias de la Computación y Tecnologías de la Información
Universidade da Coruña
Campus de Elviña s/n, 15071, A Coruña (Spain)
javier.andrade@udc.es, victor.carneiro@udc.es

En este trabajo se presenta un sistema semiautomatizado de cultivo hidropónico cerrado de interior que se ha desarrollado como elemento terapéutico para personas con discapacidad. La integración de sensores, actuadores y una arquitectura basada en Home Assistant sobre una Raspberry Pi 5 ha permitido un control completo y autónomo del entorno de cultivo. Por su parte, la app terapéutica, desarrollada para garantizar la accesibilidad cognitiva y motriz, ha permitido que usuarios con distintos perfiles de discapacidad interactuasen activamente con el huerto. El sistema actualmente está en un entorno operativo real con participación activa de usuarios y desarrollo efectivo de cultivos.

***Palabras Clave*- automatización, domótica, hidroponía, terapia**

I. Introducción

La hidroponía es una técnica de cultivo donde el sustrato no es la tierra. Los nutrientes y el soporte que necesitan las plantas son proporcionados en este caso por una solución acuosa de nutrientes y un sustrato inerte para dar soporte y retener dichos nutrientes. Así, los cultivos requieren menos espacio, regadío, dosis de abonos, fertilizantes y herbicidas, son menos propensos a plagas (sobre todo en invernadero; lo que minimiza pesticidas) y de producción más ágil.

Los orígenes de esta técnica se remontan a las civilizaciones Azteca y Maya y en 1929 Gericke publicó cultivos muy exitosos donde la tierra fue reemplazada por agua con nutrientes y abonos [1]. En la actualidad, los informes de la FAO (Organización de las Naciones Unidas para la Alimentación y la Agricultura) sobre la falta de alimentos en un contexto de crecimiento demográfico mundial han animado el interés por esta técnica de cultivo y, evidentemente, su automatización en un mundo digital. Así, son muchos los trabajos sobre automatización de cultivos hidropónicos. Un ejemplo reciente lo constituye [2], que es un trabajo de la Univ. de Zaragoza focalizado en el cultivo de la lechuga en su entorno. Este ejemplo refleja la situación actual en este ámbito: desarrollos (y automatización) particularizados y adaptados al tipo de cultivo considerado. En [3] también se han considerado apps para la automatización.

Dentro de los sistemas hidropónicos actuales destacan los denominados cerrados, donde la solución nutritiva circula por toda la extensión del cultivo y, al final, se recoge y almacena para volver a reutilizarse. En este caso, la solución debe ser controlada para reponer el agua y los nutrientes consumidos pero también para mantener las condiciones adecuadas para cada cultivo: pH, EC (electro-conductividad), salinidad, temperatura, etc. Es aquí donde la monitorización y el control telemático juegan un papel ciertamente crítico y relevante.

A. Huertos terapéuticos

Por otra parte, los huertos (en interiores o no) se están empleado profusamente de forma terapéutica con personas con discapacidad intelectual bajo diferentes denominaciones: huertos inclusivos, terapéuticos, sociales, etc. El objetivo en cualquier caso es integrar las actividades relacionadas dentro del proceso de intervención, fomentando así su autonomía personal, su empoderamiento en la gestión de actividades y su participación social en equipos de trabajo con responsabilidades asociadas. En España, por ejemplo, el CENEAM (Centro Nacional de Educación Ambiental), dependiente del Ministerio para la Transición Ecológica y el Reto Demográfico, difunde en su web un manual de trabajo con experiencias realizadas sobre huertos inclusivos comunitarios donde explícitamente recoge la diversidad funcional como un aspecto clave a contemplar [4]. Como experiencia más reciente (marzo de 2025), cabe destacar el proyecto Huerto inclusivo CiBiS/UAL [5], destinado a personas con discapacidad y/o necesidades especiales, del CiBiS (Centro de Investigación para el Bienestar y la Inclusión Social) de la Univ. de Almería. En dicho proyecto se recoge: “Las denominadas terapias ecológicas surgen como propuesta alternativa y/o de apoyo a las terapias psicológicas más tradicionales y representan un foco de gran interés en la comunidad científica […]. En este marco, los datos científicos sugieren que los huertos o jardines urbanos promueven la interacción e inclusión social, la actividad física y el acceso a la naturaleza, todo lo cual está vinculado a importantes beneficios para la salud mental. Teorías psicológicas como la Teoría de Restauración de la Atención (ART) o la Teoría de Recuperación del Estrés (SRT), defienden que la exposición a entornos naturales puede ayudar a restaurar los recursos

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

cognitivos y aliviar el estrés por lo que los huertos urbanos se postulan como recursos valiosos para promover la salud mental, el bienestar y satisfacción vital, particularmente en el contexto de la rehabilitación neuropsicológica de la discapacidad y del daño cerebral".

B. *Hidroponía y huertos terapéuticos: HATHOR*

El proyecto HATHOR (Hydroponic Automated THerapeutic ORchard) se ha desarrollado dentro de la actividad "Suministro de fabricación para el estudio y desarrollo de la aplicación de alta tecnología en centros de referencia para la atención a las personas con discapacidad", financiada por la Consellería de Política Social e Igualdade de la Xunta de Galicia. HATHOR es un huerto terapéutico hidropónico cerrado de interior y accesible con tecnología domótica y sensórica de control que se ha inaugurado en abril de 2025 en el CAPD (Centro de Atención a Personas con Discapacidad) de Coruña, como primer paso para su potencial despliegue en el resto de centros similares dependientes de la Xunta de Galicia. El objetivo del proyecto fue el desarrollo de un huerto en el que las personas con discapacidad puedan hacer actividades sencillas vinculadas al cuidado del huerto y de disfrute de las fases de crecimiento, floración, cosecha y sensaciones organolépticas asociadas). Dicho huerto es accesible vía Internet para que sus usuarios puedan participar en la responsabilidad de su mantenimiento (riego, abono, regulación de temperatura, etc.) y también en la visualización del progreso de los cultivos cuando lo deseen (en el propio CAPD—presencialmente en su localización o en otro aula—, desde casa, etc.). Para ello, como se describe a continuación, se ha dotado al huerto de sensórica y actuadores que permiten monitorizar y semiautomatizar su funcionamiento, así como de 2 apps para (i) administrar el mantenimiento y (ii) gestionar sencilla e intuitivamente la interactuación de los usuarios.

La razón de elegir para este caso un huerto hidropónico es que sus usuarios (con discapacidad) no tienen conocimientos agrícolas ni es su única actividad (el huerto tiene que ser autónomo y con muy poco mantenimiento), no pueden manejar ciertas herramientas ni tampoco productos químicos (e.g., pesticidas) y, finalmente, el huerto tiene que estar localizado en el interior (un aula). Los principales objetivos terapéuticos perseguidos son facilitar la autonomía básica en tareas asistidas, mejorar la coordinación visomotora y cognitiva, promover la adherencia a rutinas y fomentar la participación activa en contextos cooperativos. Para su evaluación se definieron indicadores clave (KPI): número de acciones autónomas ejecutadas, frecuencia de uso de la aplicación terapéutica y progresión individual medida en tiempos de respuesta y comprensión de estados del sistema. Estos datos se registran automáticamente y permiten valorar el impacto del sistema en el proceso de intervención.

II. Arquitectura del Sistema

A. *Descripción general*

El sistema ha sido desplegado en un entorno interior de condiciones ambientales controladas, empleando una arquitectura hidropónica basada en la técnica NFT (Nutrient Film Technique) integrada en un bastidor modular optimizado para espacios reducidos. La estructura presenta accesibilidad frontal y lateral, lo que facilita la ejecución de operaciones de mantenimiento, inspección y manipulación agronómica. Los usuarios pueden realizar acciones específicas como riego, fertilización y activación del sistema de enfriamiento mediante una aplicación móvil con interfaz adaptada, mientras que las tareas agronómicas manuales (e.g., poda, conducción vegetativa y cosecha) se ejecutan de forma asistida por terapeutas bajo protocolos de accesibilidad y seguridad.

B. *Sensórica y monitorización ambiental*

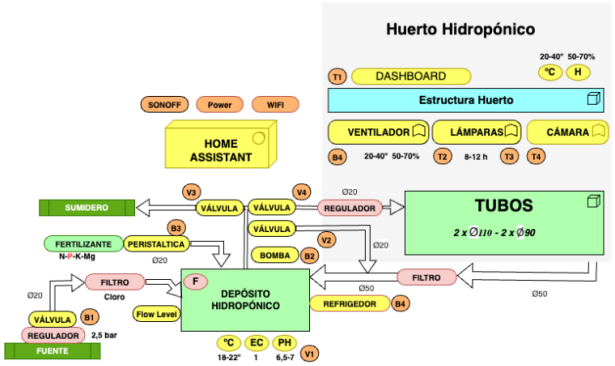


Fig. 1. Arquitectura del sistema

El sistema integra una red de sensores y actuadores interconectados mediante una arquitectura distribuida basada en dispositivos SONOFF [6] controlados a través de la plataforma Home Assistant [7] ejecutada en una Raspberry Pi 5 (Fig. 1), cuya elección se debe a su compatibilidad nativa con Home Assistant y al soporte completo para dispositivos SONOFF a través de protocolos WiFi. Aunque carece de entradas analógicas, esta limitación se resuelve mediante sensores digitales y módulos con salida serie o I2C, lo que permite una integración fiable sin necesidad de conversiones analógico-digitales locales. Así, se permite en conjunto la supervisión y automatización en tiempo real de los parámetros hidropónicos y ambientales, así como la gestión de los elementos mecánicos mediante interfaces WiFi.

Las unidades funcionales establecidas son las siguientes:

- **Bombeo y fertirrigación**: control de entrada hídrica, activación de bomba peristáltica de fertirrigación y módulos de enfriamiento.
- **Válvulas hidráulicas**: regulación de flujo en líneas de entrada/salida, drenaje y recirculación.
- **Seguridad de flujo**: detección y control de paso de agua por la línea principal.
- **Sensor multifuncional de agua**: control de temperatura, pH, EC, sólidos disueltos, salinidad, ORP y gravedad específica.
- **Iluminación y monitorización**: control segmentado de alumbrado y alimentación de cámara IP.
- **Control de luminosidad**: actuadores en persianas enrollables graduadoras de luz natural.
- **Sensor ambiental**: medición de temperatura, humedad relativa y estado de ventilación del entorno.
- **Visión remota**: transmisión de vídeo HD y control motorizado de posición.

Todos los dispositivos están integrados en la arquitectura lógica de Home Assistant como entidades individuales, lo que permite el desarrollo de automatizaciones condicionales, scripts complejos y notificaciones contextuales a través de aplicaciones móviles vinculadas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

C. Control automatizado

El sistema realiza múltiples funciones automatizadas con lógica programada en Home Assistant:

1) **Control dinámico de iluminación**: La intensidad de luz artificial y la apertura de persianas automatizadas se ajustan en función de los niveles de luminosidad ambiente. Esta regulación asegura un fotoperiodo adecuado para el desarrollo vegetal, optimizando el consumo energético y reduciendo el estrés lumínico.
2) **Inyección periódica de fertilizante**: Se realiza la mezcla automática de solución nutritiva en el depósito de fertirrigación mediante bomba peristáltica, controlando tanto volumen como concentración. La inyección se ejecuta en ciclos programados con base en la etapa fenológica del cultivo y las lecturas de EC y pH, garantizando el aporte equilibrado de nutrientes.
3) **Refrigeración adaptativa de agua y aire**: Se activa automáticamente la ventilación forzada y la refrigeración líquida al superarse umbrales definidos de temperatura para mantener rangos óptimos.
4) **Gestión y supervisión de usuarios**: Se registra la actividad de los usuarios desde la aplicación terapéutica (riego, fertilización, enfriamiento) para que los terapeutas evalúen tanto el progreso terapéutico como el uso del sistema.
5) **Notificaciones y gestión de alertas**: Se generan avisos automáticos ante eventos críticos (e.g., fallos en la inyección de nutrientes, detección de niveles anómalos de pH o EC, ausencia de agua en el sistema, bajo nivel de fertilizante o agua y desconexión de sensores clave). Las notificaciones se envían vía aplicación móvil y/o e-mail al personal supervisor.
6) **Modo mantenimiento**: Una rutina específica desactiva temporalmente las automatizaciones críticas para permitir intervenciones seguras sobre el sistema (e.g., limpieza de canaletas, purgado del circuito, recambio de nutrientes y sustitución de componentes).

D. Aplicación de administración

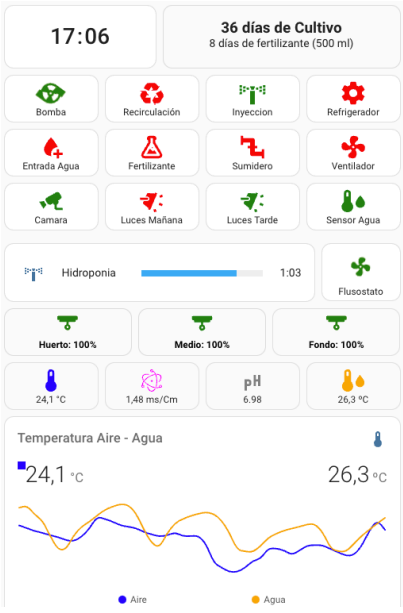


Fig. 2. App Administración

La aplicación de administración proporciona una interfaz centralizada para la supervisión y control integral del sistema hidropónico (Fig. 2). Permite la monitorización en tiempo real del estado de todos los sensores y actuadores, así como la activación manual de estos dispositivos. Integra la visualización histórica de variables críticas como temperatura, pH, EC, salinidad y sólidos disueltos, facilitando el análisis retrospectivo de las condiciones del cultivo. El control de la iluminación y de las persianas puede realizarse tanto de forma automática como manual, en función de los niveles de luminosidad registrados. Adicionalmente, incorpora un módulo de gestión para la cámara de visión remota, la renovación de la solución fertilizante, el control del nivel de reservas de nutrientes y la planificación de fechas clave, como inicio de cultivo y ciclos de fertirrigación. La aplicación también permite activar el modo de mantenimiento para intervenciones técnicas, ejecutar rutinas de reducción de salinidad, aplicar inyecciones puntuales de fertilizante y definir con precisión los tiempos y retardos de cada automatización. Todo el sistema opera sobre una lógica configurable basada en umbrales definidos por el usuario para variables como luminosidad, temperatura y humedad, y ofrece herramientas avanzadas para la creación y edición de scripts, automatizaciones y gestión de perfiles de usuario.

E. Aplicación terapéutica

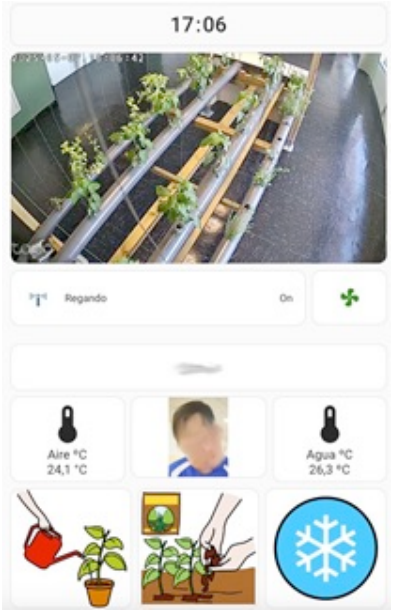


Fig. 3. App terapéutica

La aplicación terapéutica (Fig. 3) ha sido diseña para permitir el acceso controlado, mediante una interfaz adaptada a dispositivos móviles o tablets, a usuarios con discapacidad y siempre bajo la supervisión de un profesional terapéutico. Esta interfaz simplificada facilita la interacción con el sistema, permitiendo a los usuarios consultar el estado actual del huerto y visualizar parámetros ambientales clave como la temperatura del agua y del aire. La funcionalidad se restringe a la ejecución de tres operaciones básicas: activación del riego, inyección de fertilizante y activación del sistema de refrigeración. Todas las interacciones realizadas a través de esta aplicación son registradas automáticamente, lo que permite al terapeuta realizar un seguimiento detallado de la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

actividad individual de cada usuario, incluyendo el tipo de operación ejecutada y el tiempo dedicado a la misma. Esto facilita datos objetivos para la evaluación del progreso terapéutico y la planificación de futuras intervenciones.

III. Resultados

El sistema fue sometido a una fase de validación funcional durante un periodo continuo de seis semanas en un entorno real de uso, con participación activa de usuarios y desarrollo efectivo de cultivos hortícolas. Durante este periodo, se pusieron en operación todas las automatizaciones previstas, sin registrarse fallos en la ejecución de las tareas programadas ni interrupciones del servicio. La precisión de los sensores fue contrastada con instrumentos de medición de referencia, arrojando márgenes de error inferiores al 5 %, dentro de los umbrales aceptables para aplicaciones de horticultura técnica. Las automatizaciones, gestionadas a través de la lógica implementada en Home Assistant, mostraron una latencia operativa despreciable, con respuestas instantáneas ante cambios en las condiciones del entorno o interacciones de los usuarios. La arquitectura basada en Raspberry Pi 5 aseguró una alta disponibilidad del sistema, incluso frente a cortes temporales de conectividad, gracias a su capacidad de procesamiento local y la gestión autónoma de los dispositivos. Como resultado, los cultivos crecieron de forma satisfactoria alcanzando los estados fenológicos esperados, lo que confirma la estabilidad, fiabilidad y eficiencia del sistema.

Durante este periodo, usuarios con distintos perfiles de discapacidad física e intelectual participaron activamente en sesiones estructuradas de interacción con el huerto a través de la aplicación terapéutica. Las sesiones se organizaron de forma colectiva, donde cada usuario gestionó de forma individualizada una parte del sistema bajo supervisión terapéutica utilizando la interfaz adaptada (Fig.2), favoreciendo también dinámicas colaborativas puntuales en tareas comunes como la observación del crecimiento o el debate de decisiones de cultivo. Se registró una mejora progresiva en parámetros de adherencia a la rutina, autonomía operativa y coordinación visomotora. Los registros de uso permitieron al equipo terapéutico realizar un seguimiento cuantitativo de la actividad, identificando patrones de participación y evolución individual, lo que valida el sistema como herramienta complementaria en programas de intervención terapéutica basados en horticultura asistida por tecnología.

IV. Conclusiones y trabajo futuro

A diferencia de trabajos previos centrados en la automatización agrícola o en huertos terapéuticos convencionales, HATHOR integra por primera vez un sistema hidropónico cerrado, domótico y de bajo mantenimiento específicamente adaptado a personas con discapacidad, con doble interfaz (administrativa y terapéutica), control remoto completo y registro automatizado de actividad terapéutica.

El sistema fue validado en condiciones reales de operación, con usuarios finales y cultivos efectivos, confirmándose tanto la estabilidad del funcionamiento como la respuesta adecuada de todas las automatizaciones previstas. La evolución satisfactoria de los cultivos y la interacción fluida de los usuarios mediante la aplicación terapéutica evidencian el potencial del sistema como herramienta multifuncional en contextos terapéuticos y educativos. Actualmente, los KPI anteriormente comentados está siendo procesados y analizados detalladamente por el equipo de intervención terapéutica.

No obstante, durante la fase de validación se identificaron algunas limitaciones relevantes: necesidad de calibración frecuente de ciertos sensores (especialmente el pH), ausencia de integración completa con asistentes de voz y dependencia total de la red eléctrica sin fuentes alternativas de respaldo.

En función de estos hallazgos, se plantean varias líneas futuras: incorporación de algoritmos de inteligencia artificial para la detección temprana de anomalías, realización de estudios clínicos rigurosos que cuantifiquen el impacto terapéutico del uso del huerto e incorporación de fuentes de energía renovable que mejoren la autonomía y sostenibilidad.

Agradecimientos

Cofinanciación CITIC (centro con acreditación de excelencia del Sistema Universitario de Galicia y miembro de la Red CIGUS): Consellería de Educación, Ciencia, Universidades y Formación Profesional de la Xunta de Galicia a través del Programa operativo FEDER Galicia 2021-27 (Ref. ED431G 2023/01). Trabajo: financiado por el Ministerio de Ciencia e Innovación de España (proyecto PID2023-150794OB-I00). Proyecto HATHOR: financiado por la Consellería de Política Social e Xuventude de la Xunta de Galicia (76/2023-ENC01323-OTR2324).

Referencias

[1] Silvio Caputo, "History, techniques and technologies of soil-less cultivation", in *Small Scale Soil-less Urban Agriculture in Europe*, Springer, pp. 45-86, 2022.

[2] Sergio Ramón Moreno, "Greenhouse automation and control for hydroponic cultivation", *https://zaguan.unizar.es/record/128804/files/TAZ-TFG-2023-4552.pdf*, 2023, Trabajo de Fin de Grado. Accedido: 9 de mayo de 2025.

[3] Álvaro Domingo Gordo Ruiz, "Desarrollo e implementación de un invernadero automatizado con cultivo hidropónico y aplicación móvil para el seguimiento de datos", *https://biblus.us.es/bibing/proyectos/abreproy/91393/fichero/Memoria+TFG+Alvaro+Domingo+Gordo+Ruiz.pdf*, 2017, Trabajo de Fin de Grado. Accedido: 9 de mayo de 2025.

[4] Fundación Iniciativas Sur, "Huertos inclusivos comunitarios", *https://www.miteco.gob.es/content/dam/miteco/es/ceneam/recursos/materiales/manual-huertos-inclusivos.pdf*. Accedido: 9 de mayo de 2025.

[5] CiBiS/UAL, "Proyecto Huerto Inclusivo CiBiS/UAL", *https://cibis.ual.es/index.php/proyectos/huerto-inclusivo-cibis-ual*. Accedido: 9 de mayo de 2025.

[6] ITEAD Intelligent Systems Co., SONOFF Official Website, https://sonoff.tech/. Accedido: 16 de mayo de 2025.

[7] Nabu Casa Inc., Home Assistant Official Website, https://www.home-assistant.io/. Accedido: 16 de mayo de 2025.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

UEBA design for insider detection within IAM systems

V. Mayor*, R. Estepa Alonso*, J. Diaz-Verdejo†, A. Estepa Alonso*
* Dpt. Ingeniería Telemática, Escuela Superior de Ingenieros, Univ. de Sevilla
C/ Camino de los Descubrimientos s/n, 41092 Sevilla (Spain)
E-mail: {vmayor, rafa, aestepa}@trajano.us.es
† Dpt. Teoría de Señal, Telemática y Comunicaciones, CITIC, Univ. de Granada
C/ Periodista Daniel Saucedo Aranda, s/n, 18071 Granada (Spain)
E-mail: jedv@ugr.es

Insider threats represent a significant challenge in Identity and Access Management (IAM) systems, where privileged users may misuse their access for malicious purposes. This paper explores key considerations for developing of a User and Entity Behavior Analytics (UEBA) framework specifically designed for IAM environments. Our study examines the essential features of UEBA systems, identifies key challenges related to anomaly detection in the context of IAM, and reviews available datasets that could be useful for future model development.

We present an initial conceptual design based on a thorough review of the literature and available IAM-related datasets. Specifically, we propose a multi-layered UEBA architecture that integrates both individual and group-based behavioral profiling to detect suspicious activities. The system is designed to analyze IAM logs, establish behavioral baselines, and dynamically identify deviations that may indicate insider threats. Additionally, we outline a feedback mechanism aimed at refining detection models over time.

***Palabras Clave*—Insider Threats, UEBA, IAM, Anomaly Detection, Cybersecurity, Artificial Intelligence**

I. Introduction, Motivation, and Objectives

Insider threats represent a significant risk to corporate network security. Disloyal users, known as *insiders*, may engage in activities ranging from intellectual property theft, IT sabotage, fraud, and espionage, to workplace violence and aggression [1]. Their motivations vary nad may include revenge, financial gain, personal conflicts, or poor job performance [2]. Due to their privileged access and familiarity with internal systems, *insiders* can cause far greater damage than external attackers [1], [3]. A recent report [4] indicates that 48% of surveyed organizations experienced an increase in insider attacks over the past year. Furthermore, 51% reported six or more incidents, and 29% incurred remediation costs exceeding one million dollars per case.

To manage user access, medium and large enterprises commonly deploy Identity and Access Management (IAM) systems, which typically rely on Single Sign-On (SSO) protocols [5], such as OpenID Connect (OIDC) [6]. While these systems improve usability and enforce access policies (e.g., authentication factors, retry limits, password expiration), they also introduce cybersecurity challenges, particularly in detecting insider threats.

Current detection strategies [3], [4] include predefined rules and Indicators of Compromise (IoCs) within Security Information and Event Management (SIEM) systems, as well as anomaly detection through User and Entity Behavior Analytics (UEBA). Although rule-based SIEM solutions (e.g., HPE Revenue Intelligence Fraud Risk, CyberArk) support fraud detection and administrator activity monitoring, they require manual fine-tuning and prior knowledge of attack patterns. Insider threats, however, often exhibit **unpredictable** and **evolving** behavior [3], which limits the efectiveness of static rule-based approaches. Moreover, insiders may engage in non-fraudulent yet harmful actions, such as data exfiltration or workplace aggression, which traditional detection systems frequently overlook. As a result, UEBA systems are essential to complement conventional SIEM-based solutions.

UEBA leverages AI-driven behavioral models to detect anomalies based on user profiles, assessing the severity and confidence of potential security threats [7], [8]. Its core workflow involves: (i) collecting user and entity logs, (ii) establishing behavioral baselines for individual users and groups, (iii) quantifying deviations from normal behavior, and (iv) generating risk-based alerts to assist security analysts in decision-making [9].

Despite the availability of commercial UEBA solutions

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

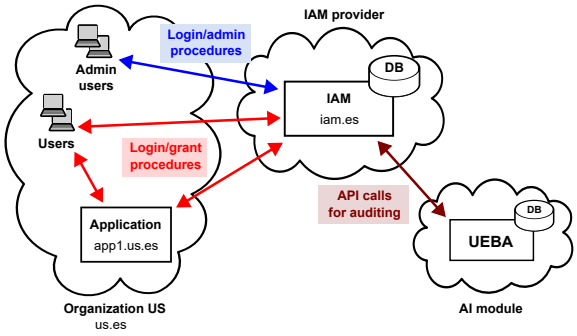


Fig. 1. UEBA integration in an IAM system.

[8] (e.g., Exabeam, Forcepoint UEBA, Fortinet, OneIdentity, Splunk, Microsoft Azure Sentinel, IBM QRadar), most are integrated into SIEM platforms, making them generic and difficult to customize. Developing a **dedicated UEBA system for IAM**, although requiring high initial investment and expertise, offers advantages: (i) tailored anomaly detection based on IAM-specific data, (ii) greater control over AI algorithms, and (iii) improved integration with IAM infrastructures for enhanced threat detection.

A. Research Objectives

Our research aims to develop an effective and flexible UEBA system integrated with IAM platforms, capable of analyzing logs from both regular and administrative users to identify suspicious activities. The intended context for this UEBA system is a multi-tenant IAM platform that manages access to multiple applications via SSO protocols [5], [6]. Fig. 1 illustrates the integration between IAM and UEBA, highlighting specific features relevant to the proposed system.

This work examines the essential features of UEBA systems, identifies key challenges in anomaly detection within IAM environments, and reviews available datasets that may support model development. It also presents an initial conceptual design based on a comprehensive literature review and an evaluation of IAM-related datasets. We propose a multi-layered UEBA architecture that incorporate both individual and group behavioral profiling to detect suspicious activities. The system is designed to analyze IAM logs, establish behavioral baselines, and dynamically identify deviations that may indicate insider threats.

More specifically, we address the following research questions:

- **Features**: What are the core characteristics and AI models used in UEBA systems?
- **Specificities**: What dhistinguishes IAM-specific UEBA systems from general-purpose ones?
- **Datasets**: What datasets are currently available for training and evaluating IAM-specific UEBA models?
- **Architecture**: What UEBA architecture is most suitable for IAM environments?
- **Scope**: What types of threats can UEBA detect, and what challenges remain unresolved?

To address these questions, the paper is organized as follows: **Section II** reviews related work on UEBA systems. **Section III** examines available datasets and proposes an IAM-specific dataset. **Section IV** presents a flexible UEBA architecture for insider threat detection. **Section V** discusses system limitations and scope, followed by conclusions and future work in **Section VI**.

II. Characteristics of UEBA Systems

The topic addressed in this work is relatively novel. After conducting an exhaustive literature review, we found no contributions that specifically define the characteristics of UEBA systems within IAM environments. In fact, UEBA systems themselves remain an emerging technology. For this reason, we have structured the literature review into several subsections, each covering key features and functionalities of UEBA systems, with particular emphasis on their relevance to the present study.

A. Key Features and Anomaly Metrics Used in UEBA

UEBA systems operate based on input data obtained from monitored events. Among the various data sources available, it is essential to select relevant input variables –commonly referred to as *features*– that are critical for anomaly detection in AI models. In [10], a state-of-the-art study categorizes and classifies all possible *features* used in user profiling. These features can be broadly grouped into the following categories:

- User interests (e.g., topics, browsing history),
- Knowledge and skills (e.g., types of IT tools used),
- Demographic information (e.g., gender, marital status, location),
- Intentions (e.g., Internet searches, queries),
- Online behavior (e.g., password creation habits, password change frequency, number of visited websites)
- Offline behavior (e.g., CPU usage, memory consumption, session duration, login frequency),
- Social network activity and network traffic (e.g., flow characteristics, HTTP connections, traffic volume, usage hours).

Some studies, such as [11] and [12], propose UEBA systems using only seven *features*, while others use twelve [13]. At the other extreme, some models employ up to 168 features [14]. A comparative study [15] analyzed results using either a restricted set of seven features or an extended set of sixteen features, finding no significant differences in performance. One key conclusion from this study is that data sources must be normalized, and *feature* selection should be guided by the nature of the anomalies being detected.

Several studies distinguish between *features* and what they term *anomaly metrics*, which are derived from one or more *features*. These metrics are typically designed to detect unseen or unexpected behaviors in a user profile [14], such as ”new device detected” or ”new login location”. Anomaly metrics are commonly defined as ”new occurrence,” ”rate per hour,” or ”total count.”

In IAM environments, where monitored data includes access logs (e.g., browser, IP, geolocation, login events,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

administrative actions), feature selection is inherently constrained. However, this limitation does not necessarily degrade system performance; rather, it retricts the types of anomaly detection models that can be applied.

B. AI Models for Baseline Profiling and Anomaly Detection

Once the input features are defined, it is necessary to determine the types of models used for profiling and anomaly detection. There are two main profiling approaches: individual user profiling and group-based profiling. This involves defining an behavioral model for each user, as well as identifying the group to which the user belongs, which informs the group-level model.

Individual profiling requires sufficient statistical data to train reliable models. However, during the initial phase, performance may be suboptimal due to the lack of historical data –a challenge known as the *cold-start* problem. Group assignment can be based on predefined roles or determined using clustering techniques [17]. Managing users with multiple roles presents additional complexity [18].

Various AI techniques are applicable to UEBA systems. A review in [20] categorizes anomaly detection methods into statistical models (20 studies), neural networks (3), hidden Markov models (3), rule-based learning (2), decision trees (6), support vector machines (7), and fuzzy computing (2), with statistical approaches being the most prevalent. In [21], the authors compare supervised learning techniques (e.g., logistic regression, neural networks, decision trees, Naïve Bayes) with unsupervised methods (e.g., K-means clustering, SVM, DBSCAN), concluding that model selection depends on the characteristics of the input data and the type of anomaly being detected [22].

Anomaly detection models typically produce scoring metrics to assess deviations from expected behavior [16], with Mahalanobis distance [14] being a commonly used measure, often combined with confidence intervals.

Temporal Dynamics in UEBA Models. User behavior evolves over time, introducing two key challenges: *data shift*, where behavioral patterns change [23], and the *cold-start* problem, where new users lack historical profiles [24]. Most studies employ time windows to adapt models dynamically, although this can lead to information loss. In [18], the authors propose adaptive sliding windows with contextual awareness, using algorithms such as FLORA4 [19]. Other approaches include multi-window modeling with adaptive weighting or the use of LSTMs for memory retention [25]. Determining when to retrain models remains an open challenge, with proposed solutions including entropy-based or average deviation metrics [26].

The cold-start problem is a specific case of *zero-shot learning* (ZSL) [27], where models classify instances without prior exposure. The most promising solutions involve leveraging pretrained models and contextual data [24], using similarity-based profiling.

Ensemble Methods and Performance Evaluation. Multiple anomaly detection models (detectors) are often combined into an *ensemble* to improve accuracy. In [18], the authors highlight the benefits of ensemble methods, advocating for specialization-based grouping and weighted majority voting. Feedback mechanisms can further enhance model performance.

Evaluating model performance requires labeled attack datasets. In the absence of such datasets, synthetic attack data must be generated. Some UEBA systems simulate behavioral variations by swapping user profiles [22]. Validation should reflect realistic operating conditions [30], [22], [12]. Given the high cost of false positives (FPs) in Security Operations Centers (SOCs), minimizing FP rates is critical. [29] recomends using False Acceptance Rate (FAR) and False Rejection Rate (FRR) metrics to assess the system's real-world viability.

C. System Output (Alarms) and Feedback

Generated alarms must effectively convey the perceived risk by combining multiple anomaly metrics of varying types and meanings. For example, [22] defines an anomaly metric based on an unusually high number of events, [16] identifies anomalous USB usage, [31] detects abnormal login frequencies outside regular hours; and [29] introduces additional refined anomaly metrics. Using multiple metrics in parallel enhances alarm interpretability and reduces FPs.

Alarms should include a calculated risk level, which considers event priority, asset value, and detection confidence [21]. When a predefined risk threshold is exceeded, the alarm must include at least two *scores*: a severity level (indicating the importance of the event) and confidence level (assessing the reliability of the detection). In [29], a percentile-based *scoring* system is used instead of absolute values, providing insights into external risk, *insider* risk, and anomaly risk. This method quantifies key factors such as the number of distinct anomaly models triggered, the total anomaly *scores*, and the variety of detected anomaly categories.

Most UEBA systems generate a single alarm type with an associated risk indicator. However, some approaches define multiple alarm levels. For instance, [16], [14], [21] propose a three-level hierarchy:

- Level 1: Policy violations and known attacks.
- Level 2: Absolute threshold anomalies.
- Level 3: Statistical deviations.

This hierarchical approach is particularly useful in real-world systems, as it enables alarm correlation mechanisms to streamline information for security analysts, reducing the number of alerts requiring manual review.

Feedback Mechanisms. Feedback is essential for improving system performance. When a false positive occurs, the model should be refined through retraining [22], parameter tuning, or adjusting the weight of specific models in the final decision process [29], [13]. This enables adaptive learning [22], allowing models to evolve alongside legitimate user behavior. [13] proposes dynamically adjusting output weights based on historical alarm analysis. Additionally, event correlation techniques can help minimize unnecessary alerts, improving overall system efficiency.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

D. Challenges of UEBA Systems in IAM Environments

To conclude this analysis, we highlight the key challenges and limitations associated with applying UEBA systems within IAM environments.

One of the most significant challenges is the limited applicability of academic research to real-world industry settings [12]. As noted int [22], academic studies often rely on assumptions that do not hold in practice—such as the availability of clean attack datasets for model training, the representativeness of datasets with respect to real-world conditions, and the stability of user behavior over time. Addressing these issues requires access to sufficiently large and realistic datasets.

Another major challenge is system usability. In non-academic environments, classification errors can be costly, as each false alarm requires investigation by specialized personnel. Therefore, performance metrics must ensure an extremely low false positive rate –ideally well below 1%. Commercial UEBA products have achieved false positive rates as low as 0.01%, with a significantly high ratio of true positives to false alarms. Minimizing false positives to acceptable levels is thus critical for successful real-world deployments.

A further limitation of IAM-based UEBA systems is the restricted availability of input data, which inherently constrains the effectiveness of anomaly detection models. For instance, determining the duration of user sessions within an application is ofter difficult, as many users do not explicitly log out. Even when logout events occur, they may not be recorded by the IAM system. Typically, IAM logs capture authentication events and minor account modifications, such as password resets, email updates, and changes to authentication factors. In contrast, administrative actions —such as user management, application assignment, and policy enforcement-— generate a richer dataset, enabling more sophisticated anomaly detection. Consequently, insider threats originating from administrators are generally easier to detect than to those involving regular users.

III. Review of Existing Datasets for UEBA in IAM Environments

The effectiveness of anomaly models depends heavily on the datasets available and the feature sets they provide. As such, the input dataset is a critical factor in determining the capabilities of a UEBA system, influencing the selection of *features*, the choice of models, and their applicability. Ideally, the collected data should closely resemble real-world IAM environments. However, as previously discussed, this is a relatively new research area, and no publicly available datasets are specifically designed for UEBA systems focused on IAM environments.

To address this limitation, we review existing datasets that may be relevant based on a minimal set of required attributes:

- **Event timestamp**.
- **User identifier** performing the action.
- **User location** (IP address and geolocation data).
- **Resource identifier** accessed or modified by the user.
- **Action performed** (e.g., login, logout, user creation, deletion).

Using these criteria, we filtered existing datasets and evaluated their relevance for IAM-integrated UEBA systems.

A. Sci-Hub Dataset

Sci-Hub is a web repository containing over 100 TB of research articles available for download. Users (authenticated or not) can retrieve documents, with authenticated users' download activities being recorded.

An anonymized log of transactions from 2017 is publicly available at https://zenodo.org/records/1158301, comprising 14 GB of data. It includes:

- Event timestamp (yyyy-MM-dd HH:mm:ss).
- DOI (document identifier).
- Anonymized IP (sequentially numbered).
- Anonymized user ID (sequentially numbered).
- GeoIP data (country, city, latitude, longitude).

Although Sci-Hub is not an IAM system, it provides historical user geolocation data, making it useful for designing location-based anomaly detection models. Additionally, the dataset volume is suitable for statistical modeling.

B. Insider Threat Test Dataset

The Insider Threat Test Dataset, published by Carnegie Mellon University, is a synthetic dataset with labeled normal and malicious activities [32]. Various versions (r1, r2, ...) have been generated, aiming to facilitate research on insider threat behaviors in organizations. It includes simulated records of activities such as:

- Email usage.
- File transfers.
- Access to corporate resources (files and systems).
- Web browsing.
- USB device usage.

The dataset also includes a response file identifying users and time periods associated with malicious activities. It spans 96 GB across six versions (2011–2021), incorporating different data generation strategies.

Despite its popularity in insider threat research, its synthetic nature and limited similarity to IAM environments restrict its direct applicability. Its most relevant use for IAM-focused UEBA is in detecting anomalies in administrator actions and correlating behavioral patterns.

C. Cloud-based User Entity Behavior Analytics Log Dataset (CLUE-LDS)

CLUE-LDS is a dataset designed for UEBA research, containing real user activity logs from a cloud-based environment [22]. Logged events include logins, file access, link sharing, and configuration changes.

This dataset comprises approximately 50 million events from over 5000 users, covering a five-year period (July 7, 2017 – September 29, 2022, spanning 1910 days). All recorded events are available except for 109 missing

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

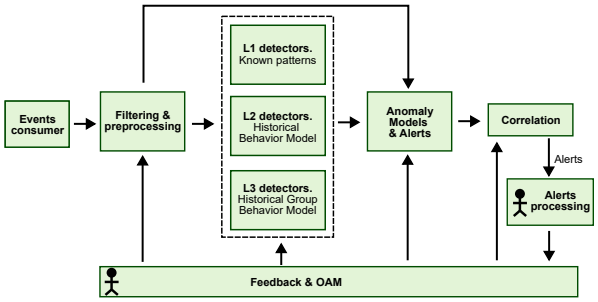


Fig. 2. Proposed UEBA system architecture.

entries due to database failures. The dataset (14.5 GB uncompressed) can be accessed at https://zenodo.org/records/7119953 and is structured in JSON format with attributes such as:

- **id**: Unique event record identifier.
- **time**: Timestamp in ISO format.
- **uid**: Anonymized user identifier.
- **uidType**: User identifier type (e.g., username, IP).
- **type**: Action performed by the user.
- **params**: Additional event parameters (e.g., file paths, user groups).
- **isLocalIP**: Boolean flag indicating internal (true) or external (false) IP.
- **role**: User role classification (*consulting, administration, management, sales, technical, external*).
- **location**: Optional geolocation data (available only for unauthenticated users).

CLUE-LDS shares multiple similarities with IAM logs, including timestamps, user identifiers, event types, and roles. However, the absence of geolocation data for authenticated users limits its utility for location-based anomaly detection.

D. Summary and Key Challenges

The reviewed datasets provide a useful starting point for developing detection models related to IAM environments. They offer essential data elements such as timestamps, user identifiers, location data, resource access logs, and user actions —components that are critical for exploring UEBA systems and potentially adaptating them for IAM-specific threat detection.

Nevertheless, to the best of our knowledge, **no publicly available datasets are explicitly designed for training and evaluating UEBA systems in IAM environments**. Therefore, real-world datasets must be collected from actual IAM deployments to enable performance assessment under realistic operational conditions.

IV. Proposed Architecture

Based on the previous analysis of the literature and the specific requirements of UEBA systems within IAM environments, we define the following key architectural **requirements**:

1) Continuous monitoring of IAM events and relevant contextual information.
2) Feature filtering and local storage of contextual data for event analysis.
3) Asynchronous detection of anomalous behaviors in users and administrators based on behavioral baselines or deterministic models.
4) Adaptive learning mechanisms to refine detection models in response to evolving user behavior.
5) Alert generation based on an anomaly score for each event, derived from detector outputs.

Considering these requirements, this section presents an initial architectural framework for the UEBA module, serving as a foundation for developing various anomaly detectors.

Fig. 2 outlines the primary components of the proposed architecture. At its core is an event ingestion module that interacts with the IAM system via APIs. These events undergo filtering and preprocessing, where only relevant events (used by the anomaly detectors) are retained, and unnecessary data is discarded. Preprocessing includes format normalization and potential data enrichment. The processed data is then analyzed across three anomaly detection layers:

- **L1: Known Attack Pattern Detection.** This layer identifies predefined behavioral patterns that match known Indicators of Compromise (IoCs) or Indicators of Attack (IoAs). For example, it can detect an administrator deleting 500 users within five minutes or a user logging in from two distant countries within a short time frame. Each detector operates independently, filtering and extracting only the relevant event features. For instance, from the *Time* attribute, features such as daily connection count or working-hour logins can be derived. The core detection mechanism evaluates the presence of specific IoC patterns, providing an assessment of their match (full/partial/none) along with contextual insights for further analysis.
- **L2: Individual Behavior Anomalies.** This layer detects deviations from an individual's learned behavioral model. An example would be a user who consistently logs in between 08:00 and 08:30 AM on weekdays but suddenly starts logging in at 03:00 AM. Such behavior deviates from the user's normal pattern. Similar to L1, multiple detectors operate in parallel, each with tailored preprocessing and AI models suited to specific anomaly types.
- **L3: Group Behavior Anomalies.** This layer identifies deviations in a user's behavior compared to the norms of their assigned group. Unlike L2, which evaluates individual patterns, L3 models group-wide behavior. In its simplest form, users can be classified into two groups: standard users and IAM administrators. The system learns typical behavioral patterns for each group and flags users who significantly deviate from their group's norms.

The architectures for L2 and L3 are structurally similar. After initial filtering and preprocessing (as in L1), various

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

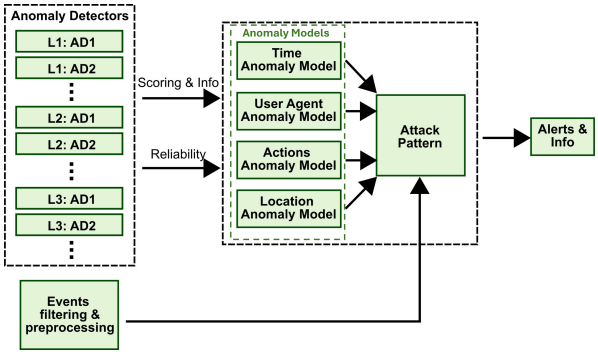


Fig. 3. UEBA Anomaly Models and Alerts.

detectors are deployed based on the nature of the anomaly being assessed. Common AI techniques include Isolation Forest, k-NN, neural networks, and statistical modeling, which are particularly effective for time-series anomaly detection. AI-based models requiring training follow a two-phase approach: a *LEARNING* phase for model training and a *TEST* phase for real-time anomaly detection. Model *Reliability* evaluation is critical to determine the need for retraining. Finally, L2 and L3 detectors generate outputs in the form of an anomaly *Scoring* index, along with contextual information to assist in interpreting the detected anomalies.

The alarms generated by the anomaly detectors in L2 and L3 are aggregated into *ensembles* within the "Anomaly Models" module (Fig. 3). The purpose of this approach is to enable logical combinations of individual detector alarms. For example, an alarm indicating multiple active sessions for a single user may only be considered anomalous if the logins originate from different locations, thus triggering a location-based anomaly alarm. Based on the state-of-the-art review and the specific IAM application context, four distinct anomaly model categories are initially defined: time-based anomalies, user-agent anomalies, location anomalies, and action-based anomalies.

Each anomaly model incorporates multiple detectors leveraging different *features*. The models assign weights to each detector's *score*, factoring in elements such as reliability and detector-specific importance (e.g., impact assessment). The model outputs are then analyzed alongside external security events (e.g., firewall, IDS alerts) to identify patterns indicative of specific attack scenarios. This information is used to generate alerts, which also include additional context to enhance interpretability, ensuring that the reasoning behind each alert extends beyond the AI models responsible for its detection.

Since alerts must be processed by the SOC, providing enriched contextual information improves analyst efficiency, leading to faster resolution times and reduced operational costs. Finally, a correlation and feedback module is incorporated to further minimize false positive rates.

The proposed architecture establishes a multi-layered detection system that leverages IAM event data to identify behavioral anomalies at different levels. By integrating rule-based, statistical, and AI-driven detection mechanisms, the system aims to enhance insider threat detection while minimizing false positives. Future work will focus on optimizing model interpretability and fine-tuning anomaly scoring thresholds to improve real-world deployment effectiveness.

V. Scope of the Initial Implementation and Roadmap

Based on the analysis of available datasets, we define an initial scope for implementing the proposed architecture.

In collaboration with the Cybersecurity Officer of a partner company, we identified the most common threats in an IAM system. From an initial list of 15 threat categories, we determined that IAM-collected data enables the detection of the following specific threats (Table I): phishing, brute-force attacks, token theft, credential harvesting, abuse of overprivileged accounts, password reuse attacks, and social engineering attacks. These define the system's maximum detection scope. Notably, insider-related behaviors within an IAM system are encompassed within this list.

For each of these threats, the UEBA module can define Indicators of Compromise (IoCs) corresponding to the proposed anomaly detection models. These indicators may be implemented at any of the three architectural levels: L1 (known attack patterns), L2 (anomalous behavior relative to a user's baseline), or L3 (anomalous behavior relative to a user's group). These indicators form the basis for the Anomaly Model and alert generation. Table I outlines, for each threat, the potential anomaly detectors to be implemented in the UEBA module, their assigned level, and a score indicating dataset suitability for evaluation (1 = low statistical significance, 5 = highly representative dataset).

This initial list represents the anomaly detectors and models to be explored throughout the project's development. Experimental validation will determine which proposed detectors yield meaningful results and which have sufficient statistical representation in the collected dataset.

As a first approach, anomaly detection modules for standard users will be based on *application usage patterns* and *login context data* (e.g., device, software, IP address, geolocation, and time-based information). For *administrators*, additional monitoring of privileged actions, such as *permission changes, account creation/deletion, and policy modifications*, is required.

It is important to note that in some cases, *high-quality data may not be available* for validation, requiring the generation of *new datasets* to ensure robust model evaluation.

For each anomaly model, an *ensemble* of different detectors will be defined, each targeting specific *features* derived from available data. For instance, in the *time-based anomaly model*, relevant *features* may include:

- Logins at unusual times,
- Daily/weekly access frequency during business hours,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
POTENTIAL IOCS FOR UEBA SYSTEM DEVELOPMENT

ID	Attack	IoC	Indicator Description (Level)	Dataset Score
A1	Phishing	i1	Unusual login location for the user (L2)	4
		i2	Unexpected configuration change (L2)	2
		i3	Unusual user-agent (L2)	3
		i4	Login attempt from an already authenticated user (L1)	-
A2	Brute-force attack	i5	Multiple failed login attempts from different users within a time window (L1)	1
		i4	Login attempt from an already authenticated user (L1)	-
A3	Token theft	i6	Unusual actions for the user (L2)	2
		i7	Multiple active IP addresses for the same user (L1)	2
		i8	Multiple active user agents for the same user (L1)	2
A4	Credential harvesting	i9	Unusual login hours for the user (L2)	5
		i10	Unusual login hours across multiple users within a time window (L2)	5
A9	Abuse of overprivileged accounts	i11	Unusual actions by an administrator (L2)	4
		i12	Unusual behavior within the administrator group (L3)	1
		i13	List of ”abusive” actions (e.g., deletion of 100+ accounts in one day) (L1)	4
A11	Password reuse attacks	i14	New IP address (L1) and unusual IP for the user (L2)	5
		i3	Unusual user-agent (L2)	3
		i9	Unusual login hours for the user (L2)	5
		i15	Unusually high number of successful logins in a short time window (L2)	5
A15	Social engineering attacks	i6	Unusual actions for the user (L2)	2
		i16	Anomalous travel speed (L1)	5
		i17	Login from a new country (based on the 5 most frequent and recent locations) (L2)	4

- Daily/weekly access frequency on weekends,
- Number of logins during vacation periods,
- Login frequency within a 24-hour window.

For each of these *features*, the most appropriate AI models will be selected based on performance with available data. Once the optimal models and feature sets are identified, they will be combined into an *ensemble* to enhance anomaly detection accuracy and reliability. The same methodology will be applied across all defined anomaly models.

Once the anomaly detection models are validated, an *attack attribution and correlation module* will be implemented. This module will:

- Identify the *type of attack* detected (where possible),
- Reduce the *number of alerts* and *minimize false positives*.

Finally, the project will conclude with the implementation of a *feedback loop* for continuous model improvement, along with a *management and configuration layer* to facilitate system integration and deployment.

VI. CHALLENGES AND LIMITATIONS

Based on the data available in the reviewed datasets, training and evaluating accurate models is expected to be challenging due to statistical limitations. Specifically, the number of records containing meaningful information is relatively low. It is anticipated that only L2-level anomalous actions can be reliably detected for administrators, while group-based anomaly detection (L3) will be difficult to implement due to insufficient statistical data for users and a complete lack of such data for administrators.

Additionally, the difficulty in determining session durations (i.e., when a user is actively logged in) and the absence of information on login attempts may significantly hinder the evaluation of the models under development. However, the datasets do provide sufficient statistical representation to support the development of *time-based anomaly models* (L2) for both users and administrators, as well as certain L1-level anomaly detectors.

Another major challenge is the *attribution of Indicators of Compromise (IoCs) to specific attack types*, as different attacks may share common IoCs depending on the models ultimately implemented.

These challenges will be progressively addressed as the implementation advances and meaningful results emerge, allowing for refinement of the research findings.

VII. CONCLUSIONS

This article presents the progress of an R&D&I project focused on integrating an AI-based module for insider threat detection within an Identity and Access Management (IAM) system. A comprehensive literature review and dataset analysis have identified key challenges in implementing UEBA systems in IAM environments. An initial UEBA-based architecture has been proposed to enhance the detection of internal threats, leveraging both anomaly detection techniques and behavioral pattern recognition at the individual and group levels.

An initial specification of anomaly models applicable to IAM has been outlined, covering both generic (e.g., geolocation and time-based anomalies) and IAM-specific models (e.g., administrator activity anomalies). This specification will guide the next implementation phase, during

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

which models will be validated against available datasets. Key considerations for future development include dataset quality, temporal window selection, detection accuracy evaluation, cold-start issues, and the integration of feedback mechanisms.

Future work will focus on refining anomaly detection models, improving system interpretability, and validating the proposed architecture in operational IAM environments. These developments aim to strengthen IAM security by providing a robust and adaptive framework for detecting and mitigating insider threats.

Acknowledgments

This work has been partially funded by the project PID2024-155581OB-C21 funded by the Spanish Ministry of Industry under the contract MICIU/AEI/10.13039/501100011033/ FEDER, UE and by the OREOS contract ("Solución integral de Gestión de Identidades de Nueva Generación") PI-2437/22/2023, funded by WhiteBearSolutions under the project 22/1081 approved by the CTA body.

Referencias

[1] CERT National Insider Threat Center, "Common Sense Guide to Mitigating Insider Threats, Sixth Edition," CMU/SEI-2018-TR-010, Dec. 2018. Available: http://www.sei.cmu.edu.

[2] M. Maasberg, X. Zhang, M. Ko, S. R. Miller, and N. L. Beebe, "An analysis of motive and observable behavioral indicators associated with insider cyber-sabotage and other attacks," *IEEE Engineering Management Review*, vol. 48, no. 2, pp. 151–165, 2020.

[3] N. Saxena, E. Hayes, E. Bertino, P. Ojo, K. K. R. Choo, and P. Burnap, "Impact and key challenges of insider threats on organizations and critical businesses," *Electronics*, vol. 9, no. 9, p. 1460, 2020.

[4] Cybersecurity Insiders, "2024 Insider Threat Report – Gurucul," Available: https://www.cybersecurity-insiders.com/portfolio/2024-insider-threat-report-gurucul/. [Accessed: Jan. 18, 2025].

[5] V. Radha y D. Hitha Reddy, "A survey on single sign-on techniques," Procedia Technology, vol. 4, pp. 134-139, 2012.

[6] N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, y C. Mortimore, "OpenID Connect Core 1.0," The OpenID Foundation, 2014. [Online]. Available: https://openid.net/specs/openid-connect-core-1_0.html

[7] M. Raut, et al., "Insider threat detection using deep learning: A review," in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, IEEE, 2020.

[8] J. Datta, et al., "Real-time threat detection in ueba using unsupervised learning algorithms," in *2021 5th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, IEEE, 2021.

[9] P. A. Savenkov and A. N. Ivutin, "Methods of machine learning in system abnormal behavior detection," in *Advances in Swarm Intelligence: 11th International Conference, ICSI 2020, Belgrade, Serbia, July 14–20, 2020, Proceedings 11*, Springer International Publishing, 2020.

[10] A. H. Lashkari, M. Chen, and A. A. Ghorbani, "A survey on user profiling model for anomaly detection in cyberspace," *Journal of Cyber Security and Mobility*, pp. 75-112, 2019.

[11] R. Yousef and M. Jazzar, "Measuring the effectiveness of user and entity behavior analytics for the prevention of insider threats," *J. Xi'an Univ. Arch. & Technol*, vol. 8, pp. 175-181, 2021.

[12] M. Shashanka, M.-Y. Shen, and J. Wang, "User and entity behavior analytics for enterprise security," in *2016 IEEE International Conference on Big Data (Big Data)*, IEEE, 2016.

[13] J. Liu, et al., "Mueba: A multi-model system for insider threat detection," in *International Conference on Machine Learning for Cyber Security*, Cham: Springer Nature Switzerland, 2022.

[14] P. A. Legg, et al., "Automated insider threat detection system using user and role-based profile assessment," *IEEE Systems Journal*, vol. 11, no. 2, pp. 503-512, 2015.

[15] T. Rashid, I. Agrafiotis, and J. R. C. Nurse, "A new take on detecting insider threats: exploring the use of hidden markov models," in *Proceedings of the 8th ACM CCS International workshop on managing insider security threats*, 2016.

[16] P. A. Legg, et al., "Caught in the act of an insider attack: detection and assessment of insider threat," in *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*, IEEE, 2015.

[17] S. Sharma and G. Makkar, "Scoring Algorithm Identifying Anomalous Behavior in Enterprise Network," in *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2020, Volume 2*, Springer Singapore, 2021.

[18] F. Beer, "A Hybrid Flow-based Intrusion Detection System Incorporating Uncertainty," 2022. Available at: https://kobra.uni-kassel.de/handle/123456789/14247

[19] Widmer, G., and Kubat, M., "Effective learning in dynamic environments by explicit context tracking ", in *Proceedings of the European Conference on Machine Learning (ECML)*, pp. 227–243, 1993.

[20] J. Peng, K. R. Choo, and H. Ashman, "User profiling in intrusion detection: A review," *Journal of Network and Computer Applications*, vol. 72, pp. 14-27, 2016.

[21] S. Khaliq, Z. U. A. Tariq, and A. Masood, "Role of user and entity behavior analytics in detecting insider attacks," in *2020 International Conference on Cyber Warfare and Security (ICCWS)*, IEEE, 2020.

[22] M. Landauer, et al., "A User and Entity Behavior Analytics Log Data Set for Anomaly Detection in Cloud Computing," in *2022 IEEE International Conference on Big Data (Big Data)*, IEEE, 2022.

[23] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla, and F. Herrera, "A unifying view on dataset shift in classification," *Pattern Recognit.*, vol. 45, no. 1, pp. 521-530, 2012.

[24] C. Xie, et al., "A unified framework for long range and cold start forecasting of seasonal profiles in time series," *arXiv preprint arXiv:1710.08473*, 2017.

[25] K. Yemets, "Time series forecasting model for solving cold start problem via temporal fusion transformer," *Computer systems and information technologies*, vol. 1, pp. 57-64, 2024

[26] M. Grbovic, et al., "Cold start approach for data-driven fault detection," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2264-2273, 2013.

[27] F. Pourpanah, et al., "A Review of Generalized Zero-Shot Learning Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4051-4070, Apr. 2023, doi: 10.1109/TPAMI.2022.3191696.

[28] D. L. Costa, et al., "An insider threat indicator ontology," Technical Report CMU/SEI-2016-TR-007, 2016.

[29] P. Hakonen, "Detecting Insider Threats Using User and Entity Behavior Analytics. 2022," Available at: https://www.theseus.fi/handle/10024/786079

[30] A. G. Martín, et al., "An approach to detect user behaviour anomalies within identity federations," *Computers & Security*, vol. 108, pp. 102356, 2021.

[31] M. Z. A. Khan, M. M. Khan, and J. Arshad, "Anomaly detection and enterprise security using user and entity behavior analytics (UEBA)," in *2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS)*, IEEE, 2022.

[32] B. Glasser and B. Lindauer, "Bridging the Gap: A Pragmatic Approach to Generating Insider Threat Data," in *2013 IEEE Security and Privacy Workshops*, San Francisco, CA, USA, 2013, pp. 98-104, doi: 10.1109/SPW.2013.37.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Detección de intrusiones basada en anomalías para alumbrado público inteligente (Resumen)

A. Lara*, R. Estepa Alonso*, A. Estepa Alonso*, J. Diaz-Verdejo†, V. Mayor*
* Dpt. Ingeniería Telemática, Escuela Superior de Ingenieros, Univ. de Sevilla
† Dpt. Teoría de Señal, Telemática y Comunicaciones, CITIC, Univ. de Granada

Los sistemas de alumbrado inteligente (SLS, por sus siglas en inglés) son clave en las ciudades inteligentes por su eficiencia energética y mejoras en la seguridad pública. Sin embargo, estos sistemas también están expuestos a amenazas de ciberseguridad que pueden generar interrupciones del servicio y riesgos de seguridad.

Este artículo presenta un resumen de un sistema de detección de intrusos basado en anomalías (A-IDS) diseñado y validado en un entorno real operativo descrito en [1]. A diferencia de otras propuestas, este A-IDS monitorea los datos de red y telemetría del servidor central de operaciones (O&M), lo que permite detectar una amplia gama de ataques y fallos operativos.

***Palabras Clave*—detección de anomalías, ciberseguridad, smart-city**

I. Introducción

Los Sistemas de Alumbrado Inteligente (*Smart Lighting Systems*, SLS) mejoran la eficiencia energética y la seguridad pública en las ciudades inteligentes. Los SLS están compuestos por luminarias LED, unidades de control de luz (LCU), sensores, actuadores y un sistema central de operación y mantenimiento (O&M). Existen diversas configuraciones, siendo la más común la que emplea LCUs para controlar líneas de luminarias sin comunicación directa entre ellas debido a su bajo costo y simplicidad.

Sin embargo, al estar conectados a redes de comunicación y sensores, también presentan un creciente riesgo de ciberataques. Las amenazas típicas a los SLS incluyen: acceso físico no autorizado, ataques a la red (DoS, escaneo, jamming), manipulación de datos (inyección de comandos, falsificación de datos) y fallos operacionales (lámparas fundidas, robo de energía). Dado que estas amenazas a menudo se reflejan en anomalías en el comportamiento de la red y los datos de telemetría, en este artículo se propone una arquitectura IDS centrada en el servidor O&M para su detección.

II. Arquitectura del sistema IDS propuesto

El sistema se compone de tres módulos principales:

- Módulo de análisis de tráfico de red. Detecta ataques mediante el análisis de series temporales de flujos IPFIX usando técnicas estadísticas (SMA y EWMA).
- Módulo de análisis de datos de telemetría. Implementa tres detectores: Anomalías puntuales entre variables correlacionadas (PCA), desviaciones persistentes en el consumo de potencia (indicador de fallos o robo), y anomalías basadas en la hora del día (modelo horario de consumo normalizado).
- Módulo de correlación. Agrupa eventos y reduce alertas redundantes para facilitar la supervisión.

III. Caso de estudio

El A-IDS ha sido validado con datos reales de flujos IPFIX de 70 LCUs y datos de telemetría de consumo de más de 1500 luminarias (nivel de datos) cedidos por una empresa con despliegues de SLS en 30 países. Los datos de los flujos han sido parametrizados usando ventanas temporales sobre las que se obtienen algunos estadísticos (p.e. número de IPs diferentes), realizándose la detección en base a 7 indicadores de compromiso. Los datos de telemetría han sido analizados en busca de anomalías individuales y grupales según diversas técnicas.

Se simularon 13 ataques (red y datos) para evaluar la detección. El sistema mostró tasas de detección superiores al 90%.

IV. Conclusiones

El sistema propuesto es efectivo, de baja complejidad y aplicable a entornos reales. Su diseño modular permite adaptarlo a otras configuraciones de SLS y amenazas. Además, el uso de datos reales y su publicación permiten avanzar en soluciones reproducibles y adaptables para la seguridad de infraestructuras críticas en ciudades inteligentes.

Referencias

[1] Lara, A., Estepa, A., Estepa, R., Díaz-Verdejo, J. E., Mayor, V. (2024). Anomaly-based Intrusion Detection System for smart lighting. Internet of Things, 28, 101427.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Codificación de acceso aleatorio aplicable a la trazabilidad de documentos

Marcel Fernandez
Network Engineering Department
Universitat Politecnica de Catalunya
Barcelona, Spain

Grigory Kabatiansky
Skoltech
Moscow, Russia

Oscar Esparza
Network Engineering Department
Universitat Politecnica de Catalunya
Barcelona, Spain

El nuevo enfoque para el acceso aleatorio masivo introducido por Yuri Polyanskii en 2017 ha atraído la atención de muchos investigadores. Las publicaciones correspondientes exploran un modelo probabilístico simplificado para evitar considerar casos en los que diferentes usuarios envían la misma información. También han considerado un modelo probabilístico de ruido que no permite combatir ruido malicioso. De hecho, tener en cuenta el caso de ruido malicioso más la posibilidad de que diferentes usuarios transmitan la misma información complica el problema correspondiente de codificación combinatoria y requiere el uso de métodos más complejos de teoría de códigos. En el presente artículo, presentamos construcciones de códigos para este último caso con parámetros asintóticamente óptimos. La motivación principal del trabajo es el uso de dichos códigos para la trazabilidad de documentos.

***Index Terms*—canal de acceso múltiple, acceso aleatorio no coordinado, detección comprimida, recuperación del soporte de vectores, códigos de firma para el canal aditivo con ruido.**

I. Introducción

Considérese una enorme red de sensores que recopila datos de sus usuarios, por ejemplo, telemetría, y su nodo central que debe recopilar y procesar estos datos. Supóngase que, a pesar de la gran cantidad de sensores, solo un número limitado de ellos están activos, digamos como máximo L, en cada intervalo de tiempo dado. Debido al tamaño de la red, no es práctico asignar diferentes códigos para diferentes sensores-usuarios. Entonces, no es posible conocer el origen de ninguna pieza de información recibida, pero sorprendentemente, es posible recopilar toda la información correctamente. Este resultado contraintuitivo fue derivado por Y. Polyanskii [1] y posteriormente se desarrollaron muchas modificaciones diferentes de este enfoque, véase [3], [2] y las referencias allí citadas. Desde [1], todos los resultados se han obtenido bajo la suposición de que todos los datos enviados son *diferentes* con una probabilidad muy cercana a 1. De hecho, esta suposición de diferencia, o suposición-D para abreviar, no siempre es muy realista y, lo que es igual de importante, dependiendo de la suposición-D (o de su ausencia) resultan dos formulaciones diferentes del problema. Describámoslas a continuación.

En ambos modelos hay un conjunto $X \subset \{0,1\}^k$ de N mensajes y todos los usuarios aplican el mismo código lineal binario $[n,k]$ $\mathcal{C} \subset \{0,1\}^n$ para codificar los mensajes en palabras de código, y luego los envían a través de un canal de acceso múltiple. Sea $\psi : \{0,1\}^k \to \mathcal{C}$ la función de codificación correspondiente. Supongamos que L' usuarios activos ($L' \leq L$) quieren enviar sus mensajes $\mathbf{x}_1, \ldots, \mathbf{x}_{L'}$ al nodo central. Entonces seleccionan los vectores de código correspondientes $\mathbf{c}_1 = \psi(\mathbf{x}_1), \ldots, \mathbf{c}_{L'} = \psi(\mathbf{x}_{L'})$ y los envían a través de un canal de acceso múltiple (MAC, por sus siglas en inglés). En este artículo nos restringimos al modelo más popular de canal aditivo binario (BAC, por sus siglas en inglés) en el que la salida del canal es igual a $\mathbf{c}_1 + \ldots + \mathbf{c}_{L'} = \mathbf{s}$, donde $+$ es la suma ordinaria de vectores n-dimensionales con valores reales. Bajo la suposición-D, el código debe elegirse de tal manera que diferentes conjuntos de mensajes produzcan diferentes salidas del MAC, y esta propiedad del código se conoce como *código L-firma*. Consideraremos que la salida del MAC puede estar corrompida por errores, lo que significa para el BAC que su salida es igual a

$$\mathbf{s} = \mathbf{c_1} + \ldots + \mathbf{c_{L'}} + \mathbf{w}, \tag{1}$$

donde $\mathbf{w}$ es el vector de ruido con norma euclídea $||\mathbf{w}|| \leq \delta$. Reescribamos (1) en una forma matricial más conveniente. Denotemos por C la matriz $n \times N$ cuyas columnas son los vectores del código $\mathcal{C}$, correspondientes al conjunto de mensajes X. Y sea $\mathbf{z} \in \{0,1\}^N$ un vector característico binario del conjunto elegido de mensajes transmitidos. Entonces (1) se puede reescribir de la siguiente forma

$$\mathbf{s} = C\mathbf{z}^T + \mathbf{w} \tag{2}$$

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Al definir el vector $\mathbf{z}$ como un vector característico, implícitamente asumimos que no hay mensajes idénticos entre los mensajes transmitidos, es decir, se cumple la suposición-D.

Ahora considere el caso en que no se cumple la suposición-D y hay mensajes idénticos entre los transmitidos. Sea z_c el número de ocurrencias del vector $\mathbf{c}$. Entonces todas las coordenadas del vector $\mathbf{z}$ son enteros no negativos, su suma es igual a L' y el vector de salida $\mathbf{s}$ del BAC viene dado como antes por (2).

Así que necesitamos encontrar una solución *dispersa* de la ecuación (2). Con la suposición-D, solo nos interesan las soluciones binarias, y este caso está bastante bien estudiado. En particular, el problema correspondiente puede considerarse como la recuperación del soporte para el problema de detección comprimida, véase [4], [5], [6], [7].

Hasta donde sabemos, el caso en que el vector $\mathbf{z}$ podría ser un vector disperso arbitrario con coordenadas enteras no negativas con la restricción adicional de que la suma de todas las coordenadas no exceda L no se ha considerado antes. Aquí presentamos algunos resultados preliminares.

II. Relaciones con el problema de detección comprimida

Recordemos que para un vector $\mathbf{z} \in \mathbb{R}^N$ con valores reales, su soporte se define como el conjunto de coordenadas no nulas, es decir, $supp(\mathbf{z}) := \{i :\in [N] : z_i \neq 0\}$, donde $[N] = \{1, \dots, N\}$.

Un vector $\mathbf{z}$ se llama K-disperso si $|supp(\mathbf{z})| \leq K$. El problema de recuperación del soporte es una versión más sencilla del problema de detección comprimida [4], [5]. Concretamente, en lugar de encontrar el vector desconocido K-disperso $\mathbf{z} \in \mathbb{R}^N$ a través de n mediciones lineales distorsionadas por el vector de ruido $\mathbf{w} \in \mathbb{R}^n$, lo que se describe formalmente como encontrar una solución K-dispersa $\mathbf{z}$ de la ecuación (2) *solo se necesita encontrar el soporte de* $\mathbf{z}$. Para nuestros propósitos, esto coincide con el caso en que el vector $\mathbf{z}$ es binario. En la ecuación (2), C es una matriz de medición $n \times N$, cuyas filas son los vectores de medición $\mathbf{c}^{(i)}$, $i = 1, \dots, n$, el vector $\mathbf{z}$ es un vector desconocido K-disperso, el vector $\mathbf{w} = (w_1, \dots, w_n)$ es un vector de ruido de medición y el vector de "síndrome" es $\mathbf{s} = (s_1, \dots, s_n)$, donde $s_i = (\mathbf{c}^{(i)}, \mathbf{z}) + w_i$. El conjunto de columnas $\mathbf{c}_j$, $j = 1, \dots, N$ se denominará código para la recuperación exacta del soporte (código RES, para abreviar).

Definition 1: Un código binario $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_N\} \subset \{0,1\}^n$ se denomina código de recuperación exacta del soporte (K, Δ_ν) ((K, Δ_ν)-código, para abreviar) si para cualesquiera dos subconjuntos distintos $A, B \subset [N]$ tales que $|A|, |B| \leq K$, se cumple la siguiente desigualdad en la métrica l_ν:

$$|| \sum_{j \in A} \mathbf{c}_j - \sum_{j \in B} \mathbf{c}_j ||_\nu \geq \Delta_\nu \tag{3}$$

Consideramos un caso de un vector de ruido malicioso $\mathbf{w}$ con la única restricción de que su norma l_ν está acotada superiormente, y los casos más interesantes son $||\mathbf{w}||_1 \leq \Delta_1$ y $||\mathbf{w}||_2 \leq \Delta_2$.

Además, las matrices de medición propuestas son binarias y pueden construirse explícitamente. También, el número de mediciones m es asintóticamente casi óptimo y garantiza una *recuperación exacta del soporte*. Para K fijo y $N \to \infty$, nuestro algoritmo tiene una complejidad $O(KN)$ con el número de mediciones $m = O(K \log N) + O(\log \log N)$ si el vector de ruido $\mathbf{w}$ cumple que $||\mathbf{w}||_1 = O(n)$ o $||\mathbf{w}||_2 = O(\sqrt{n})$. Tenga en cuenta que si no se tiene en cuenta la complejidad del proceso de mediciones lineales, entonces la complejidad de nuestro algoritmo es $O(KN)$, y para cualquier algoritmo determinista de recuperación exacta del soporte, su complejidad es al menos N.

III. Construcción del Código

Tomamos prestada la construcción de una matriz de medición deseada de [8], que se inspiró en [9]. Considere dos códigos lineales binarios U y V con sus algoritmos de decodificación Φ y Ψ. El código U tiene longitud N, una matriz de verificación de paridad $r \times N$, H, con columnas $\mathbf{h}_1, \dots, \mathbf{h}_N$, y su algoritmo de decodificación Φ corrige K errores, es decir, es capaz de evaluar la solución única de la ecuación de síndrome $H\mathbf{x} = \mathbf{s}$. Y sea el código V un código lineal binario de longitud n con r bits de información con una función de codificación sistemática (por simplicidad) $\Theta : \mathbb{F}_2^r \to V$ y un algoritmo de decodificación Ψ que corrige T errores.

La matriz binaria deseada $n \times N$, C, consta de columnas $\mathbf{c}_1, \dots, \mathbf{c}_n$, donde el vector $\mathbf{c}_j = \Theta(\mathbf{h}_j)$ es el resultado de la codificación de la j-ésima columna $\mathbf{h}_j$ de la matriz H mediante el código V. Se demostró en [8] que el código correspondiente $\mathcal{C}$ es un (K, Δ_2)-código con $\Delta_2 = \sqrt{2T+1}$. Concretamente, se demostró que

$$|| \sum_{j \in A} \mathbf{c}_j - \sum_{j \in B} \mathbf{c}_j ||_2 \geq \sqrt{2T+1}, \tag{4}$$

para cualesquiera dos subconjuntos $A, B \subset [N]$, $|A| \leq T, |B| \leq T, A \neq B$. Volvemos a demostrar (4) a continuación y probamos que también es un (K, Δ_1)-código con $\Delta_1 = 2T + 1$.

En efecto, considere los vectores

$$\mathbf{c}_A := \sum_{j \in A} \mathbf{c}_j = \Theta(\sum_{j \in A} \mathbf{h}_j),$$

$$\mathbf{c}_B := \sum_{j \in B} \mathbf{c}_j = \Theta(\sum_{j \in B} \mathbf{h}_j)$$

y sus "imágenes" binarias $\mathbf{c}_A \mod 2$ y $\mathbf{c}_B \mod 2$. Entonces

$$\mathbf{c}_A \mod 2 = \sum_{j \in A} \Theta(\mathbf{h}_j) \mod 2 = \Theta(\sum_{j \in A} \oplus \mathbf{h}_j)$$

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

ya que Θ es una codificación lineal. Y de manera similar, $\mathbf{c}_B \mod 2 = \Theta(\sum_{j\in B} \oplus \mathbf{h}_j)$. Las palabras binarias $\sum_{j\in A} \oplus \mathbf{h}_j$ y $\sum_{j\in B} \oplus \mathbf{h}_j$ son diferentes porque son sumas diferentes de L o menos columnas de la matriz de verificación de paridad H' del código U que corrige L errores. Por lo tanto, $\mathbf{c}_A \mod 2$ y $\mathbf{c}_B \mod 2$ son palabras de código diferentes del código V y, por lo tanto, difieren al menos en $2T+1$ coordenadas. Así, los vectores con valores enteros $\mathbf{c}_A$ y $\mathbf{c}_B$ difieren al menos en $2T+1$ coordenadas. Por lo tanto, $||\mathbf{c}_A - \mathbf{c}_B||_1 \geq 2T+1$ y $||\mathbf{c}_A - \mathbf{c}_B||_2 \geq \sqrt{2T+1}$.

Finalmente, la ecuación (2) no tiene más de una solución K-dispersa si $||\mathbf{w}||_1 < T + 0.5$ o si $||\mathbf{w}||_2 < 0.5\sqrt{2T+1}$.

Es importante notar que los códigos propuestos pueden construirse explícitamente con complejidad polinomial, pero la codificación aleatoria permite aumentar la tasa del código en un factor de $\log K$, véase [10].

IV. Decodificación y su complejidad

Volvamos a la ecuación (2). Para recuperar el vector K-disperso $\mathbf{z}$, cuyas coordenadas no nulas son iguales a 1, primero realizamos una decodificación dura, es decir, reemplazamos cada coordenada s_i por el entero más cercano $\hat{s}_i$. Obviamente, $\hat{s}_i \neq (C\mathbf{z}^T)_i$ solo si $|w_i| = |s_i - (C\mathbf{z}^T)_i| \geq 1/2$. Por lo tanto, la distancia de Hamming $d_H(\hat{\mathbf{s}}, C\mathbf{z}^T)$ entre dos vectores con valores enteros $\hat{\mathbf{s}}$ y $C\mathbf{z}^T$, es decir, el número de coordenadas en las que estos dos vectores difieren, es a lo sumo $2||\mathbf{w}||_1$. Más aún, $d_H(\hat{\mathbf{s}} \mod 2, C\mathbf{z}^T \mod 2) \leq 2||\mathbf{w}||_1$. Si $2\Delta_1 \leq T$, entonces el algoritmo de decodificación Ψ aplicado al vector $\hat{\mathbf{s}} \mod 2$ produce $C\mathbf{z}^T \mod 2$. A su vez, los primeros r símbolos de $C\mathbf{z}^T \mod 2$ nos dan el síndrome $H'\mathbf{z}^T \mod 2$, ya que la codificación Θ es sistemática. Luego, al aplicar el algoritmo de decodificación del código U, recibimos el vector buscado $\mathbf{s} = H\mathbf{z}^T$ como la solución única de la ecuación $\mathbf{s} = H\mathbf{z}^T$.

Esta consideración se puede repetir para la norma l_2. En efecto,

$$||\mathbf{w}||_2 = \sqrt{\sum_{i=1}^{m} w_i^2} \geq 0.5\sqrt{d_H(\hat{\mathbf{s}}, H\mathbf{x}^T)}$$

ya que $\hat{s}_i \neq (H\mathbf{x}^T)_i$ solo si $w_i = s_i - (H\mathbf{x}^T)_i \geq 1/2$. El resto de la demostración es igual que para la métrica l_1.

Consideremos la propiedad asintótica de la construcción propuesta para los dos regímenes asintóticos más comunes. El primero es cuando K es constante y N crece hasta infinito, lo que corresponde a casos prácticos en los que un vector disperso desconocido es realmente disperso. Como código "interno" U de longitud N, con una matriz de verificación de paridad $r \times N$, H, tomamos códigos BCH o códigos de Goppa clásicos para los cuales $N = 2^{r/K} - 1$ o $N = 2^{r/K}$ respectivamente. Nótese que la decodificación de estos códigos tiene una complejidad $O(K \log N)$, véase [11]. En cuanto al código "externo" V, tomamos de nuevo el código de Goppa o BCH, o si $n >> 1$, entonces un código de una familia de códigos "buenos", es decir, con una tasa de código fija y la relación distancia de código/longitud separada de cero, además de tener una complejidad de implementación polinomial, por ejemplo, la familia de códigos expansores [12], [13].

Notemos en primer lugar que asumimos que el vector $\mathbf{s}$ de (2) ya se conoce y no incluimos la complejidad de las mediciones lineales en la puntuación total. La primera etapa, la decodificación del código "externo" V con el síndrome ya conocido, tiene una complejidad polinomial (e incluso lineal) en m para los códigos BCH o expansores. La segunda etapa, la decodificación del síndrome del código "interno" U para los códigos BCH, tiene una complejidad $O(Kn)$. Dado que $m = O(\log n)$, la complejidad de la primera etapa es despreciable en comparación con la complejidad de la segunda etapa. Así, la complejidad total es $O(Kn)$ para K constante y $n \to \infty$.

El uso de códigos expansores como códigos "interno" y "externo" en el caso en que K crece linealmente con n da una complejidad asintóticamente lineal en n debido a [12], [13]. Notemos también que tomar como código "externo" V un código de una familia de códigos "buenos", es decir, códigos con tasa de código fija R y distancia de código $D \geq 2T+1$, proporciona una recuperación exacta del soporte en el caso de errores cuya longitud en la métrica l_1 crece linealmente con n, o crece como $\Omega(\sqrt{N})$ en la métrica euclídea, lo que no se puede mejorar ya que cualesquiera dos vectores binarios K-dispersos pueden transformarse uno en otro mediante el vector de error correspondiente de longitud $2T+1$ para la métrica l_1 o longitud $\sqrt{2T+1}$ para la métrica l_2.

Consideremos, siguiendo [3], el siguiente ejemplo numérico: $K = 25$, $N = 2^{18}$. Entonces, sea el código "interno" U el código de Goppa irreducible con $r = 25 \times 18 = 450$ bits de verificación de paridad. En cuanto al código "externo" V, tomemos de nuevo un código de Goppa de longitud $m = 1020$, que corrige $T = 57$ errores. Entonces, la matriz de medición resultante 1020×2^{18}, C, es capaz de encontrar exactamente el soporte de un vector binario desconocido de dispersión $K = 25$ con presencia de ruido de norma euclídea como máximo 7.5 o norma l_1 como máximo 57. Desafortunadamente, los algoritmos de decodificación mencionados solo alcanzan una cuarta parte de estas capacidades, en lugar de la mitad.

V. Acceso aleatorio no coordinado

El problema considerado de recuperación del soporte para vectores dispersos es equivalente al problema de los códigos de firma que corrigen errores de canal para el canal aditivo binario y, al mismo tiempo, equivalente (en cierta medida) a los códigos de huella digital multimedia con ruido [14], [15], [16]. En este último escenario está la motivación del presente artículo, en el uso de códigos de huella digital multimedia para la protección de documentos médicos en entornos hospitalarios. El caso sin ruido se conoce como el problema de prueba de grupo de búsqueda de monedas falsificadas mediante una balanza precisa, véase [17], [18], [19], [20]. Al mismo tiempo, este problema es casi equivalente al problema de los códigos superpuestos en el espacio euclídeo, véase

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[21], [22].

Ahora volvamos al acceso aleatorio no coordinado (URA, por sus siglas en inglés) [1].

Definition 2: Un código binario $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_N\} \subset \{0,1\}^n$ se denomina código URA (L, Δ_ν) si para cualesquiera dos vectores con valores enteros diferentes $\mathbf{l} = (l_1, \ldots, l_N)$ y $\mathbf{l}' = (l'_1, \ldots, l'_N)$ tales que $l_j \geq 0$, $l'_j \geq 0$ y $\sum_{j=1}^N l_j \leq L$, $\sum_{j=1}^N l'_j \leq L$ se cumple la siguiente desigualdad

$$||\sum_{j=1}^{N} l_j \mathbf{c}_j - \sum_{j=1}^{N} l'_j \mathbf{c}_j||_\nu \geq \Delta_\nu \tag{5}$$

Nótese de nuevo que si todos $l_j, l'_j \in \{0,1\}$, es decir, si todos los mensajes transmitidos son diferentes, entonces este problema URA *restringido* coincide con el problema de recuperación del soporte considerado anteriormente. Para el problema URA general, utilicemos la misma construcción que se consideró antes, pero no binaria. Concretamente, considere los códigos U y V p-arios, donde p es primo y $p \geq L$. Entonces se puede continuar casi literalmente reemplazando los restos módulo 2 con los restos módulo p. Se podría elegir la misma familia de códigos. A pesar de que los códigos BCH no binarios no son tan buenos como los binarios, es decir, no alcanzan el límite de Hamming asintóticamente para la redundancia del código cuando el número de errores corregidos es fijo y la longitud del código tiende a infinito, la redundancia sigue siendo $\Theta(K \log N)$.

El principal inconveniente de esta construcción es el crecimiento de la longitud (en cualquier métrica l_p) de las columnas de una matriz de verificación de paridad H y, por lo tanto, el crecimiento de la longitud (¡no de la dimensión!) de los vectores de código del código resultante C.

VI. Conclusión

Nuestro estudio está dedicado a una versión combinatoria del novedoso enfoque para el acceso aleatorio no coordinado en presencia de ruido. Existe una literatura bastante extensa sobre este tema, pero los trabajos relacionados se centran principalmente en un modelo de ruido probabilístico (no malicioso). Construimos una nueva familia de códigos que permiten encontrar todos los mensajes transmitidos para el modelo de canal aditivo binario con la presencia de ruido malicioso. Notemos que la cuestión del desarrollo de algoritmos de decodificación adecuados sigue abierta.

VII. Agradecimientos

El trabajo de Marcel Fernández está financiado por the Recovery, Transformation, and Resilience Plan, European Union, under INCIBE Cybersecurity Chair CARISMATICA.

References

[1] Yuri Polyanskiy, “A perspective on massive random-access,” in *Proc. Int. Symp. on Information Theory*,2017,pp.22523-2527.

[2] V. K. Amalladinne, A. K. Pradhan, C. Rush, J. -F. Chamberland and K. R. Narayanan, ”Unsourced Random Access With Coded Compressed Sensing: Integrating AMP and Belief Propagation,” in IEEE Transactions on Information Theory, vol. 68, no. 4, pp. 2384-2409, April 2022

[3] M. Gkagkos, A. K. Pradhan, V. Amalladinne, K. Narayanan, J. -F. Chamberland and C. N. Georghiades, ”Approximate Support Recovery using Codes for Unsourced Multiple Access,” 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 2948-2953

[4] Donoho DL (2006). Compressed sensing. IEEE Transaction on Information Theory 52(4): 1289–1306.

[5] E. J. Candes, J.K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207-1223, 2006.

[6] J. Wen, Z. Zhou, J. Wang, X. Tang and Q. Mo, ”A Sharp Condition for Exact Support Recovery With Orthogonal Matching Pursuit,” in IEEE Transactions on Signal Processing, 2017, vol. 65, no. 6, pp. 1370-1382.

[7] M. Mehrabi and A. Tchamkerten, ”Error-Correction for Sparse Support Recovery Algorithms,” 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 1754-1759

[8] E. E. Egorova, M.Fernandes, G. A. Kabatiansky, Y.Miao, “Existence and construction of complete traceability multimedia fingerprinting codes resistant to averaging attack and adversarial noise”, *Probl. Inf. Transm.*, vol. 56, no. 4, pp. 388-398, 2020.

[9] Ericson T. and Levenshtein V.I. “Superimposed codes in the Hamming space”, *IEEE Trans. Inform. Theory*. 1994. V. 40. № 6. P. 1882-1893.

[10] I. Vorobyev, “Complete traceability multimedia fingerprinting codes resistant to averaging attack and adversarial noise with optimal rate”, *Des. Codes Cryptogr.*, V. 85, 2022.

[11] R.E. Blahut Theory and Practice of Error Control Codes. *Addison-Wesley*, 1984.

[12] Sipser, M.; Spielman, D.A. (1996). ”Expander codes”. IEEE Transactions on Information Theory. 42 (6): 1710–1722.

[13] Spielman, D. (1996). ”Linear-time encodable and decodable error-correcting codes”. IEEE Transactions on Information Theory. 42 (6): 1723–31.

[14] M. Cheng and Y. Miao, “On anti-collusion codes and detection algorithms for multimedia fingerprinting,” *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4843-4851, 2011.

[15] E. Egorova, M. Fernandez, G. Kabatiansky, and M. H. Lee, “Signature codes for weighted noisy adder channel, multimedia fingerprinting and compressed sensing,” *Designs, Codes and Cryptography*, vol. 87, pp. 455-462, 2019.

[16] Fan J., Gu Y., Hachimori M., and Miao Y. Signature Codes for Weighted Binary Adder Channel and Multimedia Fingerprinting. *IEEE Trans. Inform. Theory*. 2021. V. 67. No 1. P. 200-216.

[17] Erdos and A. Renyi. On two problems of information theory. Publ. Math. Inst. Hung. Acad. Sci. V. 8, 241254, 1963.

[18] *Djackov A. G.* On a search model of false coins // Topics in Information Theory (Colloquia Mathematica Societatis Janos Bolyai 16 Keszthely Hungary). 1975. P. 163-170.

[19] N. H. Bshouty and H. Mazzawi, “On parity check $(0,1)$-matrix over $\mathbb{Z}_p$,” *Proc. 22nd Annu. ACM-SIAM Symp. on Discrete Algorithms (SODA'11)*, pp. 1383-1394, San Francisco, CA, 2011.

[20] N. H. Bshouty and H. Mazzawi, “Algorithms for the Coin Weighing Problems with the Presence of Noise”, Electronic Colloquium on Computational Complexity, Report No. 124, 2011.

[21] *Ericson T., Gyorfi L.* Superimposed Codes in $\mathbb{R}^n$ // IEEE Trans. Inform. Theory. 1988. V. 34. № 4. P. 877-880.

[22] *Füredi Z., Ruszinko M.* An improved upper bound of the rate of Euclidean superimposed codes // IEEE Trans. Inform. Theory. 1999. V. 45. № 2. P. 799-802.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Aula invertida en docencia remota a nivel de máster: experiencia y resultados obtenidos

Ramon Sanchez-Iborra
Departamento de Ingeniería de la Información y las Comunicaciones,
Universidad de Murcia
30.100
ramonsanchez@um.es

La docencia en línea sigue penetrando en la educación superior, cobrando especial relevancia a nivel de máster, dado el perfil del alumnado que cursa estas titulaciones. Para poder superar la barrera física que presenta esta modalidad de aprendizaje no presencial, se hace necesario adoptar enfoques pedagógicos innovadores, siendo el aula invertida uno de los más empleados durante los últimos tiempos. Este artículo presenta la experiencia y resultados obtenidos al implantar este sistema de enseñanza-aprendizaje en una asignatura de máster centrada en la tecnología *blockchain*. Los resultados obtenidos, tanto a nivel de satisfacción de los estudiantes como académicos, validan la propuesta, la cual facilita el aprendizaje de una tecnología novedosa de forma eminentemente práctica y permitiendo a los estudiantes el poder trabajar los contenidos de la asignatura a su propio ritmo.

***Palabras Clave*—Aula invertida, innovación docente, docencia *on-line*, máster.**

I. INTRODUCCIÓN

El nivel de adopción de la enseñanza remota u *on-line* es cada vez mayor en la educación superior, dada la flexibilidad y ventajas que ésta aporta a los estudiantes. Sin embargo, aunque este tipo de docencia elimina barreras geográficas, plantea una eminente barrera física entre docente y alumnado, limitando la interacción interpersonal entre los participantes en estas clases. Esto puede propiciar que los estudiantes pierdan la concentración o motivación durante las sesiones síncronas, tanto de teoría como de prácticas [1]. Es por ello, que desde hace unos años se está fomentando el uso de metodologías docentes activas que promuevan una mayor involucración de los estudiantes en su propio proceso de aprendizaje.

Una de las metodologías más adoptadas para fomentar una mayor participación del alumnado, tanto en las sesiones síncronas como en su trabajo autónomo, es el aula invertida [2]. Este modelo pedagógico transforma la dinámica tradicional del aprendizaje, basada en la lección magistral, al desplazar el trabajo teórico fuera del aula, permitiendo que el tiempo en clase se enfoque en actividades prácticas y colaborativas. Este enfoque fomenta la motivación de los estudiantes al permitirles trabajar a su propio ritmo y que se preparen mejor para las sesiones síncronas con el docente. El aula invertida ha demostrado ser eficaz para mejorar la comprensión y retención de contenidos, lo que se traduce en un mejor rendimiento académico. Al dedicar tiempo para el aprendizaje activo en clase, los estudiantes pueden profundizar en conceptos clave y desarrollar competencias prácticas guiados por el profesor. Esto resulta especialmente valioso en docencia remota donde las distracciones externas son más comunes que en un aula [3].

Por todo lo anterior, este trabajo presenta la experiencia llevada a cabo en una asignatura de nivel de máster impartida de forma remota y en inglés, la cual ha sido adaptada para ser impartida siguiendo la metodología de aula invertida. En concreto, la asignatura, denominada "Seguridad en Servicios Distribuidos" y enmarcada dentro del programa del Máster en Ciberseguridad de la Universidad de Murcia, aborda los aspectos fundamentales de la tecnología *blockchain*, así como sus implicaciones en el mundo de la ciberseguridad. Este trabajo presenta la propuesta metodológica implementada, así como los resultados de esta experiencia. En cuanto a la metodología docente, se hace uso de distintas herramientas digitales tanto para la impartición de la parte de teoría como de prácticas. Atendiendo a la parte de teoría, se emplea material audiovisual producido por el profesorado de la asignatura, que los estudiantes deben trabajar antes de cada sesión. Durante las sesiones síncronas de teoría, se emplea una herramienta de cuestionarios en línea para mantener la motivación y concentración de los estudiantes, además de permitir su autoevaluación en cuanto al grado de comprensión de los contenidos tratados en cada sesión. Para la parte de prácticas, se hace uso de un simulador de tecnología *blockchain*, accesible de forma libre a través de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

un servicio web, lo que permite a los estudiantes trabajar con él tanto de manera síncrona con el profesor, como asíncrona desde cualquier tipo de dispositivo conectado a Internet. Los resultados de esta experiencia se obtienen en base a una encuesta de satisfacción pasada a los estudiantes, así como al rendimiento académico alcanzado por éstos. En ambos casos, se obtienen resultados muy positivos.

El resto del documento se organiza de la siguiente forma. Los trabajo previos relacionados se discuten en la Sección II. El contexto de la asignatura foco de este estudio se presenta en la Sección III. La metodología propuesta, basada en aula invertida, se detalla en la Sección IV. Los resultados de evaluación de dicha metodología, tanto a nivel de satisfacción del alumnado como académicos, se presentan en la Sección V. Finalmente, la Sección VI cierra el trabajo, presentado las conclusiones de éste y líneas de trabajo futuras.

II. TRABAJOS RELACIONADOS

La metodología de aula invertida (*flipped classroom*) representa un cambio de paradigma en el proceso en enseñanza-aprendizaje que ha recibido mucha atención durante los últimos años. En vez de recibir los contenidos a través del profesor en el aula y realizar tareas fuera de estas sesiones, los estudiantes acceden al contenido teórico fuera del aula, a través de vídeos, lecturas u otros recursos, y utilizan el tiempo en clase para actividades prácticas, debates y resolución de problemas con el apoyo del docente. El trabajo seminal que propone esta metodología destaca su potencial para personalizar el aprendizaje y aumentar la interacción en el aula [4]. Los autores enfatizan la importancia de crear vídeos instructivos que sean claros y atractivos, a la vez de diseñar actividades en clase que fomenten la colaboración y el pensamiento crítico.

Posteriormente, son muchos los trabajos que analizan el impacto de esta metodología desde distintas perspectivas [2]. A continuación, se pone especial énfasis en aquellos trabajos que analizan la aplicación del sistema de aula invertida en entornos *on-line*. Cabe destacar el meta-análisis llevado a cabo por los autores de [5]. Este artículo presenta un recorrido exhaustivo sobre la efectividad de diferentes enfoques educativos: aprendizaje *on-line*, aprendizaje combinado, aula invertida y enseñanza tradicional, en dos poblaciones específicas: estudiantes de magisterio y docentes en activo. Las conclusiones de este estudio destacan que los enfoques de aprendizaje combinado y el aula invertida demostraron ser significativamente superiores a la enseñanza basada exclusivamente en la impartición tradicional. Esto se extrae de las mejoras detectadas en el rendimiento académico y la mejora en el desempeño general tanto de los docentes como de los estudiantes. Además, de este trabajo se desprende que el aprendizaje *on-line* es al menos tan efectivo como la enseñanza presencial tradicional, lo que refuerza su viabilidad como alternativa educativa.

El trabajo presentado en [3] examina la eficacia del aula invertida en un contexto completamente remoto durante la pandemia de COVID-19, comparando los resultados de aprendizaje con métodos de enseñanza remota tradicionales. Los autores muestran que los estudiantes en el grupo de aula invertida obtuvieron mejores resultados académicos y mostraron mayor satisfacción que los del grupo de control, observando un aumento significativo en la participación y la interacción en el entorno virtual cuando se implementó el modelo invertido. Además, los estudiantes valoraron especialmente la flexibilidad en el acceso a los materiales y la posibilidad de revisar repetidamente las explicaciones, aunque el trabajo también identifican desafíos importantes relacionados con la conectividad, la autodisciplina y la gestión del tiempo por parte de los estudiantes. Por su parte, los autores de [6] concluyen que la integración del modelo de aula invertida en educación superior *on-line* en el campo de la salud pública resulta ser una innovación efectiva, ya que los estudiantes analizados lograron resultados similares en los exámenes en comparación con el modelo tradicional, pero reportaron experiencias de enseñanza-aprendizaje más positivas y una mayor percepción de aprendizaje.

El trabajo presentado en [7], destaca la necesidad de integrar tecnologías interactivas y enfoques basados en problemas para maximizar los beneficios del aula invertida. Sin embargo, este estudio también subraya que el éxito del aula invertida depende no solo de la tecnología utilizada, sino de cómo se diseñan las actividades para involucrar a los estudiantes activamente. Finalmente, los autores de [8] analizan la efectividad de la metodología de aula invertida en cursos de corta duración. Los resultados presentados ponen de manifiesto que el uso del aula invertida en este tipo de cursos permiten personalizar el aprendizaje gracias a su enfoque adaptado a grupos pequeños y homogéneos. Esto facilita una interacción más cercana entre alumnado y profesores, aumentando la eficacia del curso y el compromiso de los estudiantes. Además, se enfatiza que el uso de vídeos y materiales para estudio autónomo antes de las clases fomenta una mejor comprensión de los conceptos clave, los cuales son reforzados posteriormente de forma síncrona con el profesor. En este trabajo, la mayoría de los estudiantes reportan altos niveles de satisfacción con este enfoque pedagógico, destacando su utilidad para entender conceptos complejos y su capacidad para fomentar la autonomía y compromiso por el aprendizaje.

A la vista de las conclusiones positivas que se pueden extraer de todos los trabajos anteriores, se opta por hacer uso de la metodología de aula invertida para dar respuesta a las necesidades particulares de la asignatura que se presenta en la siguiente sección.

III. CONTEXTO: "SEGURIDAD EN SERVICIOS DISTRIBUIDOS"

La asignatura donde se ha aplicado la metodología presentada en este trabajo se denomina "Seguridad en Servicios Distribuidos" y se enmarca en el Máster en Ciberseguridad de la Universidad de Murcia. Éste se ha planteado desde su inicio como una titulación que se imparte completamente *on-line* y en inglés.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

La asignatura foco de este estudio se sitúa en la segunda mitad del segundo cuatrimestre del máster, presentando 3 créditos ECTS. Es una asignatura optativa que permite a los estudiantes adoptar uno de los perfiles ofrecidos por la titulación: Ingeniero de Seguridad (frente al otro perfil que es el de Analista de Seguridad). De este modo, se entiende que la asignatura debe tener una visión y orientación eminentemente técnica y práctica, y no tanto de gestión o manejo de políticas de seguridad que es el foco principal del otro perfil. En este contexto, se entiende que la enseñanza de la tecnología *blockchain* se ha convertido en un pilar fundamental para comprender sistemas descentralizados seguros actuales, con amplia adopción en distintos ámbitos [9], por lo que se toma como el bloque principal de contenidos para esta asignatura.

Dada la complejidad del tema a abordar, el cual normalmente no es conocido por los estudiantes antes de acceder a esta asignatura, así como el tiempo reducido para su impartición, se decide adoptar una metodología basada en aula invertida para que los estudiantes puedan sacar el mayor provecho de ella. Además, en la memoria del título[1] se sugiere el uso de metodologías de enseñanza-aprendizaje activas para esta asignatura, atendiendo a las actividades formativas asignadas a ésta, las cuales se presentan en la Tabla I. Como se puede observar, se establecen sesiones de teoría y prácticas tanto síncronas como asíncronas, por lo que la metodología de aula invertida se adapta perfectamente a las actividades formativas determinadas para la asignatura. Además, dado el reducido tiempo de sincronía entre profesor y estudiantes (4,5 horas para teoría y otras tantas para prácticas), se hace necesario que los estudiantes hagan uso de las sesiones asíncronas planificadas, trabajando los materiales que se ponen a su disposición antes de las sesiones síncronas con el profesor.

Los contenidos tratados en la asignatura son:

1) *Blockchain*: Introducción y generalidades.
2) Algoritmos de consenso y *forks*.
3) Contratos inteligentes.
4) Ataques a *blockchain*.

Estos contenidos se trabajan de forma síncrona con el docente en cuatro sesiones teóricas de una hora, más otras cuatro sesiones prácticas de la misma duración. Además, también se dedica otra sesión para que los estudiantes reciban un seminario de un experto en la materia ajeno a la universidad y realizar casos prácticos acerca de dicho seminario. En la siguiente sección se presenta la metodología docente adoptada para las actividades formativas de la asignatura, explicando en detalle el material preparado y las actividades planteadas a los estudiantes.

IV. METODOLOGÍA DOCENTE

La metodología adoptada para las partes de teoría y prácticas, aunque en ambos casos sigue la estrategia de aula invertida, difiere en las herramientas empleadas. Así, a continuación se presentan cada una de estas partes por separado.

[1]https://www.um.es/documents/1083928/40394712/MemoriaMasterCiberseguridad.pdf/1845c241-70d5-fca5-fd1a-46dffa41bb07?t=1686816084772

Tabla I
ACTIVIDADES FORMATIVAS DE LA ASIGNATURA

Id	Denominación	Horas	% Presencialidad
AF1	Sesiones síncronas de teoría *on-line*	4,5	0
AF3	Sesiones asíncronas de teoría	4,5	0
AF2	Sesiones síncronas de laboratorio *on-line*	4,5	0
AF4	Sesiones asíncronas de laboratorio	4,5	0
AF6	Tutorías	2	0
AF5	Casos prácticos	1	0
AF8	Seminario	1	0
AF9	Trabajo autónomo	51	0
AF10	Pruebas de evaluación	2	0

A. Teoría

Dada la novedad y complejidad del tema tratado, así como el escaso tiempo disponible para su impartición de forma síncrona (ver Tabla I), abordar la asignatura con una metodología basada en clase magistral se hace inviable. De este modo, se realiza una planificación basada en las cuatro sesiones síncronas de impartición de teoría disponibles. Los estudiantes tienen acceso a dicha planificación semanal través del sistema de Aula Virtual de la Universidad de Murcia (Fig. 1). En cada una de estas sesiones se aborda una de las unidades de contenidos indicadas anteriormente. Así, para cada una de estas unidades se ha grabado un vídeo donde el profesor explica el tema en cuestión haciendo uso de unas diapositivas que también se ponen a disposición de los estudiantes, días antes de la sesión síncrona donde se trabajaran dichos contenidos junto con el docente. Para que los estudiantes puedan saber qué contenidos se van a tratar en cada sesión síncrona, durante la semana anterior se les indica el material a trabajar a través del mismo sistema de planificación semanal mencionado anteriormente (Fig. 2). De este modo, se espera que los estudiantes hayan visto de antemano el vídeo correspondiente a los contenidos a trabajar en cada sesión síncrona. Además, pueden hacer uso de las diapositivas del tema para anotar dudas o cuestiones a comentar con el profesor durante la clase. Los vídeos grabados en ningún caso superan los 20 minutos de duración, para que el trabajo previo del estudiante (visionado de vídeo, búsqueda proactiva de información y anotación de comentarios y/o dudas) no supere el tiempo de trabajo asíncrono asociado a cada sesión síncrona (una hora, tal y como se establece en la Tabla I).

Durante la clase con el profesor, éste comienza haciendo una presentación de los contenidos tratados hasta el momento en sesiones anteriores (en el caso de la primera sesión, se realiza una presentación general de la asignatura). Posteriormente, haciendo uso de las mismas diapositivas suministradas de antemano a los estudiantes, el profesor realiza un recorrido sobre los contenidos a tratar en la sesión, permitiendo a los estudiantes exponer sus dudas o comentarios. En ciertas situaciones, se permite que los estudiantes se respondan entre ellos, siempre con la moderación, y posible corrección, por parte del docente. Además, se hace uso de una herramienta de cuestionarios *on-line*, donde el profesor plantea una serie de preguntas sobre puntos clave del tema, de modo que los estudiantes

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

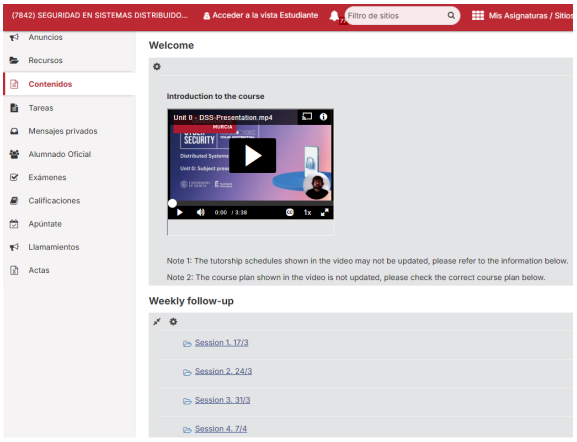


Fig. 1. Seguimiento semanal de la asignatura.

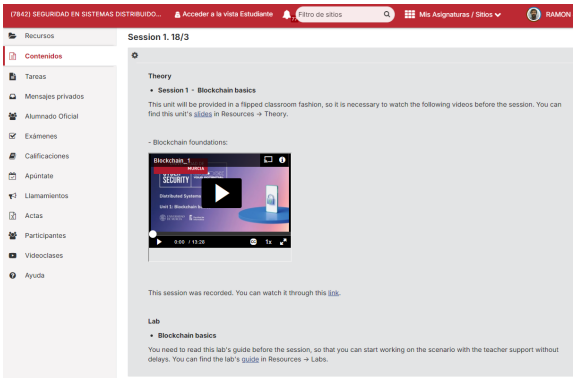


Fig. 2. Instrucciones del trabajo previo a realizar por los estudiantes para preparar una sesión síncrona.

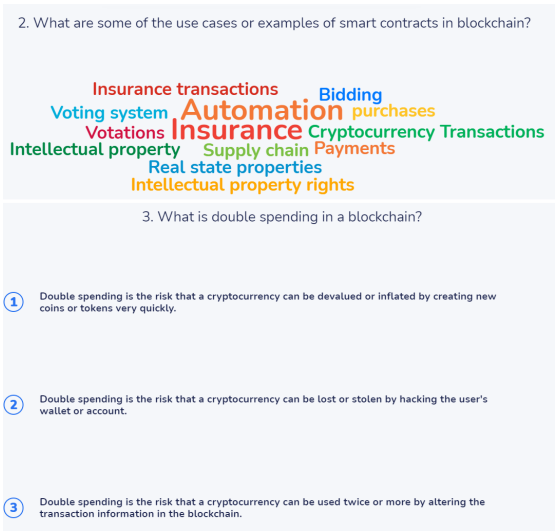


Fig. 3. Ejemplos de cuestiones planteadas durante las sesiones síncronas de teoría.

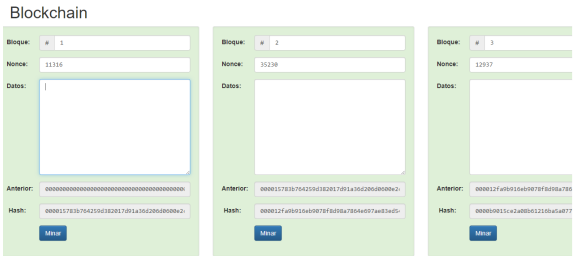


Fig. 4. Ejercicio para manipular bloques en una *blockchain*.

puedan comprobar si han asimilado correctamente dichos contenidos. Además, estas preguntas fomentan el debate porque se utilizan distintos tipos de cuestiones: nube de respuestas, preguntas tipo test, ordenación o alineación de conceptos, etc. (Fig. 3). Esto permite que la clase sea dinámica y que los estudiantes puedan plantear sus dudas abiertamente.

Finalmente, también se plantean ejercicios eminentemente prácticos para que los estudiantes puedan aplicar lo trabajado durante la sesión en un caso de uso próximo a la realidad. Estos ejercicios pueden conllevar búsquedas en la web o algún ejercicio técnico. En la Fig. 4, se muestra un ejercicio planteado durante la primera sesión síncrona de teoría en el que los estudiantes exploran y manipulan el contenido de los bloques en una *blockchain*, para comprender como impacta la alteración de este contenido en toda la cadena.

B. *Prácticas*

Tal y como se comentó anteriormente, para abordar la parte práctica de la asignatura también se adopta la metodología de aula invertida. En este caso, se ha desarrollado un guión de prácticas que abarca las cuatro sesiones síncronas de laboratorio. Tal y como se ve en la Fig. 2, en las instrucciones proporcionadas semanalmente a los estudiantes, también se les indica que deben trabajar de antemano la parte práctica para aprovechar la sesión síncrona de la mejor manera posible. Para este trabajo previo, disponen de cuatro sesiones asíncronas de prácticas de una hora cada una (ver Tabla I). En el guión de laboratorio, se les presenta a los estudiantes la herramienta a emplear, se les guía en su utilización a través de un escenario de ejemplo y, finalmente, se les plantean dos ejercicios que deben resolver durante las sesiones síncronas con el apoyo del profesor. Además, en la primera sesión síncrona, el docente también presenta un recorrido sobre la herramienta para que los estudiantes no tengan dudas a la hora de comenzar a trabajar con ella.

En cuanto a la herramienta empleada, se hace uso de un simulador de *blockchain* de acceso libre a través de Internet [10]. Esta herramienta permite desplegar cualquier topología de nodos formando una red de *blockchain* haciendo uso del algoritmo de consenso *Proof-of-Work* (el empleado en *Bitcoin*). Una vez que se despliega una red de nodos (mineros), se pueden configurar distintos parámetros de éstos, como su potencia de cómputo, ancho de banda y retardo de sus enlaces, etc. Además, cuando se lanza el proceso de minado se puede inspeccionar con detalle tanto el estado general del sistema como el de cada nodo. Así, es una herramienta muy versátil que permite explorar a fondo el funcionamiento de esta tecnología (Fig. 5).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

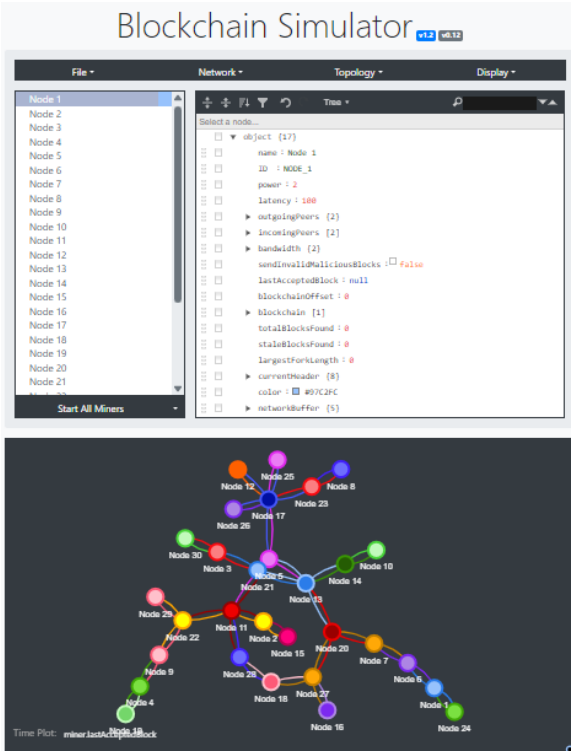


Fig. 5. Simulador de *blockchain* empleado en prácticas.

Tras una primera sesión que introduce a los estudiantes en el funcionamiento de la herramienta, se les plantea una serie de ciberataques que deben implementar, para comprender las consecuencias que éstos tienen sobre los nodos atacados. De los distintos ataques que se analizan en la parte téorica, se selecciona un conjunto representativo de estos, como el ataque de Eclipse o el del 51% [11], para que sean implementados por los estudiantes. En base a estos escenarios propuestos, se les plantean una serie de preguntas que, para ser contestadas, requieren de una profunda comprensión del modo de funcionamiento de la tecnología *blockchain* y las consecuencias de cada ataque implementado. Tal y como se comenta en la sección siguiente, los estudiantes deben realizar un informe de prácticas explicando las configuraciones desarrolladas y los resultados obtenidos respondiendo a una serie de preguntas que les van guiando en este proceso. Además, también deben "defender" dicho informe en una entrevista personal con el profesor.

C. Evaluación

En cuanto a la evaluación y calificación de la asignatura, se consideran tres ítems con distintos pesos.

- Cuestionario teórico (35%): después de finalizar las cuatro sesiones teóricas, el alumnado debe responder de forma autónoma a una serie de preguntas de desarrollo planteadas. Para resolver este cuestionario se les da un tiempo máximo de una semana y cuentan con el apoyo del profesor (de forma asíncrona) en caso de ser necesario. En la calificación de cada ejercicio planteado se valora la rigurosidad en la respuesta y la creatividad a la hora de idear ciertos escenarios planteados.
- Informe de prácticas (35%): este documento debe reportar todas las configuraciones realizadas en el simulador para resolver las cuestiones planteadas en el guión de prácticas. Se debe demostrar que la implementación realizada permite simular el comportamiento de los ataques propuestos, así como discutir el funcionamiento de los mismos. Para realizar esta tarea, el alumnado dispone de un plazo mayor que para resolver el cuestionario teórico (alrededor de dos semanas).
- Entrevista personal (30%): una vez entregado el informe de prácticas, los estudiantes deben realizar una entrevista individual con el profesor para exponer dicho informe y resolver las cuestiones de índole teórico-prácticas que se les pueda plantear acerca del temario de la asignatura. Es obligatorio superar esta entrevista para poder aprobar la asignatura.

En el caso de no aprobar la asignatura a través de los ítems anteriores, los estudiantes pueden superar las partes suspensas mediante un examen oral final que abarca todos los contenidos teórico-prácticos de la asignatura. Dependiendo de la situación particular de cada estudiante, este examen se enfoca especialmente en evaluar las partes que no haya superado.

V. RESULTADOS

La evaluación de la metodología propuesta se realiza desde dos enfoques bien diferenciados: satisfacción del alumnado y resultados académicos alcanzados.

A. Satisfacción de los estudiantes

Para conocer el grado de satisfacción del alumnado con la metodología expuesta anteriormente, se les pasó una encuesta al finalizar la impartición de la asignatura. La encuesta constó de diez afirmaciones que los estudiantes debían valorar en una escala Likert, yendo desde 1 ("totalmente en desacuerdo") hasta 5 ("totalmente de acuerdo"). En concreto, las afirmaciones que formaron la encuesta son las siguientes:

1) He visto todos los vídeos y he leído el manual de laboratorio antes de las sesiones.
2) Los vídeos previos a las clases han sido claros y fáciles de entender.
3) Ver los vídeos antes de la clase me ha ayudado a comprender mejor los conceptos teóricos.
4) El formato de aula invertida ha incrementado mi interés por la temática del curso.
5) Me he sentido seguro al hacer preguntas en clase después de ver los vídeos.
6) El manual de laboratorio ha estado bien estructurado y ha sido fácil de seguir.
7) Utilizar el simulador de *blockchain on-line* me ha ayudado a entender cómo funciona esta tecnología en la práctica.
8) La combinación de teoría y práctica ha mejorado mi comprensión general de la tecnología *blockchain*.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

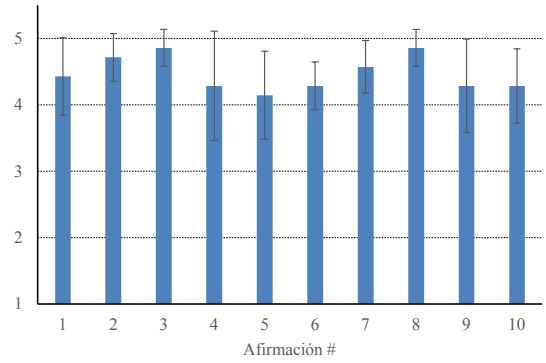


Fig. 6. Resultados de la encuesta pasada a los estudiantes.

9) Me siento más seguro de mi capacidad para explicar conceptos de *blockchain* después de finalizar esta asignatura.
10) En general, estoy satisfecho con el enfoque de aula invertida y los materiales utilizados en la asignatura.

Este cuestionario fue respondido por el 87,5% de los estudiantes matriculados en la asignatura. Los resultados de esta encuesta se presentan en la Fig. 6, que muestra la media y los intervalos de confianza (95%) de las valoraciones recibidas en cada una de las diez afirmaciones indicadas anteriormente. En general, todas las afirmaciones reciben valoraciones altas, en todo los casos con un valor por encima del 4 sobre 5. Atendiendo al grado de compromiso adquirido por los estudiantes con la asignatura y su temática (afirmaciones 1 y 4), se obtienen valoraciones medias de 4,4 y 4,3, respectivamente, lo que permite deducir que los estudiantes han realizado el trabajo previo a cada sesión, motivándose por el aprendizaje de los temas tratados en la asignatura. En cuanto a la calidad de los materiales preparados por el equipo docente, las respuestas recibidas en las afirmaciones 2 (vídeos de teoría), 6 (manual de prácticas) refrendan la buena acogida de éstos por parte de los estudiantes, con valoraciones de 4,7 y 4,3, respectivamente. Además, la utilidad de los vídeos preparados para la parte de teoría se hace evidente según las respuesta obtenidas en la respuesta 3 (4,9), orientada a entender si el visionado de estos vídeos permitía al alumnado comprender los conceptos teóricos. En cuanto a la parte de prácticas, también ha recibido una valoración muy satisfactoria, tanto a nivel de la herramienta empleada (afirmación 7: 4,6), como su utilidad en combinación con la parte teórica para entender el funcionamiento de la tecnología *blockchain* (afirmación 8: 4,9). Estos resultados vienen a confirmar que la elección de herramientas y ejercicios prácticos adecuados en este tipo de entornos de enseñanza-aprendizaje es fundamental para que los estudiantes puedan adquirir los conocimientos, competencias y destrezas trabajados en la asignatura. En cuanto a la seguridad que aporta a los estudiantes la metodología empleada, tanto para participar durante las sesiones síncronas (afirmación 5: 4,1), como para ser capaces de explicar a otros la temática tratada (afirmación 9: 4,3), también obtienen valoraciones notables en ambos casos. Finalmente, la valoración total que recibe la metodología empleada (afirmación 10) es de un 4,3 sobre 5, evidenciando una buena aceptación por parte del alumnado del enfoque adoptado para la impartición de la asignatura.

Tabla II
CALIFICACIONES OBTENIDAS EN CONVOCATORIA ORDINARIA

Item	Media	Int. confianza (95%)
Cuestionario teórico	9,16	1,28
Informe de prácticas	9,16	0,54
Entrevista personal	9,6	0,35
Examen final	-	-
Calificación total	9,3	0,53

A la vista de los resultados obtenidos, los cuales están bastante en sintonía con lo reportado en los trabajos previos revisados en la Sección II, se puede concluir que los estudiantes de la asignatura han acogido de forma positiva la metodología de aula invertida y los materiales preparados para su desarrollo. En el siguiente apartado, se analiza si esta buena acogida se ve refrendada por los resultados académicos alcanzados por los estudiantes.

B. Resultados académicos

La Tabla II muestra la calificación media obtenida por los estudiantes en los distintos ítems de evaluación mencionados anteriormente. Cabe mencionar que en la convocatoria ordinaria, el 75% de los estudiantes matriculados siguieron el mecanismo de evaluación continua. La elevada participación en esta modalidad de evaluación sugiere un alto grado de implicación y motivación del alumnado, posiblemente potenciado por la metodología basada en aula invertida. Por su parte, los estudiantes que escogieron el sistema de evaluación final no se presentaron en dicha convocatoria, por lo que no se tienen datos sobre su rendimiento, ya que este trabajo se ha desarrollado antes de producirse la convocatoria extraordinaria de la asignatura.

Atendiendo a los resultados obtenidos por los estudiantes que han seguido el proceso continuo de evaluación, se observa un rendimiento muy alto en todos los ítems considerados. En todos los casos, la calificación media supera el 9, incluyendo la calificación total de la asignatura. Destaca el desempeño en la entrevista personal, donde los estudiantes deben responder en tiempo real y sin documentación a preguntas técnicas realizadas por el profesor acerca de conceptos tratados tanto en sesiones teóricas como de prácticas. La calificación media obtenida en este ítem es incluso superior a las del cuestionario e informe de prácticas, realizados de forma asíncrona por los estudiantes.

A la vista de estos resultados, se considera que la metodología basada en aula invertida ha permitido una comprensión sólida de los contenidos y competencias trabajados en la asignatura. Por un lado, el alto rendimiento en el cuestionario teórico, así como en el informe de prácticas, evidencia que los estudiantes han sabido organizar y gestionar su tiempo de trabajo autónomo de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

forma efectiva, aprovechando los recursos y herramientas proporcionados. Además, los resultados satisfactorios en la entrevista personal ponen de manifiesto que los estudiantes son capaces de demostrar no solo conocimientos, sino también la capacidad de aplicar y relacionar conceptos y aplicar destrezas en un contexto eminentemente práctico.

VI. CONCLUSIONES

Este trabajo presenta una experiencia de innovación docente enfocada en la aplicación de la metodología de aula invertida en una asignatura de nivel de máster, impartida completamente *on-line* y en inglés. Para esta iniciativa, se ha preparado material audiovisual y diversos manuales para que los estudiantes puedan trabajar los contenidos de la asignatura antes de las sesiones síncronas con el profesor. Éstas, además de permitir a los estudiantes expresar sus dudas surgidas durante su trabajo previo, se planifican con una orientación eminentemente práctica, con ejercicios planteados a través de herramientas *on-line*, como simuladores, cuestionarios de distinta naturaleza, etc. Los resultados obtenidos, tanto desde una perspectiva de la satisfacción del alumnado, como de los resultados académicos alcanzados, confirman que el planteamiento metodológico ha sido adecuado, dada la buena acogida por parte de los estudiantes, así como por su buen rendimiento académico. Como trabajo futuro, se plantea incorporar nuevas herramientas a la asignatura, para mejorar la experiencia en el proceso de enseñanza-aprendizaje, así como adoptar la metodología propuesta en otras asignaturas de nivel de grado, lo que implica una mayor duración de la asignatura y un número más alto de estudiantes involucrados.

AGRADECIMIENTOS

Este trabajo ha sido apoyado por el proyecto PID2023-148104OB-C43 financiado por MICIU/AEI/10.13039/501100011033 y por FEDER/UE; por la Comisión Europea, a través del proyecto NANCY (101096456); y por el Ministerio de Economía y Transformación Digital, a través del proyecto CERBERUS-MINERVA (TSI-063000-2021-46). El autor quiere también agradecer el apoyo recibido a través del grupo de *Innovación docente para la enseñanza de las TIC y sus fundamentos* y por la convocatoria 2024 de *Proyectos y Acciones de Innovación y Mejora*, ambos de la Universidad de Murcia.

REFERENCIAS

[1] H. Akram and S. Li, "Understanding the Role of Teacher-Student Relationships in Students' Online Learning Engagement: Mediating Role of Academic Motivation," *Perceptual and Motor Skills*, vol. 131, no. 4, pp. 1415–1438, 2024.

[2] G. Akçayır and M. Akçayır, "The flipped classroom: A review of its advantages and challenges," *Computers & Education*, vol. 126, pp. 334–345, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360131518302045

[3] T. Tang, A. M. Abuhmaid, M. Olaimat, D. M. Oudat, M. Aldhaeebi, and E. Bamanger, "Efficiency of flipped classroom with online-based teaching under COVID-19," *Interactive Learning Environments*, vol. 31, no. 2, pp. 1077–1088, feb 2023. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10494820.2020.1817761

[4] J. Bergmann and A. Sams, *Flip your classroom: Reach every student in every class every day*, Washington, D.C., 2012.

[5] R. F. Schmid, E. Borokhovski, R. M. Bernard, D. I. Pickup, and P. C. Abrami, "A meta-analysis of online learning, blended learning, the flipped classroom and classroom instruction for pre-service and in-service teachers," *Computers and Education Open*, vol. 5, p. 100142, dec 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2666557323000204

[6] L. P. Galway, K. K. Corbett, T. K. Takaro, K. Tairyan, and E. Frank, "A novel integration of online and flipped classroom instructional models in public health higher education," *BMC Medical Education*, vol. 14, no. 1, p. 181, dec 2014. [Online]. Available: https://bmcmededuc.biomedcentral.com/articles/10.1186/1472-6920-14-181

[7] F. H. Wang, "An exploration of online behaviour engagement and achievement in flipped classroom supported by learning management system," *Computers & Education*, vol. 114, pp. 79–91, nov 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0360131517301483

[8] M. Zheng, C.-C. Chu, Y. Wu, and W. Gou, "The Mapping of On-Line Learning to Flipped Classroom: Small Private Online Course," *Sustainability*, vol. 10, no. 3, p. 748, mar 2018. [Online]. Available: https://www.mdpi.com/2071-1050/10/3/748

[9] B. Bhushan, P. Sinha, K. M. Sagayam, and A. J, "Untangling blockchain technology: A survey on state of the art, security threats, privacy services, applications and future research directions," *Computers & Electrical Engineering*, vol. 90, p. 106897, mar 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0045790620307497

[10] S. Wuthier and S.-Y. Chang, "Demo: Proof-of-Work Network Simulator for Blockchain and Cryptocurrency Research," in *2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)*. IEEE, jul 2021, pp. 1098–1101. [Online]. Available: https://ieeexplore.ieee.org/document/9546474/

[11] M. Saad, J. Spaulding, L. Njilla, C. Kamhoua, S. Shetty, D. Nyang, and D. Mohaisen, "Exploring the Attack Surface of Blockchain: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1977–2008, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9019870/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

eBPF-Hub: gestor de funciones eBPF

Jorge Gallego-Madrid, Miguel Carpe-Ferro, Ramon Sanchez-Iborra, Antonio Skarmeta
Departamento de Ingeniería de la Información y las Comunicaciones,
Universidad de Murcia
Facultad de Informática. Campus de Espinardo. 30100. Murcia.
jorgegm@um.es, miguel.carpef@um.es, ramonsanchez@um.es, skarmeta@um.es

Resumen

El crecimiento en la complejidad de las infraestructuras de red requiere de nuevas técnicas que permitan gestionar estos sistemas de forma eficiente y flexible. Una de las tecnologías que está atrayendo un mayor interés es *extended Berkeley Packet Filtering* (eBPF). Ésta permite la instanciación de programas dentro del kernel de Linux, permitiendo mejorar la eficiencia de ejecución de tareas críticas como el procesamiento de tráfico o la monitorización del propio sistema operativo. Sin embargo, desarrollar en eBPF es complejo y, además, actualmente no existe ningún repositorio centralizado para que la comunidad pueda compartir sus desarrollos. En este trabajo se presenta eBPF-Hub, una plataforma para el almacenamiento, gestión, descarga y ejecución de funciones eBPF, simplificando y optimizando su uso. La solución presentada permite manejar las funciones eBPF almacenadas en un repositorio remoto y otro local, facilitando la automatización de tareas, así como el acceso al ecosistema eBPF.

***Palabras Clave*—eBPF, repositorio, eBPF-Hub, VNF**

I. INTRODUCCIÓN

Desde el despliegue masivo de las redes de comunicaciones que se inició hace décadas, cada vez se genera y consume más tráfico, haciendo que infraestructuras y servicios de red evolucionen hacia una complejidad y exigencia mayor. Es por ello que uno de los puntos más importantes a tener en cuenta en la mejora de los sistemas de comunicación actuales y futuros es su flexibilidad. Ésta potencia la adaptabilidad de las infraestructuras de red a diferentes entornos, requisitos y demandas, la escalabilidad de las mismas para crecer y expandirse de acuerdo con las necesidades de la sociedad y la personalización de los servicios ofrecidos, según las preferencias de los usuarios [1]. En esta línea, conceptos como las redes definidas por software (*Software Defined Network*, SDN) o las funciones de red virtualizadas (*Virtualised Network Functions*, VNFs), juegan un papel clave en lograr la flexibilidad mencionada anteriormente, además de mejorar la eficiencia y portabilidad de las infraestructuras de red, así como de los elementos y funciones que las componen [2].

Una tecnología destacable en este proceso de *softwarización* de las redes de comunicaciones es *extended Berkeley Packet Filtering* (eBPF) [3]. Es una tecnología que permite ejecutar de forma segura programas dentro del kernel de Linux sin modificar el propio código del núcleo, de modo que se ha convertido en una herramienta muy poderosa y empleada a la hora de implementar funciones de red de manera eficiente y flexible [4]. Sin embargo, aunque eBPF ofrece un gran potencial para la implementación de funciones de red ágiles y poderosas [5], la complejidad de desarrollo en esta tecnología es muy alta, dadas las restricciones de seguridad impuestas por el propio kernel de Linux para poder ejecutar programas dentro de él. Además, la reutilización de funciones eBPF ya desarrolladas y validadas resulta de gran interés, ya que éstas pueden ser tomadas como unidades atómicas perfectamente funcionales, conformando por sí mismas funciones de red virtualizadas (VNFs) de gran portabilidad.

En este sentido, los gestores de VNFs son herramientas que encapsulan servicios en contenedores aislados, garantizando su correcto funcionamiento en cualquier entorno. Son fundamentales para acelerar el desarrollo de microservicios y mejorar la agilidad, eficiencia, seguridad y portabilidad en su despliegue. Acciones como la búsqueda manual de código relevante o la descarga individual de archivos y su posterior configuración pueden resultar procesos costosos y propensos a errores si no se emplean este tipo de herramientas. Considerando que eBPF es una tecnología que todavía está en sus etapas iniciales de adopción, así como sus prometedoras capacidades para habilitar la creación de funciones de red ligeras y flexibles, se identifica la necesidad de soluciones que gestionen el ciclo de vida de estas funciones y abstraigan a los administradores de infraestructuras de parte del proceso necesario para el despliegue de servicios de red virtualizados basados en eBPF.

Así, la principal contribución de este trabajo es el diseño e implementación de una plataforma centralizada que

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

facilita la descarga, gestión, compilación e instanciación de programas eBPF, denominada eBPF-Hub. A través de un programa cliente, esta solución ofrece un entorno en el que poder realizar múltiples tareas tanto en un repositorio central en la nube donde se almacenan archivos de configuración y código eBPF, como en la propia máquina local donde se van a configurar y/o desplegar las funciones de red descargadas del repositorio central. Según el análisis bibliográfico realizado por los autores, hasta la fecha no existe una herramienta específica que cumpla con las características que ofrece eBPF-Hub. Por lo tanto, las principales contribuciones de este artículo son las siguientes: (i) análisis del estado del arte de los repositorios y gestores de aplicaciones eBPF existentes, (ii) diseño e implementación de herramienta cliente-servidor para la gestión de programas eBPF, y (iii) validación del funcionamiento de la solución propuesta, evidenciando su utilidad para la gestión, configuración y despliegue de funciones de red basadas en eBPF.

El resto del documento se organiza de la siguiente forma: la Sección II sitúa este trabajo dentro de la literatura relacionada. El diseño e implementación de eBPF-Hub se explica en la Sección III. Las pruebas de validación se muestran en la Sección IV. La Sección V discute una serie de escenarios para el uso de la herramienta presentada. Finalmente, las conclusiones del artículo se ofrecen en la Sección VI, que también indica líneas de trabajo futuras.

II. ESTADO DEL ARTE

A continuación se exploran y describen los proyectos más importantes que hacen uso de eBPF, buscando tanto optimizar el desarrollo de funciones basadas en este tecnología como integrarlas en otras soluciones para mejorar su eficiencia o manejabilidad.

BPF Compiler Collection (BCC)[1], es un conjunto de herramientas de código abierto que permite crear programas eficientes para rastrear y manipular el kernel de Linux. Con BCC la escritura de programas eBPF se facilita gracias a los recursos en lenguaje C que ofrece acerca del núcleo e incluye *frontends* en Python y Lua. Una herramienta que se integra con BCC para interactuar en los sistemas eBPF es bpftrace[2], que es un lenguaje de trazado de alto nivel inspirado en awk y C que utiliza LLVM (*Low Level Virtual Machine*) para compilar código eBPF en bytecode.

Otro proyecto de gran relevancia dentro de la comunidad eBPF es Cilium[3], que es una herramienta fundamental para gestionar soluciones eBPF *cloud native* de forma eficiente y segura. Se trata de un proyecto de código abierto que ofrece una amplia gama de funcionalidades para simplificar la administración de redes, reforzar la seguridad y obtener una completa visibilidad de las aplicaciones y servicios que se ejecutan en clusters de Kubernetes y otras plataformas de orquestación de contenedores, todo ello basado en código eBPF. Por su parte, Falco[4] se centra en asegurar la seguridad de entornos en la nube soportados por sistemas Linux a través de la tecnología eBPF. Su misión principal es detectar y alertar en tiempo real sobre cualquier comportamiento anómalo o amenaza que sufra un sistema, por lo que opera como un agente de monitoreo en el kernel gracias a las herramientas provistas por eBPF.

En cuanto a la monitorización de clusters de Kubernetes, distintas soluciones hacen uso de eBPF para extraer métricas de los mismos. Por su parte, la herramienta Pixie[5] incorpora eBPF para recopilar automáticamente datos de telemetría en tiempo real de clusters de Kubernetes, incluyendo consumos de memoria o CPU de los nodos en el clúster, o incluso permite seguir la pista de las comunicaciones entre nodos. En esta línea, Calico[6] es una herramienta de código abierto que busca simplificar, escalar y proteger redes en entornos con contenedores basados también en Kubernetes. Gracias al enrutamiento y procesamiento de tráfico que ofrece eBPF, Calico es una solución que permite llevar a cabo tareas básicas como balanceo de carga, NAT, etc. para gestionar el tráfico dentro del cluster. También enfocado en tareas de gestión de red, Katran[7] es una solución para el balanceo de carga con un plano de reenvío completamente reformulado utilizando eBPF. Es remarcable destacar que esta herramienta de balanceo de carga se está utilizando en la infraestructura de un gigante tecnológico como Facebook. Desde una perspectiva más amplia, Polycube [6] es un proyecto de código abierto que proporciona funciones de red rápidas y livianas, como puentes, enrutadores, o cortafuegos, entre otros, implementadas en tecnología eBPF.

Finalmente, bpfman[8] es un gestor de funciones eBPF que se centra en facilitar los procesos de despliegue y administración de programas eBPF, tanto en una máquina local como en un cluster de Kubernetes. Este gestor también incluye un cargador integrado para poder instanciar programas eBPF además de ayudar a la implementación de programas eBPF desde imágenes OCI (*Open Container Inciative*). Se puede entender bpfman como un conjunto de herramientas utilizados para cargar, descargar, modificar y monitorizar programas eBPF.

A. Contribución

Como se puede observar, de todas las herramientas analizadas anteriormente, solo bpfman ofrece un entorno de gestión de funciones eBPF. Aunque esta solución es muy válida para la administración de programas eBPF, no dispone de un repositorio centralizado donde almacenar y hacer accesibles dichos programas a una comunidad o conjunto de desarrolladores. Tomando esta carencia como punto de partida, e inspirándose en la funcionalidad ofrecida por el conocido portal Docker Hub[9], la propuesta presentada en este artículo es una herramienta que permite

[1]https://github.com/iovisor/bcc
[2]https://github.com/bpftrace/bpftrace
[3]https://cilium.io/
[4]https://falco.org/
[5]https://px.dev/
[6]https://www.tigera.io/project-calico/
[7]https://github.com/facebookincubator/katran
[8]https://bpfman.io/
[9]https://hub.docker.com/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

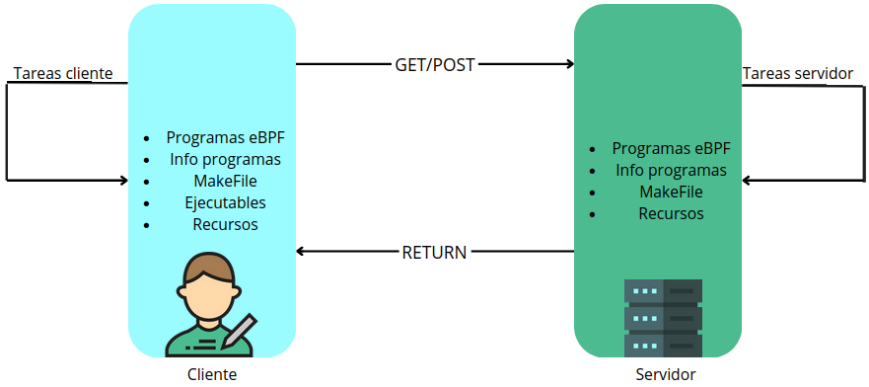


Fig. 1. Esquema de despliegue de la solución desarrollada.

la subida, gestión y descarga de programas eBPF en distintas modalidades, pudiendo hacer uso tanto de un repositorio en la nube como de otro local propio del usuario. eBPF-Hub permite trabajar tanto con archivos ya compilados, como con código fuente que puede compilado para una arquitectura determinada (compilación cruzada) en ambos repositorios bajo demanda. Finalmente, eBPF-Hub habilita la instanciación en la máquina local de cualquier función eBPF alojada en alguno de los repositorios (local o remoto), haciendo uso de distintos cargadores compatibles con diversos puntos de anclaje o *hooks*, dentro del kernel de Linux [7]. De este modo, eBPF-Hub se erige como una herramienta versátil y poderosa para la compartición y gestión de funciones eBPF, facilitando mucho las labores de manejo de este tipo de programas.

III. DISEÑO E IMPLEMENTACIÓN

Tal y cómo se ha mencionado anteriormente, eBPF-Hub adopta una arquitectura cliente-servidor (Fig. 1). Mientras que la parte servidora actúa como repositorio remoto en el que alojar y manejar programas eBPF, el cliente conforma la interfaz con la que un usuario puede interaccionar con la solución. El código de ambos desarrollos (Python 3.11.6) se puede encontrar en un repositorio público[10]. A continuación, se describe en detalle los dos componentes de eBPF-Hub.

A. Cliente

En la parte del cliente se cuenta con un repositorio local que aloja diferentes archivos eBPF (código fuente, compilados, ficheros de configuración, etc.) sobre los que realizar ciertas tareas sin interactuar con el servidor (listar, ver detalles, compilar, cargar, etc.). Además de estas tareas en local, a través de peticiones HTTP (GET y POST) se va a comunicar con el servidor remoto para poder interactuar con él, así como con los ficheros que aloja.

El cliente se ha desarrollado como una herramienta que ofrece una interfaz al usuario a través de línea de comandos, con la siguiente funcionalidad (Fig. 2):

- Descargar archivos del repositorio remoto.
- Subir archivos al repositorio remoto.
- Borrar archivos tanto del repositorio remoto como del local.
- Listar los archivos que se encuentran tanto en el repositorio remoto como en el local.
- Cargar funciones eBPF en la máquina local.
- Obtener información sobre las funciones que están cargadas en la máquina local.
- Quitar funciones cargadas de la máquina local.
- Compilar ficheros eBPF tanto en el repositorio local como remoto.
- Obtener información sobre cualquier archivo que esté en el repositorio remoto o local.
- Desplegar ayuda con información acerca de los argumentos admitidos por la herramienta.

El cliente hace uso de las librerías ofrecidas por el proyecto XDP-tools[11] para realizar la interacción con los funciones eBPF. En concreto, se hace uso de la librería *libxdp*, la cual ofrece un *Makefile* y un cargador para programas eBPF muy versátiles, que facilitan tanto la compilación como la instanciación de éstos.

Centrándonos en la función de compilación, partiendo del fichero *Makefile* base, el programa cliente es capaz de adaptarlo a las características del programa a compilar, por ejemplo, modificando el campo *TARGET* de dicho fichero. Además, en este archivo se definen las diferentes dependencias y reglas de construcción específicas para el programa eBPF a compilar.

Por su parte, la instanciación de un programa eBPF se realiza a través de un cargador; en este caso, se ha hecho uso del cargador *xdp-loader*. Ya que el foco de este trabajo está en el manejo de VNFs, la función principal de esta herramienta es cargar programas eBPF en interfaces de red específicas. Otros cargadores permiten la instanciación de programas eBPF en otros puntos (*hooks*) del sistema operativo, aunque el uso de éstos ha quedado fuera del alcance de este trabajo. No obstante, el cliente se ha programado de forma modular, por lo que que se puede usar cualquier otro cargador eBPF simplemente cambiando un argumento de entrada que recibe el programa, como se verá más adelante.

[10]https://github.com/miguelCarpe/eBPFhub

[11]https://github.com/xdp-project/xdp-tools

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

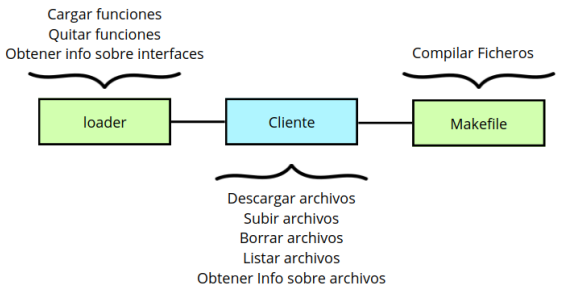


Fig. 2. Funcionalidad del cliente.

Cada programa eBPF lleva asociado un fichero en formato JSON que incluye la siguiente información:

- *Description*: campo con el cual se hace una breve descripción del programa eBPF y sus funcionalidades, añadiendo también ciertos requisitos básicos para su correcto uso.
- *Maps*: campo que únicamente tienen los ficheros que describen programas eBPF que trabajan con mapas, los cuales se utilizan para compartir datos entre el espacio de usuario y el kernel.
- *Date*: campo que determina la fecha de última modificación del programa eBPF.
- *Version*: campo que define la versión del programa eBPF. Este campo es útil si queremos trabajar con diferentes versiones del mismo programa.
- *Sources*: campo que nos indica ficheros de los que depende el programa eBPF para lograr un funcionamiento correcto y completo.

Este fichero informativo es utilizado tanto cuando el usuario solicita información acerca de un determinado programa, como cuando el programa se va a compilar o cargar, ya que ciertos campos del fichero informan de dependencias del programa o el uso de mapas de memoria. Considerando estos últimos, los mapas son un componente esencial en los programas eBPF, puesto que permiten almacenar datos de forma persistente y que éstos sean accesibles tanto desde espacio de kernel como de usuario. Para poder anclar estos mapas a la memoria de manera transparente al usuario y que sean accesibles desde el programa a cargar, el programa cliente consulta en el archivo de información del programa si éste tiene mapas asociados. En caso de que así sea, se adapta el fichero *Makefile* y el cargador de forma adecuada para que el programa instanciado tenga acceso al mapa de memoria que necesita para poder ejecutarse correctamente.

Finalmente, además de la funcionalidad descrita hasta ahora, la cual permite manejar los ficheros eBPF alojados en el repositorio local, el cliente también permite interactuar con el servidor remoto. Como se ha comentado anteriormente, para ello se hace uso de peticiones HTTP. Esta funcionalidad la habilita la librería *requests*, a través de la cual se realizan al servidor peticiones GET para manejar los diferentes ficheros alojados en él y peticiones POST para subir archivos al repositorio remoto. A continuación, se presenta la funcionalidad del servidor.

B. Servidor

El servidor ofrece un repositorio remoto en el que los clientes van a poder subir, descargar e interactuar con los programas eBPF. Las funciones que implementa el servidor son las siguientes:

- Descarga de un fichero eBPF específico (codigo fuente o ejecutable) o de todos los archivos relacionados con un programa eBPF.
- Subida de archivos.
- Obtener información de archivos.
- Listar archivos.
- Eliminar archivos.
- Compilar programa eBPF a una determinada arquitectura.

Para el desarrollo del servidor se ha partido de la librería *Flask*[12], la cual ofrece un entorno ligero, flexible y fácil de usar para la creación de esta parte de la solución. Al utilizar *Flask*, se definen distintos *@app.get* y *@app.route*. Con estas definiciones se asocia una función por cada tipo de petición realizada por el cliente. Las primeras se utilizan para peticiones de descarga, mientras que las segundas para subir archivos y para hacer peticiones que no impliquen transferencias de ficheros.

Además de la funcionalidad típica de todo repositorio, que permite descargar, subir o solicitar información de un determinado programa eBPF (a través del fichero informativo JSON asociado a éste), el servidor también ofrece la opción de descargar un programa eBPF y todos sus dependencias asociadas, así como compilar un determinado programa para la posterior descargar del binario generado. Considerando el primer caso, el campo *Sources* del fichero JSON asociado a un determinado programa presenta todos los ficheros que acompañan al mismo. Si un usuario solicita la descarga de un programa completo con todas sus dependencias, el servidor extrae de aquí la lista de ficheros a proveer, los comprime en un fichero y procede con el envío de éste al cliente. A su recepción, el cliente descomprime el archivo recibido obteniendo así los ficheros que conforman el programa completo solicitado.

En cuanto a la posibilidad de compilar funciones eBPF alojadas en el repositorio remoto, el motivo de esta funcionalidad es facilitar el acceso a estos programas, en forma de VNFs listas para ser cargadas y ejecutadas, a entidades con bajos recursos de cómputo (por ejemplo, dispositivos de Internet de las Cosas (IoC)) o en el caso de carecer de las herramientas necesarias para la compilación de un determinado programa. En la petición realizada por el cliente, éste indica un dato adicional a cuando la compilación se hace en local: la arquitectura destino. Así, mediante un compilador que soporte compilación cruzada, el servidor es capaz de proveer el binario del programa solicitado para la arquitectura específica del dispositivo destino, abriendo así un gran abanico de interacciones con toda clase de dispositivos. Una vez que se realiza la compilación, se produce la descarga del binario generado

[12]https://flask.palletsprojects.com/en/stable/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

y el cliente procede con la instanciación en local de este ejecutable a través de su cargador.

A modo de resumen, la Fig. 3 muestra un diagrama de interacción entre cliente y servidor.

IV. Validación

En esta sección se muestran algunos ejemplos de uso de la herramienta que validan su funcionamiento. En primer lugar, el Código 1 presenta la salida que ofrece el cliente cuando se le solicita que muestre todas las opciones disponibles (argumento *−−help*).

Código 1. Funcionalidad de eBPF-Hub.

```
$ python3 eBPFhub.py −−help
Available options:
−d: download file from the remote repository
−dc: download all related files from remote repository
−i: show information about a file
−le: list .c files in the remote repository
−lea: list all files in the remote repository
−ll: list .c files in the local repository
−lla: list all files in the local repository
−m: compile a local file
−ni: create virtual interfaces with the testenv tool.
−re: remove file from the remote repository
−rl: remove file from the local repository
−u: upload file to the remote repository
−a: compile and load an eBPF program
    −h: eBPF hook
    −if: eBPF program loading interface (if needed)
    −−loader: function loader
    −p: program to load
        −−section: ELF section name of the program to load
        −−progname: eBPF program name of the program to load
−de: download an executable object of a file (after its remote compilation)
    −arc: target architecture
−l: load a local eBPF program
    −h: eBPF hook
    −if: eBPF Program Loading Interface (if needed)
    −−loader: function loader
    −p: program to load
    −−section: ELF section name of the program to load
    −−progname: eBPF program name of the program to load
−s: view running eBPF functions
    −h: eBPF hook
    −if: eBPF program loading interface (if needed)
    −− loader: function loader
−un: unload an eBPF function
    −h: eBPF hook
    −if: eBPF program loading interface (if needed)
    −−loader: function loader
    −−id: ID of the function
    all: unload all functions
```

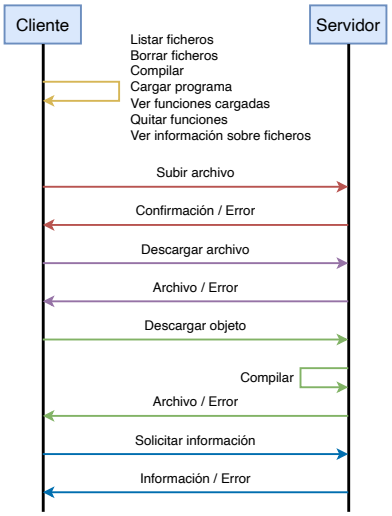


Fig. 3. Diagrama de interacción.

Como se observa, se implementa toda la funcionalidad explicada en la sección anterior a través de diversos parámetros de entrada que se le pasan al programa cliente. Las primeras opciones mostradas son aquellas que no requieren argumentos de entrada o tan solo uno, normalmente el nombre del fichero sobre el que realizar la acción. Posteriormente, se muestran aquellas funcionalidades más complejas que hacen uso de varios argumentos de entrada.

Comenzando con la funcionalidad más básica, el Código 2 muestra la salida que ofrece el cliente cuando se le solicita información sobre un programa eBPF concreto (argumento *−i*). En primer lugar, se realiza la búsqueda de este programa en el repositorio local y, en caso de no encontrarlo, se lanza la búsqueda en el repositorio remoto. Una vez encontrado el programa buscado, se muestra la información contenida en el fichero informativo (JSON) asociado al mismo.

Código 2. Información de programa eBPF.

```
$ python3 eBPFhub.py -i xdp_prog_kern_03
    Description: Description of xdp_prog_kern_03
    Date: 25/04/2024
    Maps: ['tx_port', 'redirect_params']
    Sources: executable1, executable2
    Version: 1.0.1
```

En el ejemplo mostrado, la función eBPF en concreto que se examina (*xdp_prog_kern_03*) hace uso de dos programas complementarios (por ejemplo, ejecutándose en espacio de usuario para habilitar cierta funcionalidad de alto nivel) y de dos mapas, en ambos casos información fundamental a la hora compilar y ejecutar la función.

Así, tal y como se ha comentado anteriormente, es posible realizar la compilación de un determinado programa eBPF en el propio servidor para únicamente

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

descargar el objeto eBPF ejecutable. Para ello, se hace uso del argumento *−de*, junto con el parámetro *−arc* para indicar la arquitectura destino (ver Código 3). Esto permite que equipos con pocos recursos puedan hacer uso de funciones eBPF ya listas para su ejecución sin tener que realizar la compilación por ellos mismos, siendo todo el ecosistema de IoC un claro caso de uso.

Código 3. Compilación remota y descarga de ejecutable.

```
$ python3 eBPFhub.py -de xdp_vlan01_kern -arc arm
    xdp_vlan01_kern.o succesfully downloaded
```

En este ejemplo, la arquitectura de compilación destino es *arm*, aunque podría ser otra, como *x86_64*, siempre que el compilador empleado en el servidor soporte el tipo de compilación cruzada solicitada.

Una vez que el objeto ejecutable se tiene almacenado en el repositorio local (ya sea por compilación local o descarga desde repositorio remoto), éste se debe instanciar en un punto de anclaje del sistema operativo (*hook*), haciendo uso de un cargador de funciones eBPF. Para ello se utiliza el comando mostrado en el Código 4.

Código 4. Instanciación de función eBPF.

```
$ python3 eBPFhub.py -l -h xdp -if eth0 --loader
xdp-loader -p xdp_prog_kern_03.o --section
xdp_icmp_echo
    The eBPF program has been loaded successfully,
read its description to finish the configuration
```

Como se puede observar, se requieren múltiples argumentos de entrada para proceder con la carga del programa de forma correcta. Además de indicar el programa a cargar (argumento *−p*), se debe introducir el punto de anclaje en el sistema operativo (*hook*) y la interfaz donde conectar el programa (argumentos *−h* y *−i*, respectivamente), el cargador a emplear (argumento *−−loader* y la sección de entrada al programa en cuestión (argumento *−−section*).

Dado que el proceso de descarga, compilación e instanciación de un determinado programa puede ser tedioso, el argumento *−a* permite realizar todas estas operaciones de forma automática (Código 5).

Código 5. Automatización de descarga, compilación e instanciación de función eBPF.

```
$ python3 eBPFhub.py −a xdp_prog_kern_03 −h xdp -if
eth0 −−loader xdp-loader
    The download of module xdp_prog_kern_03 has been
completed successfully
    Compilation of xdp_prog_kern_03.c has been completed
successfuly
    The eBPF program has been loaded successfully, read
its description to finish the configuration
```

En concreto, el comando mostrado busca en primer lugar la función eBPF solicitada en el repositorio local. En caso de no encontrarla, la descarga del repositorio remoto (asumiendo que se allí se encuentra, en caso contrario se mostraría un mensaje de error), la compila de forma local (por eso no es necesario indicar la arquitectura destino con el argumento *−arc*) y la carga en la interfaz y punto de anclaje indicados (argumentos *−if* y *−h*, respectivamente), haciendo uso del cargador identificado a través del parámetro *−−loader*.

Una vez cargado en local un programa por alguno de los métodos mostrados hasta ahora, se puede consultar el listado de funciones eBPF que están ejecutándose en un *hook* e interfaz concretos, a través del argumento *−s*, e indicando el punto de anclaje e interfaz a consultar. La llamada de sistema para realizar esta acción está contenida en un cargador eBPF, por lo que también hay que indicar éste, tal y como se muestra en el Código 6.

Código 6. Consulta de procesos eBPF ejecutándose.

```
$ python3 eBPFhub.py −s −h xdp −if eth0 −−loader
xdp-loader

- Iface  Prio  Name      Mode    ID   Tag     Action     -
  eth0   50    xdp_echo  native  116  094a    XDP_PASS
```

Se comprueba como el programa anteriormente cargado aparece listado dentro las funciones eBPF ejecutándose en la interfaz *eth0*, proporcionándose una serie de valores informativos acerca del mismo.

Finalmente, también se puede eliminar (argumento *−un*) una función eBPF que esté ejecutándose en un determinado punto de anclaje e interfaz, a través del identificador de ésta a nivel de sistema operativo (Código 7); dicho identificador se puede extraer de la salida del comando anterior (*116* en este caso).

Código 7. Eliminación de función eBPF en ejecución.

```
$ python3 eBPFhub.py -un -h xdp -if eth0 –loader xdp-
loader –id 116
    Function with ID 116 has been unload from the
interface eth0 successfully

$ python3 eBPFhub.py -s -h xdp -if eth0 –loader
xdp-loader

- Iface  Prio  Name  Mode  ID  Tag  Action  -
  eth0 - No eBPF program loaded!
```

Con el segundo comando mostrado se verifica que la función ha sido eliminada correctamente de la interfaz *eth0*. Si se desean quitar todas las funciones eBPF ejecutándose en un determinado *hook* e interfaz, en vez de utilizar el identificador concreto de una función eBPF, se puede emplear la cadena *all*.

Los ejemplos presentados en esta sección evidencian el correcto funcionamiento de la herramienta y su aplicabilidad en distintas situaciones, respaldando así la validez de su diseño y desarrollo. De este modo, se puede considerar eBPF-Hub como un recurso fiable y versátil, que puede tener un amplio campo de aplicación, tal y como se discute a continuación.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

V. Casos de Uso

En esta sección se identifican una serie de escenarios donde el uso de eBPF-Hub puede resultar de utilidad, facilitando el manejo y la integración de funciones eBPF en distintos entornos.

Como ya se ha comentado, en el ámbito de la gestión de red, eBPF-Hub permite explotar la eficiencia de eBPF para el procesamiento de paquetes, el balanceo de carga, la inspección profunda de tráfico y el manejo de políticas de red en tiempo real. Plataformas citadas anteriormente como Cilium y Calico utilizan eBPF para ofrecer redes de alto rendimiento en entornos de contenedores y Kubernetes, y eBPF-HUB puede actuar como un repositorio centralizado para compartir y reutilizar funciones de red optimizadas, acelerando su adopción.

Otro de los principales escenarios de aplicación de eBPF es la monitorización avanzada de rendimiento y la observabilidad de sistemas. Gracias a la capacidad de eBPF para instrumentar el kernel y recolectar métricas en tiempo real con un impacto mínimo en el funcionamiento del sistema, los administradores pueden hacer uso de eBPF-Hub para desplegar automáticamente programas que rastreen llamadas al sistema, analicen el uso de CPU, identifiquen cuellos de botella de E/S o monitoricen el comportamiento de aplicaciones y microservicios. Esto resulta especialmente útil en entornos de producción donde la monitorización granular y la capacidad de reacción rápida ante incidencias son críticas.

Otro caso de uso fundamental es la ciberseguridad. eBPF permite crear soluciones de detección de intrusos, monitorización de llamadas sospechosas al sistema, control de acceso, protección en tiempo de ejecución de contenedores y aplicaciones, etc. eBPF-Hub facilita la distribución y actualización de estos programas de seguridad, permitiendo a los equipos de operaciones y ciberseguridad implementar rápidamente nuevas políticas o adaptarse a amenazas novedosas sin necesidad de desarrollar desde cero o modificar el kernel de los equipos en producción.

eBPF también es una tecnología muy valiosa para dispositivos embebidos y sistemas basados en arquitecturas ARM, como todo el ecosistema de IoC. Estos dispositivos, como microcontroladores, *routers*, *gateways* o microcomputadoras tipo Raspberry Pi, requieren programas que permitan maximizar la eficiencia en su ejecución, garantizando la seguridad del sistema destino sin introducir una sobrecarga significativa, tal y como ofrece eBPF. Sin embargo, estos dispositivos suelen caracterizarse por presentar recursos limitados en cuanto a CPU, memoria y capacidad de almacenamiento. Así, el uso de eBPF-Hub permite el acceso de estos dispositivos a un repositorio de funciones eBPF que pueden ser configuradas bajo demanda y descargadas listas para su instanciación y ejecución sin tener que consumir recursos de cómputo, por ejemplo, para su compilación.

Con todo lo anterior, eBPF-Hub facilita el acceso al ecosistema eBPF al ofrecer una plataforma donde usuarios sin gran experiencia pueden encontrar, descargar y ejecutar funciones validadas por la comunidad, disminuyendo la barrera de entrada y fomentando la colaboración. Esto es especialmente relevante en entornos educativos, laboratorios de investigación o entidades que buscan experimentar con eBPF sin asumir los riesgos y la curva de aprendizaje asociados al desarrollo directo en el kernel de Linux.

VI. Conclusiones

El uso de eBPF cada vez se extiende más, dadas las posibilidades que habilita para gestionar de forma eficiente tareas críticas en distintos ámbitos: procesamiento de tráfico, monitorización de sistemas, etc. Sin embargo, la curva de aprendizaje de esta tecnología es muy elevada y la comunidad eBPF podría verse beneficiada si contara con una plataforma central en la que compartir desarrollos. Partiendo de estas premisas, y tomando como referencia plataformas agregadoras de contenedores como Docker Hub, este trabajo ha presentado eBPF-Hub, una herramienta que permite gestionar funciones eBPF alojadas tanto en repositorio local como remoto y automatizar de forma sencilla tareas como el compilado, instanciación y eliminado de estas funciones. La herramienta ha sido publicada en abierto para el uso de toda la comunidad y actualmente se sigue trabajando en su desarrollo. Por un lado, se pretende dotarla de un entorno gráfico para poder hacer uso de ella de una forma más visual y, por otro, se trabaja en la integración de un generador de código eBPF basado en inteligencia artificial para facilitar el desarrollo de programas ajustados a las necesidades de los usuarios.

Agradecimientos

Este trabajo ha sido financiando a través del proyecto PID2023-148104OB-C43 financiado por MICIU/AEI/10.13039/501100011033 y por FEDER/UE; por la Comisión Europea, a través del proyecto NANCY (101096456); y por el Ministerio de Economía y Transformación Digital, a través del proyecto CERBERUS-HADES (TSI-063000-2021-62).

Referencias

[1] A. M. Alberti, D. G. Pivoto, T. T. Rezende, A. V. Leal, C. B. Both, M. S. Facina, R. Moreira, and F. de Oliveira Silva, "Disruptive 6G architecture: Software-centric, AI-driven, and digital market-based mobile networks," *Computer Networks*, vol. 252, p. 110682, oct 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1389128624005140

[2] M. S. Bonfim, K. L. Dias, and S. F. L. Fernandes, "Integrated NFV/SDN Architectures," *ACM Computing Surveys*, vol. 51, no. 6, pp. 1–39, nov 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3172866

[3] L. Song and J. Li, "eBPF: Pioneering Kernel Programmability and System Observability - Past, Present, and Future Insights," in *2024 3rd International Conference on Artificial Intelligence and Computer Information Technology (AICIT)*, 2024, pp. 1–10.

[4] J. Gallego-Madrid, I. Bru-Santa, A. Ruiz-Rodenas, R. Sanchez-Iborra, and A. Skarmeta, "Machine learning-powered traffic processing in commodity hardware with eBPF," *Computer Networks*, vol. 243, p. 110295, apr 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1389128624001270

[5] J. Gallego-Madrid, A. Molina-Zarca, R. Sanchez-Iborra, J. Ortiz, and A. F. Skarmeta, "Fast traffic processing in multi-tenant 5G environments: A comparative performance evaluation of P4 and eBPF technologies," *Engineering Science and Technology, an International Journal*, vol. 52, p. 101678, apr 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2215098624000648

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[6] S. Miano, F. Risso, M. V. Bernal, M. Bertrone, and Y. Lu, "A Framework for eBPF-Based Network Functions in an Era of Microservices," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 133–151, mar 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9340283/

[7] M. A. M. Vieira, M. S. Castanho, R. D. G. Pacífico, E. R. S. Santos, E. P. M. C. Júnior, and L. F. M. Vieira, "Fast Packet Processing with eBPF and XDP: Concepts, Code, Challenges, and Applications," *ACM Computing Surveys*, vol. 53, no. 1, pp. 1–36, may 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3371038

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Evaluación de Mecanismos de Operación Multicanal para Redes Vehiculares

Felipe García Vidal, Esteban Egea López y José Santa Lozano
Escuela Técnica Superior de Ingeniería de Telecomunicación
Universidad Politécnica de Cartagena
Antigones, 30202 Cartagena, Murcia
felipe.gvidal@upct.es, esteban.egea@upct.es y jose.santa@upct.es

Con el despliegue masivo de los sistemas de transporte inteligente cooperativo (en adelante, C-ITS, del inglés *Cooperative Intelligent Transport System*), los servicios definidos en la Release 2 del instituto europeo de normas de telecomunicaciones (en adelante, ETSI, del inglés *European Telecommunications Standards Institute*) no pueden desplegarse utilizando un único canal de control, y requieren la gestión de múltiples canales. Por este motivo, se ha definido un concepto para utilizar los canales de comunicación disponibles y gestionar su funcionamiento simultáneo, conocido como operación multicanal (MCO, *Multichannel Operation*). Dos mecanismos simples han sido propuestos para la asignación de carga en múltiples canales: *load balancing* y *sequential filling*, pero hasta ahora no se habían evaluado en detalle. En este trabajo se analizan ambos mecanismos junto con sus versiones que emplean un control de congestión básico. A través de múltiples simulaciones, se estudia el comportamiento de los mecanismos en escenarios con un medio altamente congestionado, empleando distintas cargas de tráfico. Los resultados muestran que ninguno de ellos ofrece un rendimiento aceptable, especialmente bajo condiciones de alta carga. Sus principales limitaciones son analizadas y se ofrecen recomendaciones para orientar la evolución de estos mecanismos.

***Palabras Clave*—C-ITS, ETSI, MCO, mecanismos de uso de canal, redes vehiculares.**

I. INTRODUCCIÓN

En Europa, se han asignado varios canales en la banda de 5.9 GHz para las comunicaciones vehiculares, adoptando el estándar IEEE 802.11p como capa física y de acceso al medio, lo que se conoce como ITS-G5 [1], [2]. La asignación de espectro propuesta incluye cinco canales de 10 MHz *safety-related*, situados entre 5875 y 5925 MHz, y dos canales adicionales *no safety-related* entre 5855 y 5875 MHz. Para la *Release 1*, se identificó y definió un conjunto básico de aplicaciones en la primera versión del ETSI TR 102 638 [3], basadas en el *Cooperative Awareness Service (CAS)* [3] y en el *Decentralized Environmental Notification (DEN) service* [4]. Estos servicios utilizan un único canal de control en 5895 MHz, que se ha demostrado suficiente para los casos de uso de la *Release 1*.

Desde 2017, en una segunda versión del ETSI TR 102 638 [3] se han considerado un nuevo conjunto de servicios para la *Release 2*, incluyendo *Connected and Cooperative Automated Mobility (CCAM)*, *Collective Perception (CP)* y *Vulnerable Road Users (VRU)*. Estos servicios no pueden implementarse usando un único canal de control, por lo que requieren ser gestionados para operar en múltiples canales, lo que en el contexto de los C-ITS se conoce como *operación multicanal*. Varios estudios han analizado los requisitos funcionales y las limitaciones técnicas de este tipo de sistemas. En particular, el trabajo de [5] analiza los requisitos comerciales, funcionales y de seguridad de varias aplicaciones de la *Release 2*, y los autores de [6] proponen un concepto del *MCO* basado en ese análisis.

Por otro lado, las simulaciones presentadas en ETSI TR 103 439 [7] analizan un escenario donde las estaciones C-ITS operan en distintos canales de ITS-G5, evaluando principalmente los efectos de la interferencia en canales adyacentes. Las conclusiones y recomendaciones de [7] y [6] indican que distribuir el tráfico entre múltiples canales mejora el rendimiento del sistema. También destacan que, si bien la interferencia de canales adyacentes lejanos es insignificante, la interferencia del primer canal adyacente es significativa y debe ser gestionada. El enfoque sugerido para gestionar esta interferencia se basa en mecanismos de *control de congestión*. En lugar de ajustar la potencia de transmisión se establecen límites adecuados de carga, con una *carga máxima por canal* [7], [6]. Sin embargo, los mecanismos de gestión multicanal siguen siendo un tema abierto de investigación, sin que se haya adoptado aún ningún método concreto. ETSI TR 103 439 [7] examina tres estrategias potenciales de *uso de canales*: *sequential filling*, donde el tráfico se asigna a los canales en un orden

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

predefinido, llenando cada uno hasta un umbral antes de pasar al siguiente; *load balancing*, que distribuye el tráfico de forma uniforme entre los canales, siguiendo un enfoque tipo *round-robin*; y *elastic allocation*, que distribuye el tráfico de forma dinámica según distintos criterios. Además del *uso del canal*, los criterios de *asociación de canal* determinan cómo las aplicaciones asignan sus mensajes a canales específicos. Además, se proponen dos estrategias generales de asociación: *predefined* y *flexible*, que pueden combinarse con los métodos de *uso de canal* mencionados.

Sin embargo, hasta donde sabemos, ninguna de estas recomendaciones ha sido evaluada ni comparada. Por tanto, en este trabajo analizamos el rendimiento del MCO para los C-ITS basándonos en dichas recomendaciones. En concreto, evaluamos mediante simulación dos de los enfoques de *uso de canal* propuestos en [7]: *load balancing* y *sequential filling*, combinados con una *asociación de canal flexible*. El objetivo es evaluar el rendimiento de estos mecanismos de *uso de canal* en escenarios exigentes, con un medio compartido altamente congestionado. Para cada mecanismo, estudiamos dos variantes: una sin *control de congestión* y otra con un *control de congestión* básico.

El resto del artículo se organiza de la siguiente manera: en la Sección II revisamos trabajos relacionados. En la Sección III presentamos los mecanismos evaluados y seguidamente describimos el entorno de simulación en la Sección IV. A continuación, en la Sección V mostramos y analizamos los resultados. Por último, en la Sección VI ofrecemos un resumen general y discutimos nuestros hallazgos, y cerramos con la Sección VII.

II. TRABAJO RELACIONADO

Pocas propuestas han abordado específicamente los problemas del rendimiento del MCO hasta ahora. Una visión general del funcionamiento y los retos de investigación de la *operación multicanal* en los C-ITS han sido discutidos recientemente en [8], donde se identifican los desafíos de integrar el *control de congestión* y la gestión de interferencias. En el contexto de *Maneuver Coordination* (MC), los autores de [9] utilizan la *operación multicanal* pero únicamente para asignar diferentes tipos de mensajes a dos canales, es decir, no se proporciona una evaluación específica de los mecanismos del MCO. El trabajo de [10] propone una extensión del mecanismo reactivo *Decentralized Congestion Control* (DCC) de la *Release 1* [11] para decidir el canal asignado a los mensajes, en función de la congestión del canal medida mediante el *Channel Busy Ratio* (CBR). Sin embargo, su evaluación es bastante limitada y solo muestra que el CBR se distribuye entre los tres canales utilizados cuando el DCC está habilitado. Además, se observa que los mecanismos reactivos mantienen el CBR oscilando alrededor del umbral, alcanzando solo el valor medio deseado pero con grandes fluctuaciones.

En un trabajo reciente, propusimos MINOS [12], un esquema de uso *elástico* de canales, basado en un planificador óptimo, cuyo objetivo es maximizar el tráfico en los canales minimizando la interferencia en canales

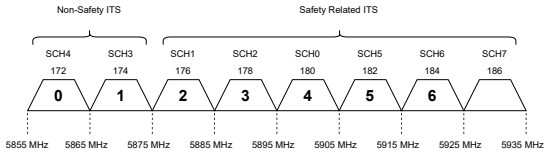


Fig. 1. Canales C-ITS

adyacentes generada por dicha carga y controlando la congestión. MINOS funciona sin problemas incluso cuando los vehículos disponen de un número diferente de interfaces de red, y permite establecer restricciones y prioridades de forma individual y dinámica. Sin embargo, ese trabajo no incluye una evaluación de los mecanismos de *load balancing* y *sequential filling*, como se realiza en este artículo.

Otros trabajos han discutido procedimientos de operación multicanal, pero no específicamente en el contexto del MCO definido por el ETSI. Por ejemplo, SAMCO [13] proporciona un heurístico para el cambio (selección) de canal basada en la priorización del servicio y la carga del canal. CarHet [14] propone un heurístico para seleccionar la interfaz a utilizar, permitiendo que los vehículos empleen múltiples tecnologías de acceso por radio, como ITS-G5 y redes celulares.

Cabe mencionar también que los mecanismos *round-robin* y *load balancing* han sido evaluados en muchos contextos distintos, y sus variantes ponderadas presentan mejor rendimiento siempre que los pesos se ajusten de acuerdo a los requisitos y objetivos del contexto particular. El *sequential filling* no se utiliza en los contextos revisados, al menos en su versión estática, posiblemente debido a su bajo rendimiento, y solo aparece en variantes más sofisticadas y dinámicas.

III. MECANISMOS DE USO DE CANAL

En esta sección describimos brevemente los mecanismos de uso de canales evaluados. Para una comprensión adecuada, es necesario previamente aclarar la numeración de los canales que se muestra en la Fig. 1. En los resultados y figuras de este artículo nos referiremos a los canales según su orden físico, comenzando por el canal 0, que corresponde a SCH4 en el estándar ETSI y al canal 172 en la nomenclatura IEEE. El último canal, SCH7, no se utiliza en las simulaciones, ya que está reservado exclusivamente para ITS ferroviario [7].

Comparamos dos variantes tanto de *sequential filling* como de *load balancing*, una sin control de congestión y otra con un mecanismo simple de control de congestión. *Sequential filling* es el mecanismo recomendado por ETSI TR 103 439 [7] para servicios de difusión avanzada, mientras que *load balancing* no lo es, principalmente porque solo resulta efectivo cuando todos los vehículos disponen del mismo número de interfaces. Para nuestro estudio asumimos que este es el caso, por lo que todos los vehículos en nuestras simulaciones cuentan con el mismo número de interfaces de red. Hemos de tener en cuenta

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

también que solo consideramos una política de *asociación flexible*, es decir, las aplicaciones o servicios no están asignados a un canal predefinido.

1) SEQUENTIAL FILLING: El mecanismo de *sequential filling* organiza los canales en un orden predefinido, y un canal determinado no se utiliza hasta que los canales anteriores en dicho orden estén suficientemente cargados. Es decir, la transmisión se realizará siempre a través del canal con mayor prioridad cuya ocupación no exceda el límite establecido para el CBR, C_{max}. Este procedimiento presenta dos características cuestionables: primero, en cuanto el CBR medido en un canal esté por debajo de C_{max}, los paquetes volverán a encolarse en él. Esto hará que algunos canales estén completamente saturados mientras que otros permanezcan infrautilizados, como veremos más adelante. Segundo, el orden de selección establecido para los canales es 4, 2, 3, 5, 0, 1 y 6, y es el mismo para todos los vehículos. Esta secuencia es la indicada en [7], basada en las regulaciones actuales del espectro. El orden de los canales afecta a la interferencia entre canales adyacentes, ya que en el orden anterior solo los dos primeros canales disponibles no son adyacentes. Para reducir la interferencia entre canales adyacentes existen secuencias mejores, pero mantenemos esta debido a las limitaciones regulatorias, como se ha mencionado.

El control de congestión para el mecanismo de *sequential filling* consiste en definir un cierto límite de CBR a partir del cual se cambia a otro canal. En el caso sin control de congestión, se desea utilizar completamente un canal antes de cambiar, pero en la práctica nunca se mide un valor de CBR igual a 1. Por tanto, es necesario establecer a partir de qué valor de CBR se considera que un canal está completamente utilizado (saturado). En nuestro caso, se fija $C_{max} = 0.82$ como el límite para cambiar a otro canal. Cuando el control de congestión está habilitado, se establece $C_{max} = 0.62$ para todos los canales, lo cual coincide con el límite indicado en el DCC en [7]. Si todos los canales superan el valor de C_{max}, los nuevos paquetes generados se descartan hasta que haya al menos un canal con un valor de CBR por debajo del límite establecido. El número de paquetes descartados en este caso se presentará en los resultados.

2) LOAD BALANCING: El mecanismo de *load balancing* distribuye los paquetes entre los diferentes canales utilizando un algoritmo de tipo *round-robin*, comenzando en un canal aleatorio y procediendo en orden ascendente a través de todos los canales. Una vez que se alcanza el último canal, el mecanismo reinicia el ciclo en el canal cero. Con este mecanismo, la carga de canal en los distintos canales es similar bajo circunstancias generales.

Cuando se emplea el mecanismo sin control de congestión, no se monitorea el CBR de los canales. Sin embargo, dado que es posible que los paquetes se descarten debido a desbordamientos en las colas del MCO, se ha monitoreado el número promedio de paquetes/s perdidos por nodo en este nivel y se ha confirmado que, en los peores casos, solo se pierde aproximadamente un 0,01% de los paquetes. En el caso con control de congestión, los canales cuyo CBR excede $C_{max} = 0.62$ se eliminan de la secuencia de canales disponibles y no se utilizan hasta que el CBR en dichos canales vuelva a estar por debajo del límite. En ese caso, cuando todos los canales han superado el C_{max}, se produce una cantidad significativa de paquetes descartados, como se analiza en los resultados.

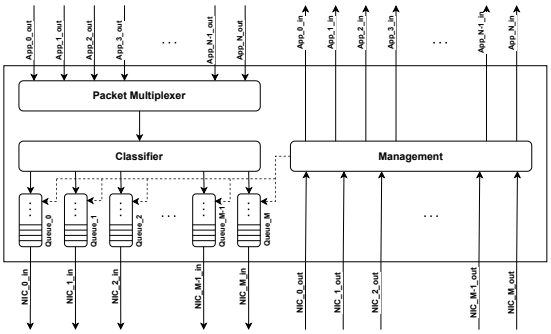


Fig. 2. Arquitectura del MCO

IV. SIMULACIÓN

En esta sección describimos y analizamos la implementación del simulador utilizado para la evaluación de los mecanismos del MCO sugeridos por el ETSI TR 103 439 [7].

A. IMPLEMENTACIÓN

La evaluación realizada en este trabajo utiliza el simulador OMNeT++ 6.1.0, junto con la biblioteca INET 4.5.4, que implementa el estándar IEEE 802.11. Esta biblioteca también incluye un modelo realista de propagación e interferencias para calcular la relación señal a interferencia más ruido (SINR, *Signal to Interference-plus-Noise Ratio*) y determinar las probabilidades de recepción de paquetes, considerando también el efecto de captura. Hemos ampliado esta biblioteca con una implementación del MCO. El código fuente está disponible en nuestro repositorio público[1].

La arquitectura considerada en nuestra implementación se muestra en la Fig. 2. En ella, varias aplicaciones generan tráfico que posteriormente es gestionado por el módulo MCO para ser asignado a los diferentes canales disponibles. Los vehículos tienen el mismo número de tarjetas de red que de canales disponibles, y la asignación de canales a las interfaces de red es estática. Cuando una aplicación genera un paquete, este se clasifica en una cola según el algoritmo del clasificador. Cada cola corresponde a una interfaz de red. El paquete esperará en esta cola hasta que la interfaz de red correspondiente esté disponible para transmitir, siendo la asignación a la cola mediada por el módulo `Management`, que también es responsable de enviar los paquetes recibidos a la aplicación destino correspondiente.

[1] https://github.com/Felipin22/MCO

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

B. CONFIGURACIÓN DE SIMULACIÓN

En nuestras simulaciones replicamos el escenario de [7]: una carretera recta de ocho kilómetros de longitud con seis carriles, cada uno de cuatro metros de ancho. La mitad de los vehículos circulan en una dirección, mientras que la otra mitad lo hace en la dirección opuesta. La posición inicial de los vehículos se distribuye uniformemente a lo largo de los carriles y de la longitud de la carretera. Cada vehículo se desplaza a una velocidad constante, seleccionada aleatoriamente a partir de una distribución gaussiana con una media determinada y una desviación estándar igual a una décima parte de dicha media. La media de la velocidad (130 km/h en nuestro escenario) depende de la densidad de vehículos, siguiendo el modelo de [7]. Cuando un vehículo llega al final de la carretera, reaparece al inicio del mismo carril. Todas las simulaciones se realizaron con 100 nodos, lo que en este escenario corresponde a una densidad de 12.5 vehículos/km.

A nivel de aplicación, el número de aplicaciones es un parámetro configurable que se varía en distintas ejecuciones de simulación para ajustar la carga del sistema. Cabe destacar que no utilizamos únicamente generación de tráfico periódica como en [7], sino una distribución más general que modela el hecho de que las aplicaciones representan diferentes servicios de la *Release 2* con distintos tipos de tráfico. Específicamente, en cada simulación hay siempre una aplicación que genera tráfico periódico, la cual simula un servicio de difusión como el CAS [7]. Dado que queremos evaluar el rendimiento de los mecanismos bajo condiciones de estrés, se configura con una tasa de generación de 50 paquetes/s, lo que correspondería a la tasa requerida por una aplicación de *Cooperative Adaptive Cruise Control* (CACC). Adicionalmente, se agregan aplicaciones cuya generación de paquetes sigue una distribución exponencial con $\lambda = 100$ paquetes/s. Esta distribución ha sido elegida para modelar aplicaciones o servicios más generales que se ejecutan en el vehículo, mientras que la tasa también está diseñada para forzar la carga de los canales. En todos los casos, los paquetes tienen un tamaño de 350 bytes.

Como en [7], adoptamos el modelo de propagación rural ECC68 modificado, pero, dado que el modelo de desvanecimiento utilizado en [7] no está suficientemente especificado, empleamos un modelo Nakagami-m sobre el modelo ECC68 rural.

La interferencia de un canal hacia uno adyacente se modela como se describe también en [7]. En particular, dada una transmisión desde un nodo genérico *j* en un canal, la interferencia percibida en un nodo genérico *i*, a una distancia d_{ji}, en un canal diferente, se calcula como:

$$P_{ji}^{MCO} = P_{ji}^{r} + L_{ji}^{adj}, \tag{1}$$

donde P_{ji}^{r} es la potencia recibida en dBm que depende únicamente del modelo de propagación y del desvanecimiento entre el nodo *j* y el nodo *i*, y L_{ji}^{adj} representa los efectos de la selectividad y el rechazo de canales adyacentes. En nuestras simulaciones, dado que consideramos la interferencia despreciable más allá del segundo canal adyacente y la descartamos, utilizamos 25.9 y 40.6 dB como los valores de L_{ji}^{adj} para el primer y segundo canal adyacente, respectivamente.

Tabla I
PARÁMETROS GENERALES

Parámetro	Valor
Tiempo de simulación	10 s
Número de vehículos	100
Potencia de ruido	-110 dBm
Nakagami shape factor	3
Pérdida por adyacencia del primer canal	-25.9 dB
Pérdida por adyacencia del segundo canal	-40.6 dB

Tabla II
PARÁMETROS VEHICULARES

Parámetro	Valor
Número de interfaces de red	7
Número de aplicaciones	1, 2, 3, 4, 5, 6
Tasa de paquetes de la 1ª aplicación (periódica)	50 paquetes/s
Tasa de paquetes de la 2ª aplicación	100 paquetes/s
Tamaño de paquete	350 bytes
Potencia de transmisión	23 dBm
Ganancia de antena	3 dBi
Tasa de transmisión	6 Mbps
Umbral de detección (D_T)	-98 dBm
Umbral SINR ($SINR_T$)	3.1 dB
Velocidad media	130 km/h

C. INDICADORES CLAVE DE RENDIMIENTO

Dentro de esta subsección, se describen brevemente los KPIs utilizados en nuestra evaluación. Además, la Tabla III presenta los valores de los parámetros KPI considerados en el funcionamiento de los mecanismos de uso de canal evaluados.

1) CHANNEL BUSY RATIO (CBR): Proporciona la fracción de tiempo en la que el canal no está inactivo y, por tanto, actúa como una estimación de su carga. Para un intervalo de tiempo de medición dado, T_m, indica la fracción de ese intervalo en la que la señal en un canal está por encima del umbral de detección, D_T. Se mide individualmente para cada vehículo y canal, y posteriormente se calcula el valor promedio entre todos los vehículos.

2) PACKET RECEPTION RATIO (PRR): Dado un valor máximo de distancia, d_m, dividido en un número de intervalos equiespaciados, d_i, para cada paquete transmitido por un vehículo, el PRR en cada intervalo se calcula dividiendo el número de vehículos que recibieron correctamente el paquete entre el total de vehículos dentro de ese intervalo. Este KPI también se mide individualmente para cada vehículo y canal, y posteriormente se calcula el valor promedio entre todos los vehículos. El PRR se considera habitualmente el principal indicador del rendimiento esperado y actúa como un indicador de los efectos de la interferencia entre canales adyacentes, que resulta difícil de medir directamente.

V. RESULTADOS

En esta sección presentamos los resultados de las simulaciones realizadas. Es importante destacar que todas

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla III
PARÁMETROS KPIS

Parámetro	Valor
PRR máxima distancia (d_m)	2000 m
PRR intervalo de distancia (d_i)	100 m
CBR intervalo de tiempo de medición (T_m)	100 ms
CBR límite con CC (C_{max})	0.62
CBR límite *sequential filling* sin CC (C_{max})	0.82

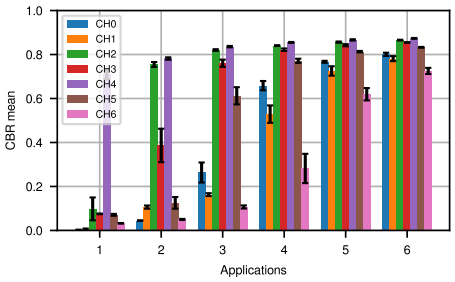


Fig. 3. CBR medio por número de aplicaciones - *Sequential Filling* sin control de congestión

las simulaciones se han replicado cinco veces. De este modo, evaluamos la incertidumbre del CBR y del PRR utilizando intervalos de confianza del 95%, los cuales se representan mediante barras de error y bandas sombreadas, respectivamente.

A. SEQUENTIAL FILLING

1) SIN CONTROL DE CONGESTIÓN: La Fig. 3 muestra el valor medio de CBR por número de aplicaciones. A primera vista, se puede observar que los canales con mayor prioridad definidos en la secuencia son los que presentan un valor de CBR más alto, como era de esperar. Asimismo, cabe destacar que cuando el número de aplicaciones es bajo, el valor medio del CBR solo es significativo en los primeros canales de la secuencia, ya que solo se transmite a través de ellos. El valor residual que aparece en los demás canales se debe a la interferencia entre canales adyacentes. Nótese cómo, incluso con una sola aplicación, la interferencia entre canales adyacentes da lugar a un CBR = 0.05. A medida que aumenta el número de aplicaciones, el valor medio del CBR en todos los canales tiende a igualarse. Para comprender mejor las mediciones de CBR, la Fig. 4 presenta un gráfico temporal del CBR para un nodo arbitrario con dos aplicaciones. El valor máximo de CBR que puede alcanzar un nodo se aproxima al umbral $C_{max} = 0.82$, que definimos como el límite a partir del cual se considera que un canal está completamente cargado y debe cambiarse a otro. Como era de esperar, el CBR medido supera este límite antes de cambiar de canal, y luego desciende cuando el nodo pasa al siguiente canal definido en la secuencia, repitiéndose este comportamiento en el tiempo.

A continuación, nos centramos en el PRR, que se muestra en la Fig. 5. Desde un principio, si nos fijamos en el escenario con un sola aplicación, se puede observar que la curva para el primer canal de la secuencia (CH4) es

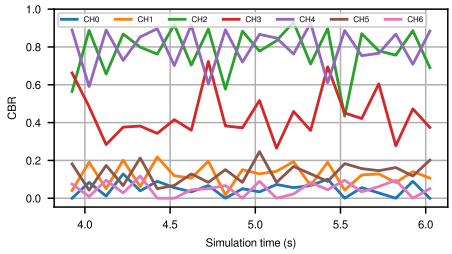


Fig. 4. CBR instantáneo para el nodo 0 - *Sequential Filling* sin control de congestión - 2 aplicaciones y 100 nodos

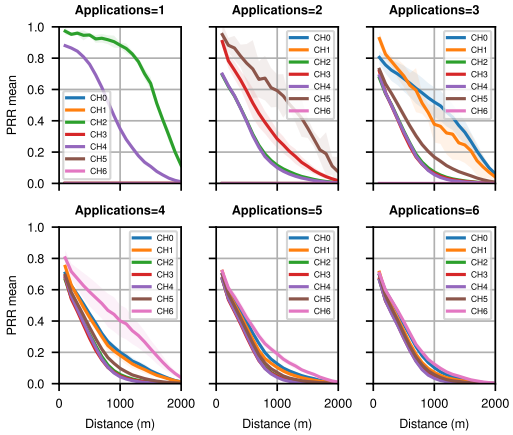


Fig. 5. PRR medio - *Sequential Filling* sin control de congestión

peor respecto a la del segundo canal de la secuencia (CH2), esto es así porque el primer canal se lleva a la saturación mientras que el segundo se encuentra más liberado. En el caso de dos aplicaciones, el efecto es aún peor, ya que tanto CH4 como CH2 se mantienen completamente cargados, lo cual es perjudicial para su PRR y, además, generan una cantidad considerable de interferencia en los canales adyacentes, afectando al siguiente canal en la secuencia, CH3, que no está suficientemente separado, como ya mencionamos, reduciendo también su PRR. Estos efectos se mantienen al aumentar el número de aplicaciones. Finalmente, el PRR muestra claramente el efecto de la interferencia entre canales adyacentes, ya que los canales extremos (CH0 y CH6) siempre presentan un mejor PRR cuando se utilizan todos los canales.

2) CON CONTROL DE CONGESTIÓN BÁSICO: En este caso, el comportamiento del CBR es muy similar al del enfoque sin control de congestión, como puede observarse en la Fig. 6. La única diferencia destacable es que, en este enfoque, con un límite de CBR más bajo, el cambio al siguiente canal en la secuencia definida se produce antes, lo que da lugar a que se midan valores promedio de CBR más altos en los canales posteriores. No obstante, los valores promedio de CBR también son muy similares a los del caso sin control de congestión, cuando lo esperable sería que el promedio se aproximara más al límite de CBR establecido de 0.62. Esto se debe

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

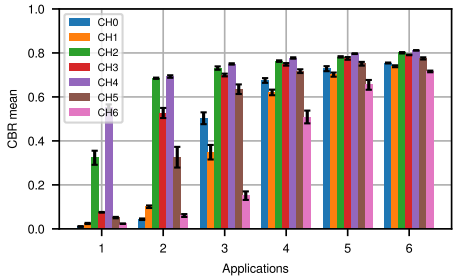


Fig. 6. CBR medio por número de aplicaciones - *Sequential Filling* con control de congestión

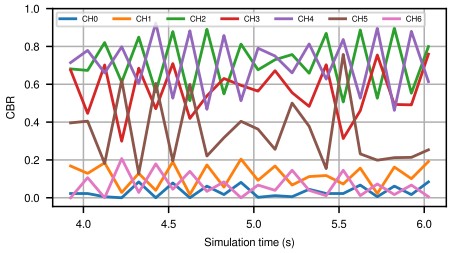


Fig. 7. CBR instantáneo para el nodo 0 - *Sequential Filling* con control de congestión - 2 aplicaciones y 100 nodos

principalmente a la naturaleza reactiva del mecanismo de *sequential filling*. Así, por ejemplo, cuando se toma una muestra con un valor inferior al límite establecido, incluso si este valor está muy próximo al límite, se interpretará que el canal no está saturado. Como consecuencia, el canal se utilizará para transmitir, *y se encolarán paquetes*, lo que provocará que en el siguiente intervalo de muestreo se mida un valor de CBR por encima del límite. Esta situación puede observarse en la Fig. 7, donde vemos que este comportamiento se repite a lo largo del tiempo, haciendo que finalmente el valor medio del CBR sea superior al límite establecido. A medida que se incrementa el número de aplicaciones, este comportamiento se acentúa, ya que se inyectan más paquetes durante los intervalos de muestreo del CBR.

Para el PRR, la Fig. 8 muestra que las curvas son muy similares al caso sin control de congestión, salvo que al establecer un C_{max} más bajo, cuando el número de aplicaciones es reducido, se utilizan más canales en la secuencia. Parece que, en última instancia, la posible mejora al reducir la carga en algunos canales se ve contrarrestada por el aumento de la interferencia entre canales adyacentes causada por la mayor carga en otros canales.

B. LOAD BALANCING

1) SIN CONTROL DE CONGESTIÓN: Con el mecanismo de *load balancing*, el valor medio de CBR en todos los canales es similar, siendo siempre más alto en los canales centrales, que experimentan una mayor interferencia de canal adyacente, como se puede observar en la Fig. 9. A diferencia de *sequential filling*, los nodos

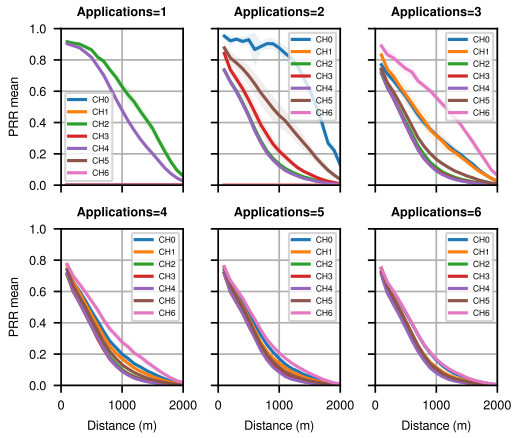


Fig. 8. PRR medio - *Sequential Filling* con control de congestión

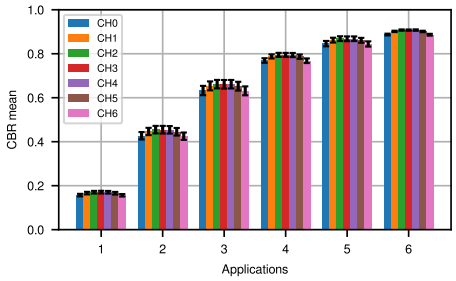


Fig. 9. CBR medio por número de aplicaciones - *Load Balancing* sin control de congestión

miden un valor de CBR estable para todos los canales durante prácticamente toda la simulación, sin oscilaciones, como se muestra para uno de los nodos en la Fig. 10, y dicho valor se asemeja mucho al promedio mostrado en la Fig. 9.

En cuanto al PRR, el paralelismo inherente de este mecanismo hace que las curvas de PRR sean mejores en los casos donde el número de aplicaciones es bajo, ya que los canales no están excesivamente saturados y la carga se distribuye uniformemente entre ellos. Se puede observar que, a medida que aumenta el número de aplicaciones, las curvas empeoran progresivamente, ya que en este caso no hay descarte de paquetes debido al control de congestión.

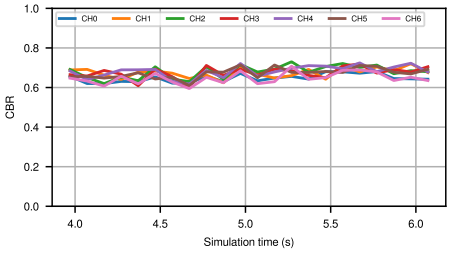


Fig. 10. CBR instantáneo para el nodo 0 - *Load Balancing* sin control de congestión - 3 aplicaciones y 100 nodos

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

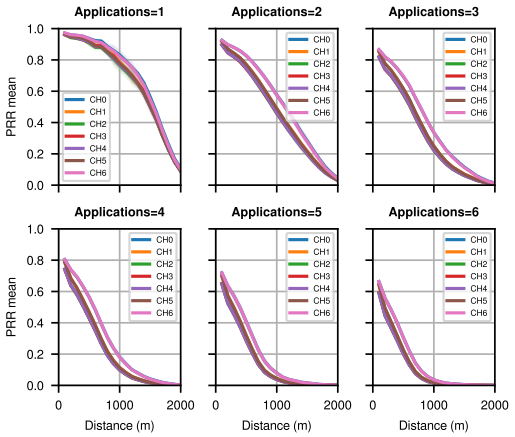


Fig. 11. PRR medio - *Load Balancing* sin control de congestión

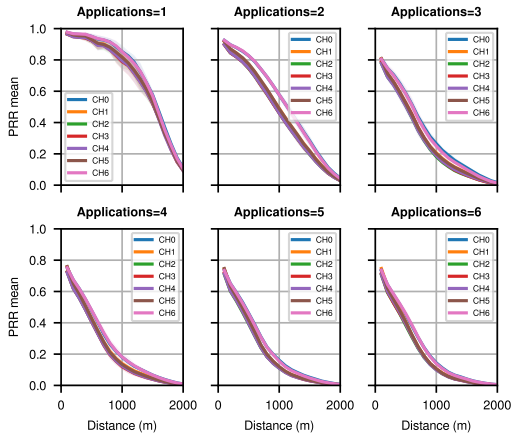


Fig. 13. PRR medio - *Load Balancing* con control de congestión

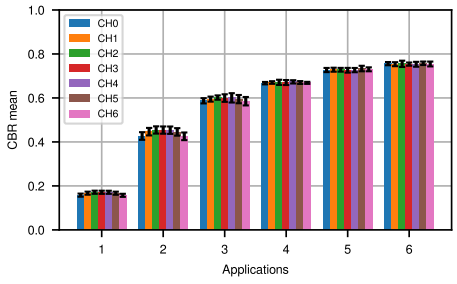


Fig. 12. CBR medio por número de aplicaciones - *Load Balancing* con control de congestión

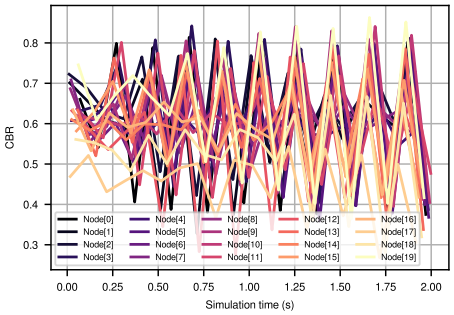


Fig. 14. CBR instantáneo en el canal 0 para varios nodos - *Load Balancing* con control de congestión - 3 aplicaciones y 100 nodos

Todo esto puede apreciarse en la Fig. 11. El rendimiento es claramente superior al de *sequential filling*. De hecho, el PRR de *load balancing* sin control de congestión es mejor que el de *sequential filling* con control de congestión hasta tres aplicaciones. A medida que aumenta el número de aplicaciones, ambos mecanismos se comportan de manera muy similar, ya que todos los canales están igualmente cargados, y el PRR de *sequential filling* resulta mejor en los canales menos cargados.

2) CON CONTROL DE CONGESTIÓN BÁSICO: Usar *load balancing* con control de congestión implica que, si todos los canales miden un CBR por encima de C_{max}, los paquetes generados son descartados por el MCO hasta que uno de los canales tenga un valor de CBR por debajo del límite. La Fig. 12 muestra el valor medio de CBR para este enfoque, donde se puede observar que el comportamiento es muy similar al anterior, es decir, el control de congestión en realidad no mantiene el CBR en el nivel de $C_{max} = 0.62$. En estos casos, se produce la misma situación que explicamos previamente para *sequential filling* con control de congestión, ya que ambos mecanismos son reactivos. La única diferencia notable respecto al enfoque sin control de congestión se observa cuando el número de aplicaciones es alto y se produce descarte de paquetes. En estos casos, el valor medio de CBR es algo menor y más uniforme.

La Fig.13 muestra las curvas de PRR cuando se utiliza control de congestión, y son muy similares a las del caso anterior cuando se utilizan menos de tres aplicaciones. Sin embargo, a partir de tres aplicaciones, sorprendentemente, el PRR es peor que el de *load balancing* sin control de congestión, especialmente a distancias cortas, mientras que solo es ligeramente mejor a distancias largas. Para entender la causa, debemos observar la evolución temporal del CBR para varios nodos, representada en la Fig.14, donde se puede ver que los vehículos utilizan los canales de forma sincronizada, es decir, los mismos canales se cargan y descargan aproximadamente al mismo tiempo en todos los vehículos. La razón es la siguiente: no hay muchos canales disponibles para seleccionar, y todos los vehículos que seleccionaron un canal inicial concreto miden simultáneamente un CBR por encima del límite y derivan el exceso de carga al mismo siguiente canal de la secuencia. Esos paquetes llenan la cola y provocan una sobrecarga de ese canal en el siguiente intervalo de medición, lo que a su vez hará que ese canal se descarte, reduciendo el CBR, pero produciendo el mismo efecto en el canal siguiente de la secuencia. El PRR es ligeramente mejor cuando el canal está infrautilizado y mucho peor cuando está sobrecargado. El resultado global es un PRR peor, excepto para los intervalos más lejanos.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

VI. DISCUSIÓN

El ETSI TR 103 439 [7] recomienda el uso de *sequential filling* para la difusión avanzada de aplicaciones de la *Release 2*. Sin embargo, nuestros resultados muestran un rendimiento bajo de este mecanismo, incluso en condiciones de baja carga de tráfico y baja densidad: los resultados del PRR en la Fig. 5 muestran que saturar un canal antes de cambiar a otro no mejora realmente el PRR *en ninguno de ellos*. Se podría argumentar que el canal no debería estar completamente saturado antes de hacer el cambio, pero, de hecho, eso es precisamente lo que se ha evaluado en nuestras simulaciones con control de congestión, y los resultados no son mucho mejores. La razón es que se trata de un mecanismo puramente reactivo: cuando el canal está cerca, pero por debajo del umbral C_{max}, los paquetes siguen encolándose y provocan que el CBR se dispare considerablemente por encima del límite en el siguiente intervalo, degradando todas las métricas de rendimiento.

Load balancing no es recomendado por [7] para difusión avanzada porque los vehículos podrían no estar utilizando el mismo número de interfaces de red (NICs). Esta es una preocupación válida, pero nuestras simulaciones muestran que, en realidad, es un mecanismo mejor cuando la carga no es especialmente alta. Sorprendentemente, cuando se utiliza control de congestión con *load balancing*, los resultados son incluso ligeramente peores. La razón vuelve a ser la naturaleza reactiva del control de congestión simple implementado, pero también su combinación con un proceso de sincronización entre vehículos que hemos comentado, y que provoca ciclos de carga y descarga simultánea de los canales.

A la luz de nuestros resultados, *load balancing* sin control de congestión sería una mejor opción que cualquiera de los otros mecanismos evaluados. Una posible solución si los vehículos no utilizan el mismo número de NICs es que estos negocien, mediante un procedimiento dedicado, el número máximo de NICs a utilizar. En cualquier caso, cuando la carga es alta, ninguno de los mecanismos proporciona un rendimiento aceptable, y los controles de congestión reactivos simples no mejoran la situación: el CBR oscila alrededor del valor deseado, que solo se alcanza como media, pero con grandes fluctuaciones, lo cual es típico de los procedimientos reactivos, como también se muestra en [10]. Se necesitan más investigaciones y mecanismos de control de congestión más sofisticados y no reactivos, como el que proponemos en [12], para poder soportar eficientemente escenarios de alta carga y de carácter general.

VII. CONCLUSIONES

En este trabajo, presentamos un entorno de simulación para evaluar mecanismos de operación en múltiples canales en redes vehiculares. Hemos llevado a cabo un amplio conjunto de simulaciones de dos mecanismos de uso de canales, *sequential filling* y *load balancing*, tanto en presencia de un control de congestión básico como sin él. En las simulaciones, cada nodo dispone de siete canales para transmisión y recepción, y se ha modelado la interferencia entre canales adyacentes. Los resultados revelan que, si bien estos mecanismos son sencillos de implementar, ninguno de ellos logra ofrecer un rendimiento aceptable, especialmente en condiciones de alta carga. Además, la introducción de un control de congestión reactivo simple no consigue aliviar el problema. Por tanto, es necesario proponer modificaciones más elaboradas a estos mecanismos que logren ajustar mejor su naturaleza reactiva cuando el tráfico generado es relativamente alto.

AGRADECIMIENTOS

Este trabajo ha sido financiado por las ayudas CNS2022-136150 y PID2023-148104OB-C42, concedidas por el MICIU/AEI/10.13039/501100011033 y por el FEDER/UE; y el proyecto GREENTBEAT BG-RRP-2.005-0002, financiado por la Unión Europea NextGenerationEU a través del Plan Nacional de Recuperación y Resiliencia de la República de Bulgaria.

REFERENCIAS

[1] ETSI, "Intelligent Transport Systems (ITS); Access Layer Specification for Intelligent Transport Systems Operating in the 5 GHz Frequency Band," ETSI, Standard EN 302 663 V1.3.1, 2020.

[2] IEEE, *IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. IEEE Std 802.11-2020, Feb. 2021.

[3] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions," ETSI, Technical Report TR 102 638 V1.1.1, 2009.

[4] ——, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specification of Decentralized Environmental Notification Basic Service," ETSI, Standard EN 302 637-3 V1.2.1, 2014.

[5] Car 2 Car Communication Consortium, "Multi-Channel Operation (MCO); Part 1: Functional Requirements," Car 2 Car Communication Consortium, Tech. Rep., 2020.

[6] ——, "Multi-Channel Operation (MCO); Part 3: MCO Concept," Car 2 Car Communication Consortium, Tech. Rep., 2022.

[7] ETSI, "Intelligent Transport Systems (ITS); Multi-Channel Operation Study; Release 2," ETSI, Technical Report TR 103 439 V2.1.1, 2021.

[8] A. Bazzi *et al.*, "Multi-Channel Operation for the Release 2 of ETSI Cooperative Intelligent Transport Systems," *IEEE Communications Standards Magazine*, vol. 8, no. 1, pp. 28–35, 2024.

[9] K. Garlichs, A. Willecke, A.-C. Hagau, and L. C. Wolf, "Maneuver Coordination for CAVs: ITS-G5 Channel and Service Performance Evaluation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 14 210–14 224, 2023.

[10] O. Vászi and L. Bokor, "Modeling and Evaluation of a Dynamic Channel Selection Framework for Multi-Channel Operation in ITS-G5," *Telecom*, vol. 4, no. 2, pp. 313–333, 2023.

[11] ETSI, "Intelligent Transport Systems (ITS); Decentralized Congestion Control Mechanisms for Intelligent Transport Systems Operating in the 5 GHz Range; Access Layer Part," ETSI, Technical Specification TS 102 687, 2011.

[12] E. Egea-Lopez and P. Pavon-Mariño, "Adjacent Channel Interference and Congestion Control for Multi-Channel Operation in Vehicular Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 1, pp. 1–15, 2025.

[13] M. Boban and A. Festag, "Service-Actuated Multi-Channel Operation for Vehicular Communications," *Computer Communications*, vol. 93, pp. 17–26, 2016.

[14] M. Sepulcre and J. Gozalvez, "Heterogeneous V2V Communications in Multi-Link and Multi-RAT Vehicular Networks," *IEEE Transactions on Mobile Computing*, vol. 20, no. 1, pp. 162–173, 2021.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Comparación de técnicas de calibración para el sensor O_3 del módulo ZPHS01B de calidad del aire en Nodos IoT

E. Meneses-Albala*, S. Felici-Castell*, JJ. Perez-Solano*, G. Montalban-Faet*, J. Segura-Garcia*, P. Salva-Garcia†, G. Mas‡
* ETSE, Universitat de València,† University Western of Scotland (Reino Unido),‡ BSG Ingenieros S.L.
emeal2@uv.es, felici@uv.es, jjperezs@uv.es, guimonfa@uv.es, jsegura@uv.es,
pablo.salva-garcia@uws.ac.uk, gemma@bsg.es

En los entornos urbanos, la calidad del aire es un indicador relevante dado que está vinculado con diferentes patologías y problemas de salud, y se recomienda mejorar su monitorización con mayor densidad espacial/temporal. Dado el elevado coste de las estaciones de monitorización de calidad del aire, los sensores de bajo coste combinados con Internet de las Cosas (*Internet of Things*, IoT) son una excelente alternativa, pero sus prestaciones son limitadas en precisión y fiabilidad. La integración de dichos sensores en nodos equipados con comunicaciones inalámbricas (o Nodos IoT) y el uso de técnicas tanto estadísticas como basadas en inteligencia artificial, pueden mejorar sus prestaciones. En este artículo, analizamos las características de estos nodos, equipados con el multisensor ZPHS01B, y comparamos diferentes técnicas de calibración para ozono terrestre aplicando estas técnicas mencionadas. De todas las técnicas evaluadas, los métodos basados en aprendizaje automático (algoritmos *ensemble*), en particular el modelo *Gradient Boosting* (GB), ofrece los mejores resultados, con errores de precisión menores del 10%.

***Palabras Clave*—calibración, calidad del aire, sensores de bajo coste, regresión, aprendizaje automático, redes neuronales**

I. INTRODUCCIÓN

La calidad del aire es un factor crucial para la salud pública en las ciudades, ya que la contaminación atmosférica está estrechamente vinculada con enfermedades respiratorias, cardiovasculares y diversos tipos de cáncer [1]. Según la Organización Mundial de la Salud (OMS), el 99% de la población mundial está expuesta a niveles de contaminación del aire que superan las directrices de seguridad recomendadas [2]. Para prevenir estos problemas, la OMS proporciona directrices que incluyen niveles máximos permisibles de contaminantes para proteger la salud pública [3].

A nivel europeo, las Directivas 2008/50/CE [4] y recientemente la Directiva 2024/2881/CE [5] del Parlamento Europeo, establecen el marco legal para evaluar y gestionar la calidad del aire ambiente definiendo los límites legales para diversos contaminantes del aire (centrándose fundamentalmente en $PM10$, $PM2.5$, NO_2, SO_2, CO, O_3), y obliga a los Estados miembros a realizar un seguimiento continuo de la calidad del aire mediante estaciones de monitoreo distribuidas estratégicamente. En la Directiva 2008/50/CE, se fijan los criterios para la ubicación y densidad mínima de los puntos de muestreo, en función de la población y la superficie de las zonas o aglomeraciones. En particular esta normativa establece la obligación de instalar estaciones de monitoreo (*Air Quality Monitoring Stations*, (AQMS)) en zonas urbanas con al menos un punto de muestreo por cada 2 millones de habitantes o uno por cada 50,000 km^2, el que resulte en un mayor número de puntos de muestreo.

En este contexto, los sensores de bajo coste para calidad del aire son una alternativa interesante para mejorar la resolución espacio/temporal de la calidad del aire. No obstante, estos sensores, por sí solos, no ofrecen una buena precisión [6], [7], [8], y aún menos para ser considerados como una referencia. Entre los principales problemas que presentan son: a) baja precisión y exactitud especialmente en condiciones ambientales variables, b) alta variabilidad entre sensores de mismo tipo que pueden registrar valores distintos bajo las mismas condiciones, c) sensibilidad cruzada, con dificultades para distinguir entre diferentes tipos de gases o partículas, pudiendo reaccionar a gases interferentes y d) deriva con el tiempo por el desgaste de los sensores [7], [9].

Sin embargo, el uso de tecnologías como la Inteligencia Artificial (*Artificial Intelligence*, (AI)), mediante técnicas de aprendizaje automático (*Machine Learning*, (ML)) y redes neuronales (*Neural Networks*, (NN)) y/o con apren-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I: Características de los sensores integrados en el módulo ZPHS01B: contaminante detectado (unidades), rango y resolución

Contaminante [unidad]	Rango	Resolución
$PM2.5$ [μg/m^3]	0-1000	$\pm$15 <100, 15% >100
O_3 [ppm]	0-10	$\pm$0.01
CO [ppm]	0-500	$\pm$0.1
NO_2 [ppm]	0.1-10	$\pm$0.05
CO_2 [ppm]	0-5000	$\pm$500ppm
CH_2O [mg/m^3]	0-6.250	$\pm$0.03 <0.2, $\pm$20% >0.2
TVOC [4 levels]	0-3	
T [º]	-20, 65	$\pm$0.5º
HR [%]	0-100	$\pm$3%

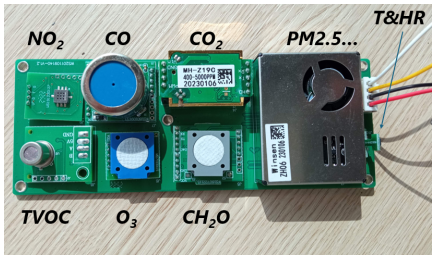
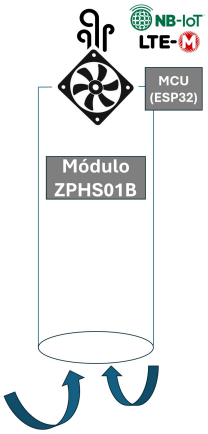


Fig. 1: Detalle de los sensores del módulo ZPHS01B para monitorización de la calidad del aire.



(a) Diseño del Nodo

(b) Emplazamiento

Fig. 2: Diseño del Nodo IoT de calidad del aire basado en MCU (ESP32) con módulo ZPHS01B y emplazamiento encima de estación de calidad del aire.

dizaje profundo (*Deep Learning*, (DL)), junto con modelos de regresión lineal (*Linear Regression*, (LR)), combinados con comunicaciones inalámbricas desarrolladas para estos escenarios, en particular en el contexto de Internet de las Cosas (*Internet of Things*, (IoT)), abren nuevas oportunidades para desplegar este tipo de redes con soporte a la monitorización de calidad del aire.

Un ejemplo de caso de uso en un escenario de alta contaminación, se describe en [10], detallando el diseño de un Nodo IoT para calidad del aire, que integra el módulo multisensor ZPHS01B [11] de Winsen como se muestra en la Figura 1, que integra 9 sensores diferentes: Temperatura (T, en °C), Humedad Relativa (HR, en %), así como CO, CO_2, NO_2 y O_3 medidos en partes por millón (ppm), formaldehído (CH_2O) medido en mg/m^3, material particulado (*Particulate Matter*, (PM) medido en μg/m^3, y compuestos orgánicos volátiles totales ($TVOC$), cuya concentración se clasifica en 4 niveles (muy bajo, bajo, intermedio y alto). La Tabla I resume toda esta información y las características principales de dichos sensores. Este módulo se integra en dicho Nodo IoT, basado en el MicroControlador (MCU) ESP32 y para soporte de comunicaciones inalámbricas, usa el modem SIM7600 para transmisión en *Narrow Band* (NB)/IoT. Dicho Nodo IoT, su diseño y encapsulado para monitorizar la calidad del aire, se muestra en la Figura 2.

En este artículo, basándonos con las experiencias previas con este Nodo IoT, evaluamos métodos para la calibración del sensor de ozono (O_3), comparando métodos estadísticos (también conocidos como *cajas blancas*) y métodos basados en AI (*cajas grises*) [12]. Los resultados obtenidos, apuntan a que los métodos basados en ML para este tipo de sensor integrado en este módulo, son los que mejor funcionan. Para evaluar el ajuste de los modelos utilizamos, por un lado el coeficiente de determinación (R^2), así como el Error Cuadrático Medio (*Root Mean Square Error*, (RMSE)) y Error Medio Absoluto (*Mean Absolute Error*, (MAE)).

El resto del trabajo se estructura en las siguientes secciones. En la Sección II, mostramos el trabajo relacionado. En la Sección III, describimos la recogida de datos. En la Sección IV y V, analizamos tanto las técnicas basadas en AI como estadísticas. Finalmente, en la Sección VI, resumimos las principales conclusiones y trabajo futuro.

II. ESTADO DEL ARTE

El módulo ZHPS01B ha sido utilizado previamente en otros trabajos. En [13], se muestra la implementación de un dispositivo para la evaluación de la calidad del aire en exteriores, sin calibración, con el fin de mapear los contaminantes en una zona urbana. En [14], este módulo se emplea para para monitorización en interiores, calculando un índice de calidad del aire. En [15], los autores explican brevemente el uso de una red neuronal para determinar (clasificar) tipos de aire: con o sin contaminación. Además, en [16], se presenta un prototipo para medir la calidad del aire utilizando este módulo en un dron.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

El proceso de calibración de estos sensores de bajo coste es un reto, en el que se pueden utilizar tanto modelos basados en AI (ML y DL) como estadísticos. En [17], los autores muestran modelos de calibración (utilizando datos de 16 semanas) para mejorar el rendimiento del sensor, destacando que el enfoque *Random Forest* (RF), es más robusto ya que tiene en cuenta las sensibilidades cruzadas de los contaminantes y consiguen un Error Medio Relativo (*Mean Relative Error*, MRE) de 15% para O_3. En [18], se lleva a cabo la calibración de un sensor de aerosoles para $PM2,5$ comparando modelos de regresión lineal simple con *Gradient Boosting* (GB), utilizando el sensor de partículas de PPD42. El estudio concluye que GB se comporta mejor y mejora significativamente el rendimiento de los sensores, alcanzando un R^2 de hasta 0,76. Asimismo, en [9], los autores utilizan redes neuronales dinámicas para la calibración logrando modelos con R^2 (MAE) (en ppb) de 0,69 (7,45), con un MRE del 49%.

Finalmente, en [19], basándose en sensores de O_3 y NO_2, junto con T y HR, los autores analizaron diferentes opciones de calibración tanto basados en AI como estadísticos. Concluyen que multiple RF alcanza la mayor precisión durante la Fase I (pre-despliegue), con un R^2 de 0,98 y MAE (MRE) de 4,31 (5,74%), en un escenario con una media de O_3 de 75 $\mu g/m^3$. Y en la Fase II (validación de campo) realizada en una ubicación diferente, el rendimiento empeora, con el MAE (MRE) 22,22 (29,62%) mientras que la regresión lineal múltiple (*Multiple Linear Regression*, (MLR) ofrece 12,96 (17,28%). En este caso, MLR proporciona mejores resultados. Los autores concluyen que MLR pueda ser una solución más adecuada para representar modelos físicos más allá del conjunto de datos de calibración de la Fase I, demostrando una mejor transferibilidad a través de diversos entornos espaciales y temporales, destacando que los modelos paramétricos como MLR tienen una ecuación definida con sólo unos pocos parámetros, lo que los hace más fáciles de ajustar a los posibles cambios en el tiempo.

Destacar que en los trabajos [17], [9] y [19], utilizan sensores de O_3, en particular el sensor Alphasense Ox-B431, Alphasense B4 Electrochemical y MiCS-2614 respectivamente. Estos sensores tienen un precio mucho mayor (aproximadamente 10 veces) que el propio módulo ZPHS01B, que dispone simultáneamente de 9 sensores diferentes.

III. RECOGIDA DE DATOS CON EL NODO DE CALIDAD DEL AIRE.

Para la generación del *datasets*, hemos realizado un despliegue con dos Nodos IoT diferentes (Nodo 1 y 2) junto la estación AQMS, llamada Molí del Sol (Valencia, España) situada en 39.48113875, -0.40855865, gestionada por la Generalitat Valenciana (GVA) y cuyos datos se recuperan de https://rvvcca.gva.es/estatico/46250048. El *dataset* recoge datos desde el 14 de Marzo de 2024 hasta el 25 de Marzo de 2025, un total de 376 días. Esta estación de calidad del aire recoge información de SO_2, NO_2, NO, CO, O_3, $PM1$, $PM2.5$ y $PM10$. Tomamos como valores de referencia, los de la estación que se dan en $\mu g/m^3$ promediando periódicamente cada 10 minutos. La conversión de unidades de O_3 en ppm (del sensor) a $\mu g/m^3$ (de la estación), es 1 ppm=1960 $\mu g/m^3$. Destacar que las lecturas del módulo ZPHS01B en el Nodo IoT se toman a razón de 10 muestras por minuto, una muestra cada 6 segundos.

Una ejemplo de la captura de los valores de ozono de estos nodos frente la estación, lo podemos ver en la Figura 3, del 13 al 20 de Junio del 2024, donde observamos un patrón repetitivo y las mismas tendencias de la AQMS. Este patrón repetitivo es debido al ciclo de formación del ozono con la radiación solar y reacción fotoquímica de otros contaminantes (óxidos de nitrógeno y azufre de la combustión de motores térmicos). Durante la campaña de medida, el valor medio del ozono es de 55,72 $\mu g/m^3$, valores relativamente bajos (asociados a baja contaminación) comparados con los umbrales de la OMS que fija en 100 $\mu g/m^3$ [3]. Además se observa en la Figura 3, que el sensor de O_3 del módulo ZPHS01B tiene un límite mínimo de detección de 0,02 ppm, lo que corresponde aproximadamente a 40 $\mu g/m^3$, frente el valor mínimo registrado por la AQMS de 4 $\mu g/m^3$. Por tanto, estos sensores de ozono no tiene suficiente resolución para detectar concentraciones por debajo de este umbral.

IV. EVALUACIÓN DE TÉCNICAS BASADAS EN INTELIGENCIA ARTIFICIAL.

En este sección evaluamos diferentes técnicas para mejorar las medidas de los sensores de bajo coste, basados en inteligencia artificial: aprendizaje automático (ML) y redes neuronales (NN). Estos métodos requieren de *datasets* para poder realizar las fases de entrenamiento así como una fase previa de exploración de dichos datos.

A. PREPARACIÓN DE LOS DATA-SETS.

En la secciones siguientes, el proceso de aprendizaje se realiza a través de la división del *dataset* en tres partes, entrenamiento (*training*), validación y testeo (*testing*). Probamos diferentes porcentajes de *training-testing* de estos conjuntos de datos: 60%-40%, 70%-30%, 80%-20% y 90%-10%. Cuando dividimos el conjunto de datos para entrenamiento y testeo, ambos conjuntos permanecen independientes y aislados. Sin embargo, durante el proceso de entrenamiento propiamente dicho, el conjunto de datos se divide a su vez en dos partes: una para entrenamiento y otra para la validación. Por defecto, destinamos el 80% de los datos al entrenamiento y el 20% a la validación. En este proceso, los conjuntos de datos de entrenamiento y validación se combinan en diferentes iteraciones.

Comentar que las métricas de rendimiento mostradas en la siguientes secciones son la media ponderada de cada métrica a lo largo de 100 iteraciones, cambiando el contenido del conjunto de *training-testing* para obtener resultados con el mínimo sesgo posible.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

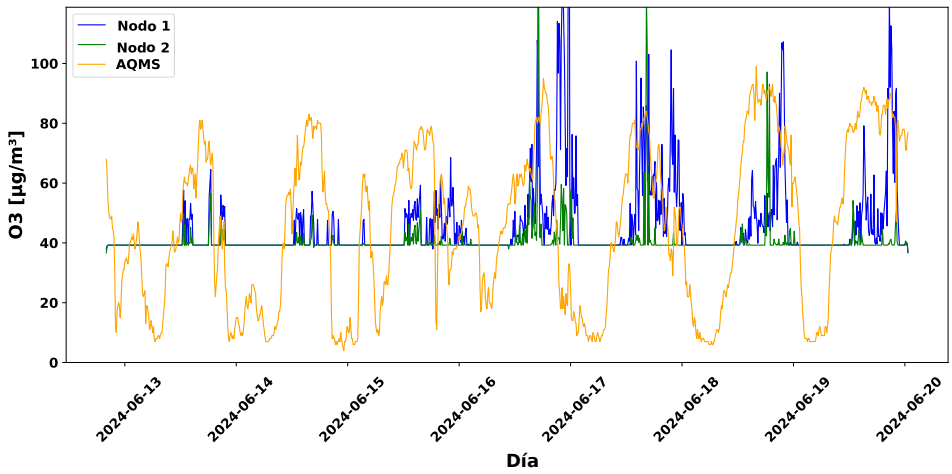


Fig. 3: Medidas del sensor de O_3 de los Nodos IoT 1 y 2, frente la estación AQMS en $\mu g/m^3$

B. ANÁLISIS EXPLORATORIO DE DATOS

Las características de diseño y fabricación de los propios sensores de bajo coste del módulo ZPHS01B, repercuten en la calidad de las medidas como hemos indicado previamente [7], [9], siendo la interferencia de ciertos contaminantes un factor que a tener en cuenta. Esta información, puede ser extraída analizando la correlación entre los diferentes sensores.

Además, el análisis de importancia de características es una técnica para identificar qué variables de entrada (características o *features*) tienen mayor influencia en las predicciones de un modelo. Este análisis puede ser independiente del modelo, y evalúa el impacto de cada característica en el rendimiento del modelo sin importar el algoritmo utilizado. Este análisis mejora la interpretabilidad del modelo, guía la selección de características y apoya la toma de decisiones basada en datos [12].

A partir del *dataset*, realizamos estos análisis sobre las características (variables) que proporcionan los mejores resultados, concluyendo que para O_3 son `[date, O3, T, HR]`. En estos modelos, la `date` (fecha) se incluye como metadato para tener en cuenta el efecto del envejecimiento y mejorar los modelos que siguen el patrón de tráfico rodado relacionado con los patrones del ozono.

Destacar que a menor número de características a utilizar en los modelos, se obtienen mejores resultados. Es decir, si añadimos más características que no son tan significativas, se empobrece el conjunto de datos.

A continuación evaluamos los diferentes modelos.

C. MODELOS BASADOS EN REDES NEURONALES (NN).

Para ello, usaremos un modelo basado en una red neuronal secuencial simple de TensorFlow/Keras, utilizando un optimizador de descenso de gradiente estocástico, con una entrada de 4 características y dos capas. Estas dos capas son: una capa densa con cuatro neuronas y activación lineal, seguida de una segunda capa con una sola neurona

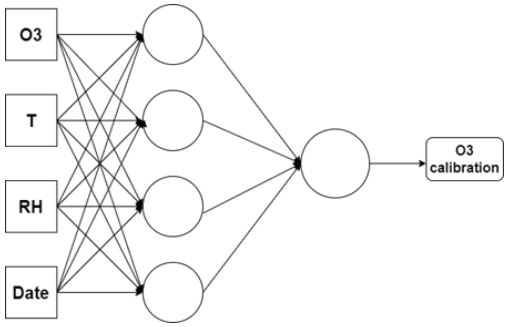


Fig. 4: Red neuronal secuencial simple de Tensor Flow/Keras con una entrada de 4 características y dos capas para calibración de O_3.

que proporciona la salida. El esquema de la red se muestra en la Figura 4.

Los resultados de esta red neuronal secuencial son: coeficiente de determinación R^2=0,9999, $RSME$=3,5144e-05 y MAE=2,9925e-05 para los diferentes combinaciones de *training-testing*. Como podemos observar, estos resultados muestran que estas técnicas aprenden y memorizan el *dataset* completo (sobre-ajuste) y no podemos generalizar.

D. MODELOS BASADOS EN APRENDIZAJE AUTOMÁTICO (ML).

En esta sección, nos centramos en los modelos *Decision Tree* (DT) y los algoritmos *ensemble*: *Random Forest* (RF), *Adaptive Boosting* (ADA) y *Gradient Boosting* (GB), que utilizan árboles de decisión como base y están implementados en las librerías *scikit-learn*. La capacidad de estos modelos para combinar múltiples árboles y generar predicciones más estables y precisas los hace ideales para calibraciones precisas y generalizables, incluso en condiciones ambientales variables.

DT es un modelo simple y fácil de interpretar, pero propenso al sobreajuste. RF mejora esto al combinar mu-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

chos árboles entrenados de forma independiente mediante técnicas de muestreo aleatorio, lo que reduce la varianza. ADA y GB, en cambio, aplican técnicas de impulso (*boosting*), entrenando árboles de forma secuencial. ADA ajusta el peso de las muestras mal clasificadas, mientras que GB corrige los errores del modelo minimizando la función de pérdida. De forma resumida, estos modelos buscan mejorar la precisión del modelo combinando múltiples árboles de manera diferente.

Estos algoritmos son especialmente interesantes porque permiten corregir las imprecisiones y no linealidades propias de estos sensores. Además estos modelos, al ser robustos frente al ruido y a la multicolinealidad, pueden trabajar con datos imperfectos y variables correlacionadas.

Como paso previo, para estos modelos debemos optimizar las configuraciones internas (o hiperparámetros), utilizando para ello *GridSearch*, que muestran mejores resultados en R^2, RMSE y MAE. En particular estos hiperparámetros son según modelo:

- RF: `número de estimadores`: 100, `Max. depth`: ninguno, `Max. features`: 1
- GB: `número de estimadores`: 900, `Max. depth`: ninguno, `Max. features`: 1, `Learning rate`: 0,1, `Subsample`: 1 `Loss`: error cuadrático.
- ADA: `número de estimadores`: 50, `Learning rate`: 0,01, `Loss`: exponencial.
- DT: `Max. depth`: ninguno, `Max. features`: 1, `Splitter`: best.

Modelo	*Nodo*	R^2	**RMSE**	**MAE**
GB	1	0,94	6,11	4,34
	2	0,95	5,33	3,84
RF	1	0,93	6,65	4,54
	2	0,94	5,80	3,91
ADA	1	0,92	6,80	3,94
	2	0,94	5,93	3,40
DT	1	0,87	8,92	5,24
	2	0,88	7,79	4,52

Tabla II: Evaluación de métricas (R^2 y RMSE, MAE en $\mu g/m^3$) de los modelos ML (GB, RF, ADA, DT) con *dataset* 90% y 10% para entrenamiento (*training*) y testeo

En la Tabla II, mostramos los resultados de la evaluación de las métricas (R^2, RMSE y MAE) de los modelos ML con *dataset* 90% y 10% para entrenamiento (*training*) y testeo. En particular podemos observar que el modelo que presenta mejores resultados es GB con un ajuste R^2 de al menos 0.94 y con un MAE entre 3,84 y 4,34 en $\mu g/m^3$, seguido de RF, ADA y DT. Observamos que teniendo en cuenta el valor medio del ozono durante la campaña, tal como se comentó en Sección III, de 55,72 $\mu g/m^3$, el error de predicción es menor de 5,57$\mu g/m^3$, un error relativo medio menor del 10%. En la Figura 5, podemos ver el resultado del ajuste con el modelo de GB frente al valor de ozono de referencia, utilizando modelo GB optimizado e inicial entre 16:00 y 20:00 horas, 16 Junio 2024.

Durante el proceso de entrenamiento (*training*), podemos observar en la Figura 6, la convergencia de R^2, que proporciona información sobre el sobre-ajuste, considerando tanto el conjunto de datos de entrenamiento como el de validación a través de diferentes iteraciones, con un mejor ajuste en el entrenamiento (*training*) que con la validación. Destacar que el proceso de convergencia con entrenamiento no alcanza un ajuste perfecto en ningún caso, lo que justifica y apoya inicialmente la conclusión de que no existe sobre-ajuste en los modelos. Además, vemos que los valores de R^2, tanto para el entrenamiento como para la validación son mejores que los valores mostrados para el testeo en la Tabla II, debido a que el conjunto de datos de testeo no participa en el proceso de entrenamiento.

V. EVALUACIÓN DE TÉCNICAS BASADAS EN MODELOS ESTADÍSTICOS: REGRESIONES LINEALES.

Ventana	*Nodo*	R^2	**RMSE**	**MAE**
24	1	0,68	26,35	21,42
	2	0,67	26,96	22,18
12	1	0,56	66,96	39,64
	2	0,54	53,43	38,40
8	1	0,51	46,71	35,97
	2	0,49	47,59	35,86
4	1	0,44	37,08	25,60
	2	0,42	37,15	24,94
1	1	0,40	22,97	12,67
	2	0,49	27,80	14,00

Tabla III: Evaluación de métricas (R^2 y RMSE, MAE en $\mu g/m^3$) del modelo de regresión lineal para calibración de O_3 con diferentes tamaños de ventana: 1, 4, 8, 12 y 24 horas

Ventana	*Nodo*	R^2	**RMSE**	**MAE**
24	1	0,62	28,55	24,43
	2	0,60	27,36	25,24
12	1	0,52	72,36	43,62
	2	0,50	63,33	48,40
8	1	0,47	52,71	55,97
	2	0,43	57,69	55,86
4	1	0,42	47,08	26,60
	2	0,42	32,07	24,94
1	1	0,49	24,37	22,67
	2	0,44	24,50	15,44

Tabla IV: Evaluación de métricas (R^2 y RMSE, MAE en $\mu g/m^3$) del modelo de regresión cuadrática para calibración de O_3 con diferentes tamaños de ventana: 1, 4, 8, 12 y 24 horas

Finalmente, planteamos métodos estadísticos para ajuste y calibración basados en regresión lineal. Para ello, hemos diseñado un proceso de calibración que se realiza automáticamente cada 24 horas, ya que los nodos descargan todos los días los valores diezminutales de la estación AQMS y a partir de ellos, ajustan la calibración para las siguientes 24 horas. Cada nodo ajusta los modelos de regresión por separado.

Para ello, aplicamos modelos de regresión simples, con diferentes tamaños de ventana para el ajuste, en particular tanto lineal como cuadrático. En concreto, hemos comparado diferentes tamaños de ventana de 1, 4, 8, 12 y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

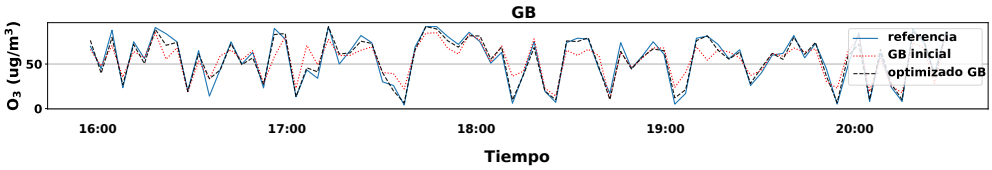
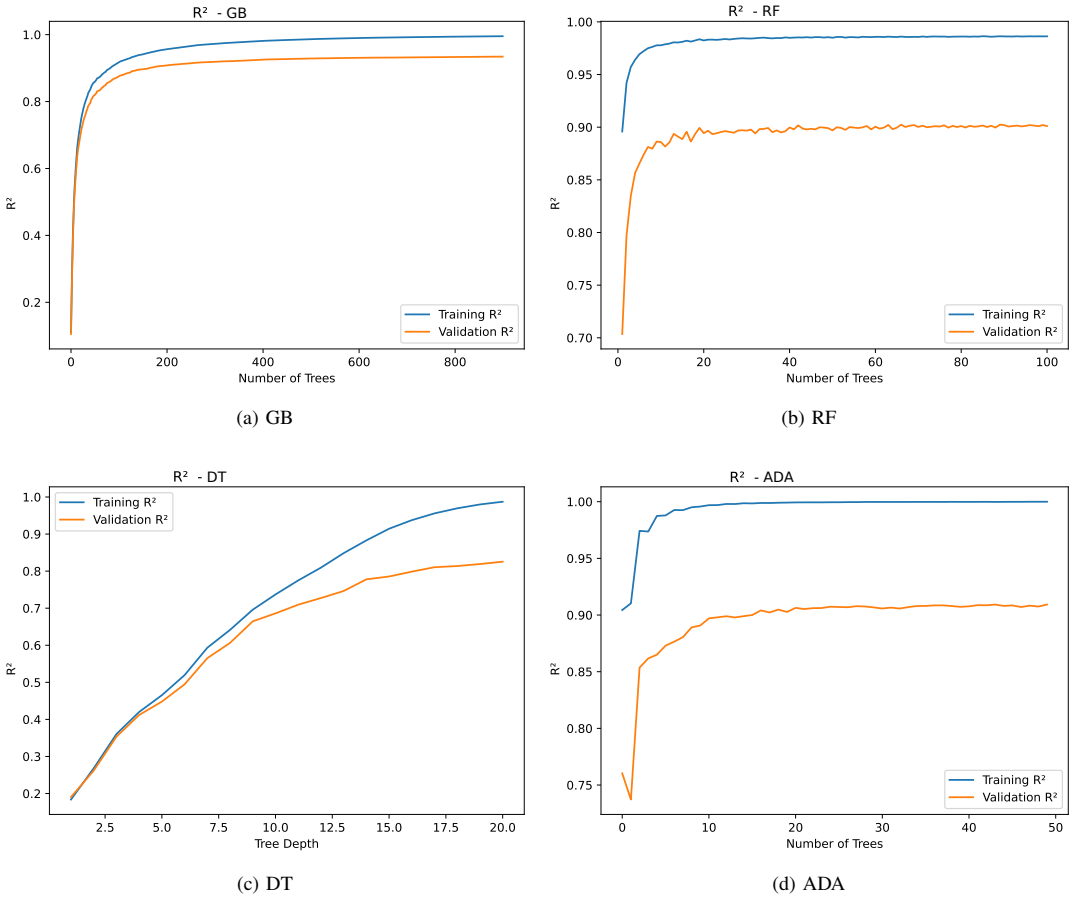


Fig. 5: Ajuste del modelo de GB para calibración de O_3 en $\mu g/m^3$, con modelo GB optimizado e inicial entre 16:00 y 20:00 horas, 16 Junio 2024.



(a) GB

(b) RF

(c) DT

(d) ADA

Fig. 6: Convergencia de coeficiente de determinación R^2 con diferentes iteraciones para entrenamiento (*training*) y validación para los modelos ML (GB, RF, ADA y DT) con sus principales hiperparámetros.

24 horas, y presentamos los resultados de estos ajustes analizando tanto RMSE, MAE y R^2, como métricas de rendimiento, como se muestra en las Tablas III y IV para cada uno de los nodos. Cabe destacar que los mejores resultados se obtienen con regresión lineal (Tabla III), comparados con regresión cuadrática (Tablas IV) con una ventana de 24 horas. Observamos que estos resultados son peores que los mostrados con los modelos ML, mostrados en la Tabla II. Además, se observan diferencias entre los Nodo 1 y 2, debido a su proceso de fabricación de bajo coste.

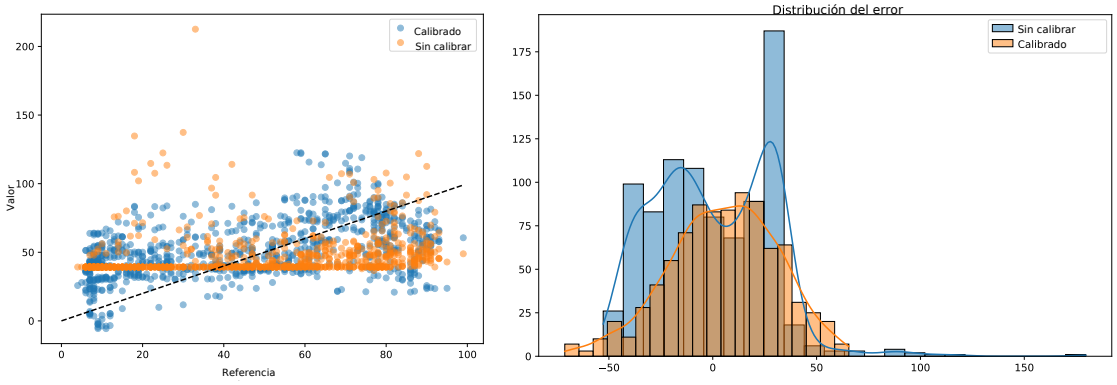
En la Figura 7, mostramos los puntos de la regresión en el ajuste lineal (7a), junto con la distribución del error (7b) y el resultado del ajuste con ventana de 24 horas del ozono de referencia frente la versión calibrada y sin calibrar con ajuste lineal (7c).

VI. CONCLUSIONES

La calidad del aire se va afectada en las ciudades por los motores térmicos, donde los contaminantes generados por la combustión pueden ser extremadamente peligrosos a partir de determinadas concentraciones y deben observarse con mayor detalle y resolución, donde los sensores de bajo coste desempeñan un papel crucial.

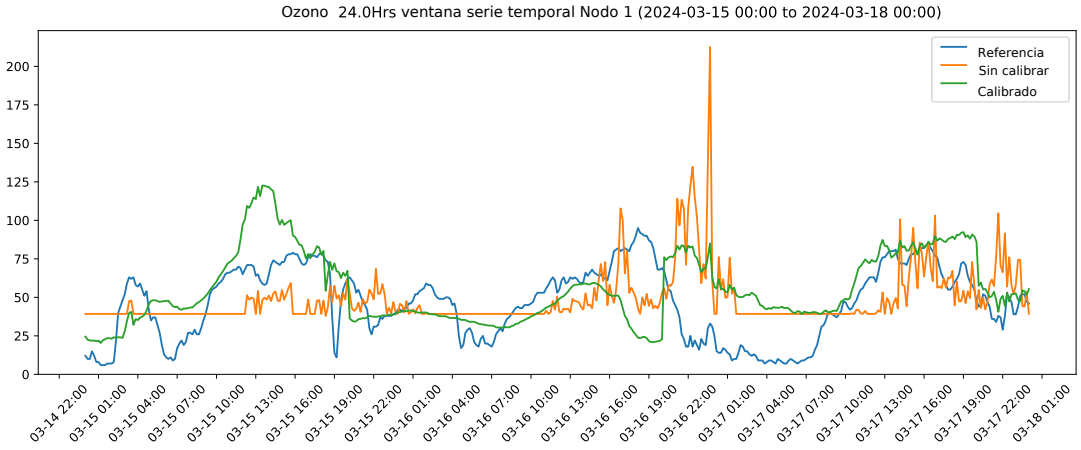
Para ello, hemos presentado un Nodo IoT basado en el módulo multisensor ZPHS01B con capacidad de sensar

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) Puntos de regresión

(b) Distribución de error



(c) Serie temporal

Fig. 7: Puntos de regresión lineal (7a), junto con la distribución del error (7b) y visualización temporal del ozono de referencia frente la versión calibrada y sin calibrar con ajuste lineal (7c).

los contaminantes atmosféricos de calidad del aire a bajo coste, en particular el ozono terrestre, pero con limitaciones, debido fundamentalmente a la falta de precisión y sensibilidad, deriva e interferencias cruzadas.

Para superar estas limitaciones, se han comparado diferentes técnicas. Observamos que los modelos basados en aprendizaje automático (ML) y algoritmos *ensemble* permiten obtener los mejores resultados, con un error de precisión menor del 10% para una atmósfera poco contaminada, con valores medios de ozono de 55,72 $\mu g/m^3$, lo que limita el rango de operación del sensor. En particular el modelo *Gradient Boosting* (GB) consigue los mejores resultados, seguido de RF, ADA y DT. Estos modelos pueden considerar tanto las interferencias cruzadas, como las no-linearidades de los sensores. Los métodos basados en modelos estadísticos y regresiones lineales, son peores porque tiene limitaciones para estimar valores del sensor en zonas de poca sensibilidad, en rangos inferiores a 40 $\mu g/m^3$. Finalmente los modelos basados en redes neuronales, dadas las características del *dataset* (derivadas del sensor de bajo coste para medición de ozono en un escenario (poco contaminado)), generan sobre-ajuste y no permiten su generalización.

Actualmente estamos trabajando con estos modelos (y diferentes mejoras) así como modelos alternativos (XGBoost y LSTM por citar algunos), aplicando otras técnicas, con un análisis de exploración de datos más exhaustivo, utilizando *datasets* de mayor duración (más de un año, para recoger las diferentes estaciones), teniendo en cuenta la limitación de envejecimiento de estos sensores.

AGRADECIMIENTOS

El artículo ha sido parcialmente financiado por el proyecto PID2021-126823OB- I00 por MCIN/AEI/ 10.13039/ 501100011033 y por European Union NextGeneration EU/ PRTR; por Generalitat Valenciana con ayudas CIAICO/2022/ 179, CIACIF/2023/ 416 y CIAEST/2024/ 71 así como el Ministerio de Educación en las becas de movilidad PRX23/00589.

REFERENCIAS

[1] H. Adair-Rohani, "Air pollution responsible for 6.7 million deaths every year," https://www.who.int/teams/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants, 2023, accessed: 25/1/2025.

[2] P. Orellano, J. Reynoso, N. Quaranta, A. Bardach, and A. Ciapponi, "Short-term exposure to particulate matter (pm10 and pm2.5), nitrogen dioxide (no2), and ozone (o3) and all-cause and cause-specific mortality: Systematic review and meta-analysis," *Environment International*, vol. 142, p. 105876, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160412020318316

[3] WHO agency, "WHO global air quality guidelines," https://apps.who.int/iris/bitstream/handle/10665/345329/9789240034228-eng.pdf, 2023, accessed: 15/03/2023.

[4] DIRECTIVE 2008/50/EC, "OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 May 2008 on ambient air quality and cleaner air for Europe." *Official Journal of the European Communities*, vol. L 152, pp. 1–44, 2008.

[5] The European Parliament and the Council of the European Union, "Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe (recast)," Official Journal of the European Union, L Series, Nov. 2024, oJ L, 20.11.2024, pp. 1–70. ELI: http://data.europa.eu/eli/dir/2024/2881/oj. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024L2881

[6] e. a. C. Borrego, "Assessment of air quality microsensors versus reference methods: The eunetair joint exercise," *Atmospheric Environment*, vol. 147, pp. 246–263, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1352231016307610

[7] F. Karagulian, M. Barbiere, A. Kotsev, L. Spinelle, M. Gerboles, F. Lagler, N. Redon, S. Crunaire, and A. Borowiak, "Review of the performance of low-cost sensors for air quality monitoring," *Atmosphere*, vol. 10, no. 9, 2019. [Online]. Available: https://www.mdpi.com/2073-4433/10/9/506

[8] M. R. García, A. Spinazzé, P. T. Branco, F. Borghi, G. Villena, A. Cattaneo, A. D. Gilio, V. G. Mihucz, E. G. Álvarez, S. I. Lopes, B. Bergmans, C. Orłowski, K. Karatzas, G. Marques, J. Saffell, and S. I. Sousa, "Review of low-cost sensors for indoor air quality: Features and applications," *Applied Spectroscopy Reviews*, vol. 57, no. 9-10, pp. 747–779, 2022. [Online]. Available: https://doi.org/10.1080/05704928.2022.2085734

[9] E. Esposito, S. De Vito, M. Salvato, V. Bright, R. Jones, and O. Popoola, "Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems," *Sensors and Actuators B: Chemical*, vol. 231, pp. 701–713, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092540051630332X

[10] E. Meneses-Albala, G. Montalban-Faet, S. Felici-Castell, J. J. Perez-Solano, and R. Fayos-Jordan, "Assessment of a multisensor zphs01b-based low-cost air quality monitoring system: Case study," *Electronics*, vol. 14, no. 8, 2025. [Online]. Available: https://www.mdpi.com/2079-9292/14/8/1531

[11] L. Zhengzhou Winsen Electronics Technology Co., "Multi-in-one sensor module (model: Zphs01b) manual," https://www.winsen-sensor.com/d/files/zphs01b-english-version1_1-20200713.pdf, Zhengzhou Winsen Electronics Technology Co., LTD, 07 2024, [Accessed 2/6/2025].

[12] J.-J. Zhu, M. Yang, and Z. J. Ren, "Machine learning in environmental research: Common pitfalls and best practices," *Environmental Science & Technology*, vol. 57, no. 46, pp. 17 671–17 689, 2023, pMID: 37384597. [Online]. Available: https://doi.org/10.1021/acs.est.3c00026

[13] H. Coto-Fuentes, F. Valdés-Perezgasga, K. Guevara-Amatón, K. Limones-Ríos, and C. Calderón-Ibarra, "Integración de estaciones KNARIO con un sistema de información geográfico para el monitoreo de la calidad del aire en la zona metropolitana de La Laguna," *Revista Ciencia, Ingeniería y Desarrollo*, vol. 1, no. 8, pp. 109–114, 2022.

[14] S. Vaheed, P. Nayak, P. S. Rajput, T. U. Snehit, Y. S. Kiran, and L. Kumar, "Building iot-assisted indoor air quality pollution monitoring system," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 2022, pp. 484–489.

[15] A. Antonenko, V. Boretskij, and O. Zagaria, "Classification of Indoor Air Pollution Using Low-cost Sensors by Machine Learning," in *EGU General Assembly Conference Abstracts*, ser. EGU General Assembly Conference Abstracts, May 2023, pp. EGU–14 856.

[16] Z. Kennedy, D. Huber, H. R. Xie, J. E. Sohl, J. Page, and W. Dowell, "Miniature Multi-Sensor Array (mini-MSA) for Ground-to-Stratosphere Air Measurement, Phase II," *Mechanical Engineering Commons*, 10 2021. [Online]. Available: https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1600&context=spacegrant

[17] N. Zimmerman, A. A. Presto, S. P. N. Kumar, J. Gu, A. Hauryliuk, E. S. Robinson, A. L. Robinson, and R. Subramanian, "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring," *Atmospheric Measurement Techniques*, vol. 11, no. 1, pp. 291–313, 2018. [Online]. Available: https://amt.copernicus.org/articles/11/291/2018/

[18] N. E. Johnson, B. Bonczak, and C. E. Kontokosta, "Using a gradient boosting model to improve the performance of low-cost aerosol monitors in a dense, heterogeneous urban environment," *Atmospheric Environment*, vol. 184, pp. 9–16, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1352231018302504

[19] A. Cavaliere, L. Brilli, B. P. Andreini, F. Carotenuto, B. Gioli, T. Giordano, M. Stefanelli, C. Vagnoli, A. Zaldei, and G. Gualtieri, "Development of low-cost air quality stations for next-generation monitoring networks: calibration and validation of no_2 and o_3 sensors," *Atmospheric Measurement Techniques*, vol. 16, no. 20, pp. 4723–4740, 2023. [Online]. Available: https://amt.copernicus.org/articles/16/4723/2023/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Simulador web mulsemedia en 3D

Fernando Boronat, Erika Villashagñay, Lluc Simó y Juan González
Departamento Comunicaciones,
Universitat Politècncia de València – Campus de Gandia
IIM R&D Group – Calle Paraninfo, 1, 46730, Grao de Gandia (Valencia-España)
fboronat@dcom.upv.es, {evillas, lsimmar, juagons4}@upv.es

Abstract. La comunidad investigadora lleva años trabajando en el desarrollo de sistemas *mulsemedia*, es decir, sistemas multimedia que incluyen efectos sensoriales asociados a los contenidos AV, que son capaces de estimular sentidos de los usuarios más allá de los tradicionales visual y auditivo. Como ejemplo de efectos se pueden citar aromas (olfato), sabores (gusto), vibraciones, presión, viento (tacto), iluminación especial, temperatura, humo, vaporizaciones, etc. Existen soluciones y estándares (p.ej., *MPEG-V*) que posibilitan la integración de dispositivos reales generadores de efectos en los sistemas multimedia. Sin embargo, disponer de escenarios físicos reales completos con múltiples dispositivos físicos reales con el fin de probar su correcto funcionamiento es costoso y poco flexible. Un método alternativo más rápido y económico consiste en el uso de simuladores. En este artículo se presenta un simulador 3D basado en web, compatible con el estándar *MPEG-V*, que permite comprobar dicho funcionamiento de forma visual.

***Palabras Clave- MPEG-V*, *Mulsemedia*, Simulación, Simulación *Mulsemedia*, Simulador de Efectos Sensoriales**

I. Introducción

Los sistemas multimedia tradicionales normalmente incluyen contenidos audiovisuales (AV) que estimulan únicamente los sentidos de la vista y el oído. Sin embargo, la comunidad científica y la industria llevan años trabajando en el desarrollo de sistemas multimedia que incluyen metadatos de efectos sensoriales (SEM, *Sensory Effects Metadata*), asociados a los contenidos AV, siendo capaces de generar dichos efectos, estimulando todos los sentidos de los usuarios. Como ejemplo de efectos, se pueden citar los de aromas (olfato), sabores (gusto), vibraciones, presión, efectos de viento (tacto), iluminación especial, temperatura, humedad, humo, vaporizaciones, etc. El objetivo de estos sistemas es estimular otros sentidos más allá de la vista y el oído y, por tanto, conseguir proporcionar experiencias de usuario más realistas e inmersivas ("*ver y escuchar es creer, pero sentir y experimentar es la realidad*" [1]). Para ofrecer a los usuarios nuevas sensaciones explorando otros sentidos más allá de los dos tradicionales en aplicaciones multimedia, se ha propuesto un nuevo concepto conocido como ***Mulsemedia*** (*Multiple Sensorial Media*). Dicho concepto está tomando mucha fuerza en multitud de áreas [2], tales como telecomunicaciones, educación [3], comercio electrónico, publicidad, entretenimiento (p.ej., en cines 4D), turismo [4], salud, integración social o la TV inmersiva [5]. A través de la inclusión de efectos sensoriales (en adelante, contenido *mulsemedia*), se intenta proporcionar experiencias más realistas e inmersivas a los usuarios [6]. En el pasado, se ha demostrado que dicha práctica mejora la Calidad de la Experiencia (QoE, *Quality of Experience*) de los usuarios [7] [8]. Actualmente, está creciendo el número de empresas que comercializan dispositivos generadores de efectos sensoriales (también denominados actuadores), como, p. ej., Sensory Acumen, Inc.[1], Olorama Technology[2] (aromas) o bHaptics[3] (vibración). Además, en algunos parques temáticos (ej, Portaventura[4]) y cines 4D[5] ya se están incorporando los efectos sensoriales a las experiencias de los usuarios.

El término *mulsemedia*, apareció por primera vez en 2014 en [1]. Desde entonces, la investigación en este campo ha avanzado significativamente, incluso apareciendo estándares relacionados (p. ej., *MPEG-V*[6] [9]) y arquitecturas de Internet de las cosas (IoT [10]) que permiten integrar dispositivos reales generadores de efectos en los sistemas multimedia. En [11] se presentó un exhaustivo estudio sobre sistemas *mulsemedia*, destacando la poca exploración del campo *mulsemedia*, los desafíos y su aplicabilidad en las áreas mencionadas. En [6] se recopilan muchos de los avances más destacados en el área hasta el año 2020. En [12] se comprobó el efecto enmascarador que los contenidos multisensoriales proporcionan al contenido tradicional. Uno de los efectos con el que más se ha experimentado en el pasado es el de aromas. En [13] se demuestra su relevancia para provocar y evocar emociones, sentimientos, cambios de actitud y recuerdos, y se enfatiza la necesidad de sistemas que los puedan recrear. En [14], se proponen una serie de recomendaciones con el fin de orientar en el diseño y la evaluación de sistemas de consumo de contenido *mulsemedia*, además de describir algunos retos clave para integrar dichos efectos satisfactoriamente, como la integración de contenido *mulsemedia* basada en aromas, la

[1] https://sensoryacumen.com/ (últ. acceso: septiembre 2025)
[2] https://www.olorama.com/ (últ. acceso: septiembre 2025)
[3] https://www.bhaptics.com/ (últ. acceso: septiembre 2025)
[4] https://www.portaventuraworld.com (últ. acceso: septiembre 2025)
[5] https://www.oceanografic.org/actividad/cine-4d/ o https://www.hero-ncity.com/valencia/heron-city-paterna/4dx (últ. acc.: septiembre 2025)
[6] https://www.mpeg.org/standards/MPEG-V/ (últ. acc.: septiembre 2025)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

sincronización, estandarización, desarrollo de generadores de aromas, intensidad y duración de los efectos, aplicabilidad en diferentes áreas (salud, educación, turismo...) o su distribución remota. En [15], se presentan los antecedentes, la historia y las tecnologías esenciales de codificación, decodificación y comunicaciones que sustentan este campo emergente, centrándose en las aplicaciones de la Realidad eXtendida (XR) y las comunicaciones holográficas. Por otro lado, en cuanto a dispositivos generadores de efectos, en [16] se presenta un recopilatorio de dispositivos existentes y una guía para construir entornos *mulsemedia*, tanto para escenarios de escritorio como para escenarios inmersivos.

A pesar de la existencia de muchos trabajos sobre *mulsemedia* y de la existencia del estándar *MPEG-V*, aún quedan muchos temas por resolver: cómo se señalizan y cómo se integran los efectos sensoriales en los flujos multimedia en la cadena de distribución, el control efectivo de los efectos, la sincronización con los contenidos AV, las posibles combinaciones entre varios efectos, su integración en entornos domésticos, etc. Además, el número de efectos sensoriales a generar y sincronizar puede ser elevado. Por otra parte, la inclusión de contenido *mulsemedia* junto con contenido AV omnidireccional (p. ej., vídeo 360) o de RV presenta retos adicionales, como, p. ej., la dependencia de los efectos sensoriales con respecto al campo de visionado (FoV, *Field of View*).

En la Fig. 1 se muestra un ejemplo de la infraestructura necesaria para disponer de un entorno local *mulsemedia*. Sin embargo, montar un escenario completo de este tipo, con dispositivos físicos reales colocados en diversas posiciones alrededor del usuario, con el único fin de comprobar el correcto funcionamiento de la experiencia *mulsemedia* diseñada, y de la correspondiente generación de efectos en los instantes adecuados es costoso y poco flexible. Un método más rápido y económico consiste en el uso de simuladores que permitan comprobar la corrección de los metadatos *mulsemedia* generados asociados al contenido AV. Dichos metadatos deberán seguir un formato concreto y contener información precisa sobre los efectos sensoriales a des/activar desde determinadas posiciones y en instantes específicos durante la reproducción. Los simuladores suelen incluir elementos gráficos representando a generadores de efectos (o actuadores) virtuales. De este modo, cuando se deba activar un actuador de un efecto sensorial concreto, se resalta su representación gráfica para indicar que en dicho instante se debería estar generando dicho efecto y debería estar siendo percibido por los usuarios del sistema *mulsemedia*.

En este artículo se presenta una aplicación de simulación en 3D de la integración de múltiples efectos sensoriales en una experiencia de consumo multimedia, con la finalidad indicada anteriormente. Está basada en tecnologías web y es compatible con el estándar *MPEG-V*. Incluye un reproductor de contenidos AV almacenados (en un servidor de contenidos) y un entorno web de simulación en 3D. La simulación permite comprobar si la des/activación de efectos sensoriales se realiza en los instantes oportunos (es decir, de forma sincronizada con la reproducción de los contenidos AV) y desde las posiciones correctas. Sin embargo, no permite evaluar la idoneidad, ni la calidad, ni si las características de los efectos son las adecuadas (intensidad, realismo…) para el contenido AV mostrado, para lo cual se necesita un entorno mulsemedia real.

Fig. 1. Ejemplo de entorno local *mulsemedia*

Para poder utilizar la herramienta con dicha finalidad, se ha incluido en la misma un módulo de comunicaciones para poder conectarse y enviar mensajes, compatibles con el estándar *MPEG-V*, de des/activación de efectos, a un dispositivo denominado "*Controlador Mulsemedia*" (ver Fig. 1), si existe. Dicho dispositivo será el encargado de realizar las funciones de gestión de dispositivos físicos reales para la correcta generación de los efectos sensoriales existentes en el entorno local *mulsemedia*.

Las principales contribuciones de este artículo son las siguientes:

- App web de simulación *mulsemedia* en 3D, incluyendo reproductor de contenidos AV e interpretación de metadatos SEM compatibles con el estándar *MPEG-V*.
- Diseño y desarrollo de entorno web 3D de simulación.
- Emulación de la des/activación de hasta 8 tipos de efectos sensoriales en tiempo real, desde todas las posiciones contempladas en *MPEG-V*.
- Módulo de comunicaciones integrado para la comunicación con dispositivo *Controlador Mulsemedia* mediante mensajes compatibles con *MPEG-V*.

La estructura del artículo es la siguiente: en la sección 2 se presenta un resumen de trabajos relacionados previos; en la sección 3 se describe la aplicación desarrollada, su arquitectura, el entorno 3D desarrollado y su interface gráfica. Finalmente, se presentan unas breves conclusiones.

II. Trabajos Relacionados

Para crear experiencias *mulsemedia*, es necesario generar previamente metadatos SEM para un contenido AV y asociarlos. Así, ambos pueden ser distribuidos y utilizados posteriormente en aplicaciones y reproductores que puedan interpretarlos. Existen varias herramientas que permiten crear metadatos SEM siguiendo lo especificado en la parte 3 de *MPEG-V*. En esta sección, se resume lo que incluye dicha parte, se presentan dos herramientas creadas con este objetivo, SEVino [7] y *STEVE* [17], y, finalmente, se presentan otras herramientas de simulación *mulsemedia* y sus limitaciones.

II.A. Parte 3 del estándar MPEG-V

MPEG-V [9] proporciona herramientas para describir mundos reales y virtuales. Especifica un formato para intercambiar datos entre el mundo real (escenario físico donde se ejecuta la aplicación *mulsemedia* y es percibida por los usuarios) y el mundo virtual (la propia aplicación). Entre otras muchas cosas, ofrece una gran cantidad de herramientas para describir y representar efectos y dispositivos sensoriales. El

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

estándar consta de 7 partes. La parte 1 comprende la descripción de su arquitectura. La parte 2 presenta los metadatos que pueden utilizarse para caracterizar dispositivos generadores de efectos sensoriales y sensores, y también define metadatos para describir las preferencias del usuario. La parte 3 presenta herramientas para describir efectos sensoriales, así como la sintaxis y la semántica de los efectos. La parte 4 especifica características de objetos del mundo virtual (p. ej., avatares). La parte 5 define los metadatos que pueden usarse para transmitir órdenes de des/activación a dispositivos generadores, así como la información proveniente de sensores. La parte 6 especifica herramientas y tipos de datos que se utilizan en diferentes partes del estándar. Finalmente, la parte 7 presenta software de referencia y conformidad con el estándar. Este subapartado se centra en la parte 3, que define las estructuras básicas de información sensorial para poder generar metadatos SEM. Dichas estructuras consisten en los bloques de construcción y los atributos comunes (p. ej., el tipo de efecto, la duración, el desvanecimiento o *fading*, la intensidad, etc.) que se utilizan en todos los efectos. También define un lenguaje de descripción de efectos sensoriales (SEDL, *Sensory Effect Description Language*), basado en XML, para describir las estructuras. Los efectos en sí se deben describir utilizando el denominado vocabulario de efectos sensoriales (SEV, *Sensory Effect Vocabulary*), que permite especificar todos los efectos sensoriales en detalle, sus atributos generales (información de des/activación, duración, prioridad y posición) y, además, según el efecto, sus atributos específicos (p. ej., color y aromas concretos, para los efectos de luces y olfativos, respectivamente). Según el estándar, se deben utilizar el SEDL y el SEV, de forma conjunta, para crear descripciones SEM que puedan ser entendidas e interpretadas por el denominado MPE (*Media Processing Engine*), un decodificador o un ordenador capaz de analizar dichas descripciones y controlar dispositivos generadores de efectos compatibles. Para describir los comandos del generador de un efecto, y la información detectada por sensores, *MPEG-V* proporciona el lenguaje de descripción de interface de interacción (IIDL, *Interaction Interface Description Language*). También define el lenguaje de descripción de información de control (CIDL, *Control Information Description Language*) para ejecutar los comandos de los generadores de efectos, de des/activación, según las preferencias del usuario. El CIDL también permite especificar las capacidades de los generadores/sensores. En lo que respecta al modelo espacial de *MPEG-V*, los efectos sensoriales tienen un atributo común denominado ubicación (*location*), que especifica la región desde la que un usuario debe percibir un efecto sensorial. En el estándar se definen 3 planos (frontal, medio y trasero), 3 alturas en cada plano (baja, media y alta) y 5 posiciones en cada altura (izquierda, centro-izquierda, centro, centro-derecha y derecha). De esta manera, se distinguen un total de 45 posiciones (Fig. 2). Cada posición se representa mediante la concatenación de tres palabras. Por ejemplo, la posición '*right:bottom:front*' indica que el efecto se des/activará en el plano frontal, en la parte baja (suelo) y a la derecha del espacio 3D alrededor del usuario.

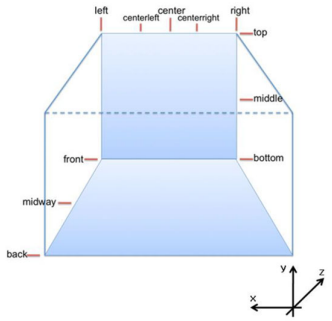


Fig. 2. Modelo de localización y sistema de coordenadas en *MPEG-V*[18]

II.B. Generación de metadatos de efectos sensoriales (SEM)

Existen varias herramientas que permiten la generación de metadatos SEM, como, *SEVino2* (*Sensory Effect Video Annotation tool* [7], *STEVE* (*Spatio-Temporal View Editor* [17]), *RoSE Studio* (*Representation of Sensory Effects* [19]) , *SMURF* (*Sensible Media aUthoRing Factory* [20]), H-*Studio* [21], o *Real 4D Studio* [22]. A excepción de *STEVE*, todas ellas se basan en el paradigma de sincronización temporal basado en líneas de tiempo y permiten a los usuarios definir, de forma gráfica, metadatos SEM para contenido AV y guardarlos en ficheros. En [23] se presenta una revisión y comparación más detallada de las herramientas de autoría *mulsemedia* mencionadas, así como también de varias propuestas para la representación de efectos sensoriales y sus características. Sin embargo, la mayoría están obsoletas y/o ya no están disponibles. Aquí se describen dos herramientas compatibles *MPEG-V*: *SEVino2* y *STEVE 2.0*.

a) Herramienta SEVino2 [7]. Está escrita en *Java* y utiliza *VLCJ*[7]. Su interface gráfica se muestra en la Fig. 3. Permite generar, a partir de contenido AV, metadatos SEM en ficheros con formato XML, compatibles con *MPEG-V*. Sólo permite incluir 7 efectos (viento, vibración, luces, temperatura, rociador o difusor de agua, aromas y niebla) para ser generados junto con la reproducción de contenidos AV. Tras seleccionar un contenido AV permite definir, de forma muy rápida e intuitiva, los instantes de des/activación de los efectos mencionados a lo largo de la reproducción del contenido AV. Para cada efecto se pueden definir varios parámetros (p.ej., duración, desvanecimiento o *fading*, prioridad, localización, valor y rango de intensidad del efecto), tal y como muestra la Fig. 3. Incluye un reproductor de contenidos (SEMP, *Sensory Effect Media Player*) y un simulador (SESim, *Sensory Effect Simulator*) para comprobar la corrección de los metadatos SEM generados. En la Fig. 4 se muestra un ejemplo de fichero XML con los metadatos SEM generados, que puede modificarse de forma manual posteriormente. *SEVino2* tiene varias limitaciones. Por un lado, no permite definir la activación de un mismo tipo de efecto sensorial en un mismo instante de tiempo desde diferentes localizaciones. Todos los efectos que se activan en un mismo instante los agrupa dentro de la etiqueta 'grupos de efectos' (*GroupOfEffects,* Fig. 4) con una única marca de tiempo de activación común a todos. Ello implica la necesidad de realizar un procesado posterior por parte de las aplicaciones para poder procesar los datos de cada efecto de forma individual.

[7] Java Framework for the VLC Media Player. https://github.com/caprica/vlcj (últ. acceso: septiembre 2025)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

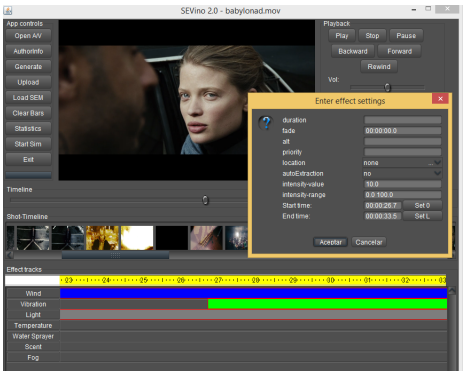


Fig. 3. Pantalla principal de *SEVino2*. Definición de un efecto sensorial

```
<?xml version="1.0" encoding="UTF-8"?>
<ns2:SEM xmlns:ns2="urn:mpeg:mpeg-v:2010:01-SEDL-NS" xmlns:ns1="urn:mpeg:mpeg7:schema:2004"
xmlns:ns3="urn:mpeg:mpeg-v:2010:01-SEV-NS" xmlns:ns4="urn:mpeg:mpeg-v:2010:01-CT-NS" xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance" xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
si:timeScale="90000" si:puMode="ancestorsDescendants" si:absTimeScheme="mp7t" xsi:schemaLocation=
"urn:mpeg:mpeg21:2003:01-DIA-XSI-NS
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG-21_schema_files/dia-2nd/XSI-2nd.xsd
urn:mpeg:mpeg-v:2010:01-SEV-NS
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG-V_schema_files/MPEG-V-SEV.xsd
urn:mpeg:mpeg7:schema:2004
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG-7_schema_files/mpeg7-v2.xsd">
<ns2:Effect xsi:type="ns3:LightType" autoExtraction="both" activate="true" si:anchorElement="true"
si:pts="0"/>
<ns2:Effect xsi:type="ns3:WindType" intensity-value="0.0" intensity-range="0.0 100.0" activate=
"true" si:anchorElement="true" si:pts="927000"/>
<ns2:Effect xsi:type="ns3:WindType" intensity-value="60.0" intensity-range="0.0 100.0" activate=
"true" fade="90000" si:anchorElement="true" si:pts="936000"/>
<ns2:Effect xsi:type="ns3:WindType" activate="false" si:anchorElement="true" si:pts="1062000"/>
<ns2:Effect xsi:type="ns3:WindType" intensity-value="5.0" intensity-range="0.0 100.0" activate=
"true" si:anchorElement="true" si:pts="1107000"/>
<ns2:Effect xsi:type="ns3:WindType" intensity-value="70.0" intensity-range="0.0 100.0" activate=
"true" fade="1890000" si:anchorElement="true" si:pts="1116000"/>
<ns2:GroupOfEffects si:anchorElement="true" si:pts="2403000"><ns2:Effect xsi:type=
"ns3:TemperatureType" activate="true"/>
<ns2:Effect xsi:type="ns3:VibrationType" intensity-value="10.0" intensity-range="0.0 100.0"
activate="true"/>
</ns2:GroupOfEffects><ns2:Effect xsi:type="ns3:TemperatureType" activate="false" si:anchorElement=
"true" si:pts="2628000"/>
```

Fig. 4. Ejemplo de fichero XML con metadatos (SEM) generados por la herramienta *SEVino2*

b) Herramienta STEVE [17]. Permite crear presentaciones o aplicaciones interactivas multimedia para sistemas web y de TV digital de una manera sencilla y amigable. Su modelo de sincronización está basado en un modelo propio denominado SIMM (*Simple Interactive Multimedia Model*). Utiliza un paradigma de sincronización temporal basado en eventos. Permite a los autores la edición de vistas espaciotemporales de documentos hipermedia y la creación de relaciones temporales causales entre sus elementos multimedia. Soporta la definición de interacciones del usuario y la simulación de estos eventos asíncronos para previsualizar la presentación. Los usuarios, además, pueden definir las propiedades de la presentación y verificarlas dentro de la interface de visualización espacial. Para ello, incluye un reproductor propio. Las aplicaciones creadas en *STEVE* se pueden exportar a documentos *HTML5* y *NCL*[8]. La interface gráfica de *STEVE 2.0* se muestra en la Fig. 5 y contempla 9 efectos: viento, difusión de agua, vibración, frío, calor, aromas, luces, destellos y niebla.

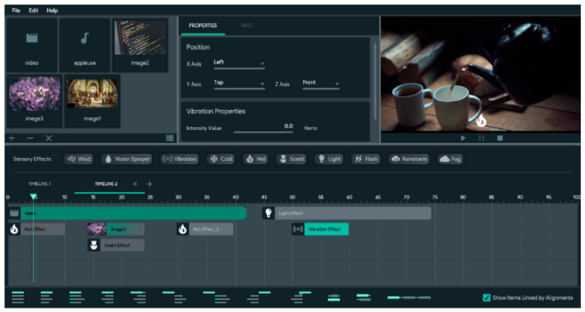


Fig. 5. Interface gráfica *STEVE 2.0* [24]

II.C. Simuladores de efectos mulsemedia

A continuación, se comentan brevemente varios simuladores existentes.

a) Sensory Effect Simulator (SESim). Está incluido en la herramienta *SEVino2*. En la Fig. 6 se muestra su interface gráfica en la que se aprecian los 7 efectos que se pueden simular (viento, luces, niebla, temperatura, vibración, difusión de agua y aromas). Alrededor del reproductor se ubican unas cajas representando los dispositivos generadores de efectos y alguna característica adicional, como su posición. En la parte derecha de la figura se muestra el simulador en funcionamiento. Se resaltan en rojo los efectos activados y datos adicionales (p. ej. el color de las luces o la intensidad del sonido o de los ventiladores). Incluye un panel con mensajes de texto indicando los efectos activos. *SESim* tiene varias limitaciones. Sólo simula los 7 efectos indicados. No permite parar la simulación en un instante dado ni saltar a un punto concreto del contenido AV, para ver qué efectos hay activados en dicho instante. Todos los efectos están colocados alrededor del display por lo que no permite simular efectos desde otras posiciones alrededor del usuario.

b) Simulador integrado en STEVE. Permite comprobar el comportamiento temporal y espacial de las aplicaciones multimedia creadas en *STEVE*, proporcionando una vista gráfica temporal y espacial de forma sincronizada. Sobre el reproductor de contenidos van apareciendo iconos de los efectos que se van activando (en la Fig. 5 se aprecia el icono del efecto calor -*hot*- en color blanco sobre el vídeo mostrado, en la parte superior derecha, mientras se sirve una taza de café/té). Este simulador está limitado a 9 efectos y tampoco muestra desde qué posición se está generando cada efecto.

c) Sensible Media Simulator [25]. Está desarrollado en Flex y también es compatible con *MPEG-V*. Su interface gráfica se muestra en la Fig. 7. Está diseñado para simular efectos sensoriales en un automóvil y testear un sistema diseñado para dicho entorno que aproveche los dispositivos generadores existentes en el mismo, lo cual limita bastante su utilidad.

d) Simulador mulsemedia en 3D basado en Maxon Cinema 4D [26]. La Fig. 8 muestra su interface gráfica. Permite que un usuario añada, elimine y reposicione actuadores en una habitación 3D, distribuidos en las paredes y el techo de la habitación. Están representados por círculos negros con líneas negras en forma de cono (indicando la dirección del actuador). En el simulador la dirección siempre es hacia el centro de la habitación 3D. El color de un actuador se ilumina para indicar que está activo en un momento determinado. Cada color representa un tipo de actuador de efecto diferente: blanco para efectos de luz; rojo para efectos de calor; y verde para el efecto de viento frío. La intensidad de la representación del color está asociada con la intensidad del efecto que se está reproduciendo. Implementa una extensión de *MPEG-V* para permitir una especificación más precisa de la localización de los efectos mediante coordenadas esféricas. También tiene varias limitaciones: sólo simula 3 tipos de efectos (luz, calor y viento frío); muestra información muy limitada de los efectos (su intensidad a través de 3 tonalidades de un mismo color); muestra un único efecto por ubicación; y requiere disponer del software profesional *Maxon Cinema 4D*.

[8] *Nested Context Language* (NCL) es un lenguaje declarativo basado en XML para crear documentos hipermedia para sistemas de TV digital. También es un estándar de la UIT para servicios y aplicaciones multimedia para sistemas IPTV.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

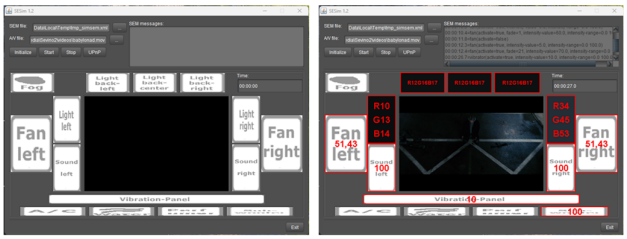


Fig. 6. Herramienta *SESim* incluida en *SEVino2*

Fig. 7. Interface gráfica de *Sensible Media Simulator* [25]

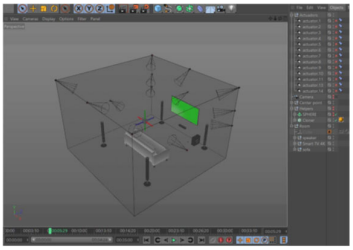


Fig. 8. Simulador en 3D presentado en [26]

e) Simulador en 3D de Real4DAStudio [22]. En la Fig. 9 se muestra su interface gráfica. Herramienta propietaria de autoría que permite crear metadatos SEM basados en *MPEG-V* y que, además, permite realizar una simulación 3D con hasta 9 tipos efectos diferentes (luz, destellos o flash, temperatura, viento, vibración, difusión de agua, aromas, niebla y movimiento de cuerpo rígido) a partir de dichos metadatos. Se emplean interfaces basadas en contenidos y un paradigma basado en eventos para la sincronización de los efectos. Su principal limitación es que la herramienta se ofrece comercialmente[9] y no está disponible de forma abierta.

III. Simulador web *MULSEMEDIA* en 3D

La aplicación desarrollada se ha implementado utilizando tecnología *HTML5* y librerías para trabajar en 3D en la web.

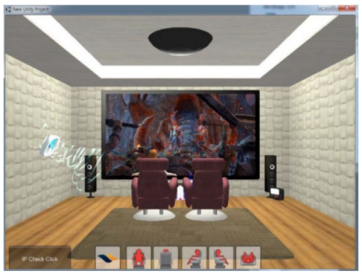


Fig. 9. Ejemplo de simulación 3D de *Real4DAStudio* [22]

Mediante *JavaScript* se maneja la lógica de la aplicación, facilitando la interacción del usuario, controlando la reproducción del contenido AV, y la comunicación, en caso de existir, con el dispositivo *Controlador Mulsemedia* en un entorno *mulsemedia* local. Para su correcto funcionamiento es necesario que, previamente, se hayan generado metadatos SEM correspondientes a los contenidos AV a emplear. Para ello, de las soluciones indicadas en la Sección II, se propone utilizar la herramienta *SEVino2*, que hace uso del lenguaje SEDL y permite crear un archivo XML con metadatos SEM siguiendo las especificaciones de la parte 3 de *MPEG-V* [18]. Los ficheros SEM XML generados son fácilmente modificables por el usuario con un simple editor de texto, permitiendo añadir nuevos efectos o modificar los existentes. Ello será necesario debido a las limitaciones comentadas anteriormente de la herramienta de *SEVino2*. Tanto los contenidos AV (*MP4* o *MPEG-DASH*) a visualizar en la simulación *mulsemedia* como los ficheros de metadatos SEM correspondientes deben estar accesibles vía web. El usuario elige un contenido AV, con sus metadatos SEM, y puede reproducirlo a pantalla completa o en el simulador 3D.

La aplicación desarrollada contempla lo estipulado en el estándar *MPEG-V*. No obstante, se puede adaptar de forma sencilla en el futuro para admitir cualquier tipo de efecto nuevo que se quiera incluir en el fichero de metadatos SEM, aunque no esté definido en dicho estándar. A continuación, se presenta la arquitectura modular de la aplicación.

III.A. Arquitectura de la aplicación

La Fig. 10 muestra la arquitectura modular de la aplicación desarrollada, cuyos módulos se explican a continuación.

a) Reproductor de contenido AV (AV media player). Se han incluido dos reproductores de video: uno basado en el uso del elemento *video* de *HTML5*, para reproducir contenido en formato *MP4*; y otro basado en el reproductor *DashJS*, para reproducir contenido en formato *MPEG-DASH*. Se utilizarán dinámicamente según el contenido escogido por el usuario[10].

b) Analizador XML (XML Parser). Recibe el fichero con metadatos SEM y extrae las características de cada uno de los efectos disponibles a ser activados durante la reproducción del contenido AV seleccionado. Dichas características (tipo de efecto, tiempo de des/activación del efecto, intensidad de activación y si tiene activación o desactivación gradual) se guardan en estructuras de datos internas.

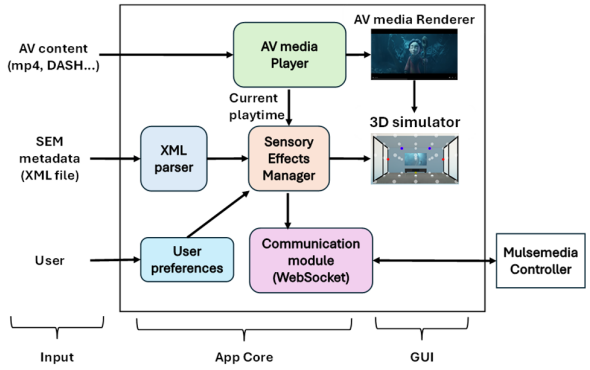


Fig. 10. Arquitectura de la aplicación desarrollada

[9] *Real4DStudio* http://www.real4dhub.or.kr (últ. acceso: septiembre 2025)

[10] La aplicación mostrará de forma automática en una lista desplegable todos los contenidos AV con metadatos SEM disponibles en el servidor.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

c) Gestor de Efectos Sensoriales (Sensory Effects Manager). Las estructuras de datos internas se pasan a este módulo, que también recibe información periódica del módulo reproductor de contenido AV sobre el punto de reproducción. Será el encargado de comprobar si, en cada punto de reproducción recibido, es necesaria la activación de algún efecto en alguna ubicación determinada. En modo simulación, le pasará la indicación pertinente al Módulo de simulación 3D y, en caso de existir dispositivo *Controlador Mulsemedia*, también pasa la información al módulo de comunicaciones para que le envíe el correspondiente mensaje de activación o desactivación.

d) Módulo de Comunicaciones. Se encarga de establecer y controlar la conexión (utilizando tecnología *WebSocket*, a través de un servidor intermedio) con el dispositivo *Controlador Mulsemedia* (ver Fig. 1), si existe, y crear y enviarle los correspondientes mensajes, compatibles con *MPEG-V*, de des/activación de efectos, con los parámetros adecuados y siempre teniendo en cuenta las preferencias del usuario. Dispone de un mecanismo de control de conexiones que gestiona reconexiones automáticas con el servidor en caso de caídas esporádicas de la red de conexión o del canal *WebSocket*.

e) Módulo de preferencias de usuario. En la plataforma se pueden almacenar las preferencias del usuario, de tal manera que se puede ajustar la configuración de los dispositivos generadores de los efectos para mejorar la calidad de su experiencia de consumo *mulsemedia* a través de servicios personalizados y aumentando su disfrute y satisfacción. La aplicación permite al elegir qué efectos, de entre los disponibles, desea que se activen o desactiven durante la sesión y sus características, descartándose los no deseados.

f) Módulo de simulación 3D. El módulo de simulación 3D, explicado más detalladamente en la siguiente Sección, se encarga de representar en un espacio 3D, que simula un recinto o habitación rectangular, los diferentes efectos que se van activando y desactivando durante la reproducción de los contenidos AV y desde qué posiciones concretas. El espacio contiene una pantalla virtual en la que se renderiza el contenido AV, además de elementos 3D que representan la des/activación de efectos en las posiciones contempladas en *MPEG-V* (Fig. 2), alrededor del usuario.

III.B. Entorno de simulación 3D.

Para poder ver, de forma simulada, la des/activación de los efectos sensoriales (uno o varios) en las 45 posiciones alrededor del usuario contempladas en *MPEG-V*, se ha creado un entorno web 3D, utilizando *Blender*[11] y *three.js*[12]. Por un lado, el entorno 3D simula un recinto o habitación en 3D con una pantalla o display 2D virtual en la pared del fondo, en la que se renderizará el contenido de vídeo 2D (Fig. 11). Se supone que el usuario está situado en el centro de la habitación mirando de frente al display 2D virtual. Por otro lado, para poder simular, de forma visual, la des/activación de los efectos, se han colocado esferas trasparentes en dichas 45 posiciones en el entorno en 3D (Fig. 11).

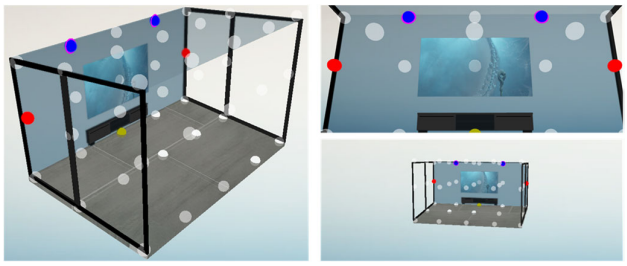


Fig. 11. Entorno de simulación 3D (rotación y zoom)

Cada esfera cambiará de color conforme se active o desactive la generación de efectos sensoriales desde su posición. Si se activan varios efectos en una posición, el primer efecto activado define el color de la esfera de esa posición, y para el resto de efectos que se vayan activando en ella, solapados en el tiempo, se dibujarán anillos alrededor de la esfera, del color correspondiente a cada uno, tal y como se observa en la figura.

El display 2D virtual está colocado en la pared frontal delantera de la habitación y las esferas pueden dificultar la visión correcta del vídeo. Esto no supone problema alguno pues el usuario puede interactuar con el escenario (rotándolo, desplazándolo, acercándolo o alejándolo, creando un efecto de *zoom in/out*) para ver mejor la simulación (Fig. 11). Para la interacción del usuario con el entorno 3D se ha utilizado la herramienta de *three.js* denominada *Raycaster*[13].

III.C. Aplicación web de simulación desarrollada

a) Interface gráfica

La Fig. 12 muestra la interface gráfica de la aplicación desarrollada, indicando las partes principales de la misma. En la parte superior, existe un menú desplegable donde el usuario puede seleccionar un contenido AV de entre los disponibles en el servidor web, junto con controles de reproducción (botón de *play/pause*, control del volumen y barra de progreso).

La versión actual de la aplicación contempla la simulación de la des/activación de 8 efectos sensoriales: niebla, luz, aromas, rocío, temperatura, viento, vibración y destellos o flash. En la parte superior también se muestra un panel informativo donde destacan dichos tipos de efectos y el color asignado a cada uno de ellos: temperatura (frío/calor) - rojo; aromas - verde; luces - azul; viento - purpura; rociador de agua - aqua; niebla - oliva; vibración - negro; destellos - rosa.

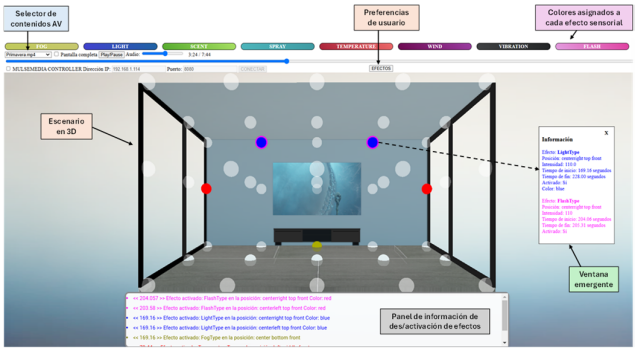


Fig. 12. Interface gráfica de la aplicación web de simulación *mulsemedia*

[11] https://www.blender.org/ (últ. acceso: septiembre 2025)
[12] https://threejs.org/ (últ. acceso: septiembre 2025)
[13] https://threejs.org/docs/#api/en/core/Raycaster (últ. acceso: septiembre 2025)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Se ha incluido el botón '*Efectos'* para la configuración de las preferencias del usuario. Tras pulsarlo aparece una ventana con la lista de efectos indicados en los metadatos SEM (del contenido seleccionado) con cuadros de selección (checkbox) para elegir cuáles se desea incluir en la simulación y cuáles no, descartándose los no deseados.

En la parte central de la página web se muestra el entorno de simulación 3D descrito. Inicialmente, se muestra el escenario desde la parte trasera de la habitación para que se puedan ver todas las posiciones (esferas) representadas, aunque, como ya se ha comentado, el usuario puede interactuar y cambiar la perspectiva, tamaño, etc., tal y como se aprecia en la Fig. 11.

b) Funcionamiento de la aplicación

El funcionamiento de la aplicación se basa en los siguientes pasos[14]:

1. Selección de contenido AV: En primer lugar, el usuario debe seleccionar un contenido AV (fichero de vídeo con extensión '*.mp4*' o archivo índice con extensión '*.mpd*', de *MPEG-DASH*) de la lista desplegable con todos los contenidos disponibles en el servidor. Al seleccionarlo, se descargan automáticamente los metadatos SEM asociados al contenido AV seleccionado. Dichos metadatos son procesados por el módulo *Analizador XML*, que extrae las propiedades de cada uno de los efectos y las guarda en estructuras de datos internas.

2. Reproducción o visualización del contenido de vídeo y simulación 3D. El usuario puede iniciar/parar la reproducción del contenido AV, con el botón '*Play/Pause*', así como saltar hacia adelante o hacia atrás durante la reproducción (pulsando sobre la barra de progreso), sin afectar a la des/activación de los efectos sensoriales. La aplicación, por defecto, muestra el entorno de simulación 3D descrito, en el que se visualiza la reproducción del vídeo en la pantalla 2D virtual incluida y, además, la simulación de la des/activación de efectos de forma dinámica en cada posición (esfera) alrededor del usuario, mientras avanza la reproducción. Para poder conocer qué efectos están activos en un momento y en una posición (es decir, en una esfera coloreada) determinados, se puede pulsar con el ratón sobre la esfera correspondiente y se mostrará, en una ventana emergente, a la derecha de la interface (Fig. 12), un cuadro con información de todos los efectos activos en dicha posición.

Además, la aplicación también permite ver el contenido de vídeo a pantalla completa (activando, en la parte superior izquierda, a la derecha del selector de contenido AV, el *checkbox* de la opción '*Pantalla completa*', ver la Fig. 12), sin ver el emulador de efectos en 3D. Se ha previsto esta opción para que, en caso de disponer de un escenario *mulsemedia* local, y del correspondiente dispositivo *Controlador Mulsemedia* y de dispositivos generadores de efectos reales, se pueda visualizar el contenido AV utilizando toda la pantalla del usuario y tener una experiencia *mulsemedia* lo más realista y completa posible, mientras se activa y desactiva la generación de efectos reales alrededor del usuario.

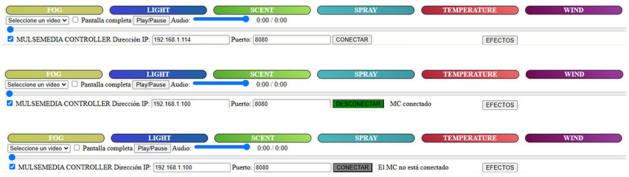


Fig. 13. Conexión con dispositivo *Controlador Mulsemedia*

3. Activación/desactivación de efectos reales. En caso de disponer de un entorno real *mulsemedia* , se hará uso del módulo de comunicaciones, para enviarle (vía canal *WebSocket*) los mensajes pertinentes, compatibles con *MPEG-V*, de des/activación de efectos al dispositivo *Controlador Mulsemedia*. Para ello, en la página web se han incluido varios elementos: un *checkbox* con el texto '*Mulsemedia Controller*', dos cuadros para incluir la dirección o nombre IP del servidor de *WebSocket* y el puerto del servicio, y un botón para realizar la conexión con el mismo. Mientras no se active el *checkbox* los demás elementos no están activos. El color del botón 'Conectar' cambia de color para indicar la conexión con el servidor, además de aparecer un texto contiguo indicando el estado de la conexión (Fig. 13).

La aplicación, aparte de simular la activación en el entorno 3D, en el momento de mostrar la des/activación (o incluso antes, dependiendo de las características de los dispositivos generadores de efectos[15]) enviará los mensajes correspondientes, compatibles con *MPEG-V* al dispositivo *Controlador Mulsemedia*. El protocolo de comunicaciones entre la aplicación y dicho dispositivo, a través del servidor *WebSocket*, está fuera del alcance de este artículo[16].

IV. Conclusiones

En este artículo se ha presentado una herramienta de simulación *mulsemedia* en 3D basada en web y compatible con lo especificado en el estándar *MPEG-V*. Permite ahorrar tiempo y costes a los creadores de sistemas y aplicaciones *mulsemedia* basadas en dicho estándar a la hora de comprobar su correcto funcionamiento, sin necesidad de disponer equipamiento real. A partir de un contenido AV y los correspondientes metadatos SEM, la herramienta permite comprobar, de forma visual y en 3D, la correcta des/activación de los efectos sensoriales y su localización, de forma sincronizada con la reproducción del contenido AV. La aplicación, además, permite al usuario especificar qué efectos incluir en la simulación. El principal inconveniente del uso de simuladores como el presentado es que no se pueden evaluar la calidad de los efectos ni si sus características son las adecuadas (intensidad, realismo…) según el contenido AV mostrado. Para ello se necesita un entorno *mulsemedia* real y, por ello, la versión actual de la herramienta incluye un módulo de comunicaciones basadas en tecnología *WebSocket,* para poderla utilizar en dichos entornos. Así, contempla el envío de los correspondientes mensajes de des/activación, compatibles con *MPEG-V*, a un dispositivo *Controlador Mulsemedia*, si existe. La implementación de dicho dispositivo y del protocolo de comunicaciones con el mismo se presentan en [27]. Por otro lado, el modelo de posicionamiento adoptado en *MPEG-V*

[14] En https://youtu.be/TVm9APxoE3o se muestra un vídeo explicativo.

[15] Algunos dispositivos generadores de efectos tardan a iniciar la activación o generación del efecto con las características deseadas por lo que deben ser activados algunos segundos antes del instante en el que el usuario debe notar el efecto (ejemplo, en un ventilador que debe generar efecto de viento, la velocidad de giro de las aspas empieza desde cero cuando se activa y tarda un tiempo en conseguir la intensidad de viento deseada).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

limita el número de posibles ubicaciones para la aparición de un efecto. En el futuro también se pretende modificar el entorno 3D para incluir un modelo de posicionamiento que también contemple el uso de coordenadas esféricas. Ello permitirá comprobar la generación de efectos desde ubicaciones más específicas, además de las preestablecidas por *MPEG-V*.

AGRADECIMIENTOS

Esta publicación es parte de los proyectos de I+D+i con refs. PID2021-126645OB-I00, financiado por MICIU/AEI/10.13039/501100011033/ y "FEDER Una manera de hacer Europa", "FEDER/UE"; CIAICO/2022/025, financiado por la Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital de la Generalitat Valenciana (DOGV 8919/05.10.2020); ACIF/2021/192 para investigadores predoctorales, financiado por GVA bajo el "Programa I+D+i de la GVA" y Fondo Social Europeo (FSE).

REFERENCIAS

[1] G. Ghinea, C. Timmerer, W. Lin, and S. R. Gulliver, "Mulsemedia: State of the Art, Perspectives, and Challenges.," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 11, no. 1s, pp. 1–23, Oct. 2014, doi: 10.1145/2617994.

[2] Y. Sulema, "Mulsemedia vs. Multimedia: State of the art and future trends," in *International Conference on Systems, Signals, and Image Processing*, IEEE, May 2016, pp. 1–5. doi: 10.1109/IWSSIP.2016.7502696.

[3] C. H. Muntean, I. Tal, D. Bogusevschi, M. Bratu, T. Bi, and G. M. Muntean, "Mulseplayer: A Multi-Sensorial Media Content Delivery Solution to Enhance End-User Quality of Experience," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, 2024, doi: 10.1109/BMSB62888.2024.10608351.

[4] M. Melo *et al.*, "Immersive multisensory virtual reality technologies for virtual tourism: A study of the user's sense of presence, satisfaction, emotions, and attitudes," *Multimed Syst*, 2022, doi: 10.1007/s00530-022-00898-7.

[5] D. Marfil, F. Boronat, J. Gonzalez, and A. Sapena, "Integration of Multisensorial Effects in Synchronised Immersive Hybrid TV Scenarios," *IEEE Access*, vol. 10, pp. 79071–79089, 2022, doi: 10.1109/ACCESS.2022.3194170.

[6] C. Velasco and M. Obrist, "Multisensory Experiences: Where the senses meet techonology," *Oxford*, p. 112, 2020, doi: 10.1093/oso/9780198849629.001.0001.

[7] M. Waltl, B. Rainer, C. Timmerer, and H. Hellwagner, "A toolset for the authoring, simulation, and rendering of sensory experiences," in *MM 2012 - Proceedings of the 20th ACM International Conference on Multimedia*, 2012, pp. 1469–1472. doi: 10.1145/2393347.2396522.

[8] Z. Yuan, G. Ghinea, and G. M. Muntean, "Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery," *IEEE Trans Multimedia*, vol. 17, no. 1, pp. 104–117, Jan. 2015, doi: 10.1109/TMM.2014.2371240.

[9] "ISO/IEC 23005. 2011. Information technology - Media context and control (MPEG-V)," 2011.

[10] L. Jalal, M. Anedda, V. Popescu, and M. Murroni, "Internet of Things for Enabling Multi Sensorial TV in Smart Home," in *2018 IEEE Broadcast Symposium, BTS 2018*, IEEE, Oct. 2018, pp. 1–5. doi: 10.1109/BTS.2018.8550959.

[11] A. Covaci, L. Zou, I. Tal, G. M. Muntean, and G. Ghinea, "Is multimedia multisensorial? - A review of mulsemedia systems," Aug. 01, 2018, *Association for Computing Machinery*. doi: 10.1145/3233774.

[12] O. A. Ademoye, N. Murray, G.-M. Muntean, and G. Ghinea, "Audio Masking Effect on Inter-Component Skews in Olfaction-Enhanced Multimedia Presentations," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 12, no. 4, pp. 1–14, Aug. 2016, doi: 10.1145/2957753.

[13] M. Obrist, A.N. Tuch, and K. Hornbaek, "Opportunities for odor: experiences with smell and implications for technology," in *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, May 2014.

[14] N. Murray, O. A. Ademoye, G. Ghinea, and G.-M. Muntean, "A Tutorial for Olfaction-Based Multisensorial Media Application Design and Evaluation," *ACM Comput Surv*, vol. 50, no. 5, pp. 1–30, Sep. 2017, doi: 10.1145/3108243.

[15] I. F. Akyildiz, H. Guo, R. Dai, and W. Gerstacker, "Mulsemedia Communication Research Challenges for Metaverse in 6G Wireless Systems," *ITU Journal on Future and Evolving Technologies*, vol. 4, no. 4, Dec. 2023.

[16] E. B. Saleme, A. Covaci, G. Mesfin, C. A. S. Santos, and G. Ghinea, "Mulsemedia DIY: A survey of devices and a tutorial for building your own mulsemedia environment," Jun. 01, 2019, *Association for Computing Machinery*. doi: 10.1145/3319853.

[17] D. P. de Mattos and D. C. Muchaluat-Saade, "STEVE: a Hypermedia Authoring Tool based on the Simple Interactive Multimedia Model," in *DocEng '18: Proceedings of the ACM Symposium on Document Engineering 2018*, Halifax NS Canada, Aug. 2018, pp. 1–10.

[18] "International standard ISO/IEC 23005-3. Information technology — Media context and control — Part 3: Sensory information (MPEG-V Part 3)," 2013.

[19] B. Choi, E. -S. Lee, and K. Yoon, "Streaming Media with Sensory Effect," in *2011 International Conference on Information Science and Applications*, Jeju, Korea (South), 2011, pp. 1–6.

[20] S-K. Kim, "Authoring multisensorial content," *Signal Process Image Commun*, vol. 28, no. 2, pp. 162–167, 2013.

[21] F. Danieau, J. Bernon, J. Fleureau, P. Guillotel, N. Mollet, and M. Christie, "H-Studio: An Authoring Tool for Adding Haptic and Motion Effects to Audiovisual Content," in *Proceedings of the adjunct publication of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 83–84.

[22] S. H. Shin, K. S. Ha, H. O. Yun, and Y. S. Nam, "Realistic media authoring tool based on MPEG-V international standard," in *International Conference on Ubiquitous and Future Networks, ICUFN*, IEEE, Jul. 2016, pp. 730–732. doi: 10.1109/ICUFN.2016.7537133.

[23] D. P. De Mattos, D. C. Muchaluat-Saade, and G. Ghinea, "Beyond Multimedia Authoring: On the Need for Mulsemedia Authoring Tools," *ACM Comput Surv*, pp. 1–31, 2021.

[24] D. P. De Mattos, D. C. Muchaluat-Saade, and G. Ghinea, "An Approach for Authoring Mulsemedia Documents Based on Events," *2020 International Conference on Computing, Networking and Communications, ICNC 2020*, pp. 273–277, Feb. 2020, doi: 10.1109/ICNC47757.2020.9049485.

[25] S-K. Kim, Y-S. Joo, and Y. Lee, "Sensible Media Simulation in an Automobile Application and Human Responses to Sensory Effects," *ETRI Journal*, vol. 35, no. 6, pp. 1001–1010, 2013.

[26] M. Josué *et al.*, "Modeling sensory effects as first-class entities in multimedia applications," *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, pp. 225–236, 2018, doi: 10.1145/3204949.3204967.

[27] F. Boronat, C. Sáez, Ll. Simó, and J. Pastor, "Plataforma web mulsemedia," in *Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025)*, Cáceres (España), Nov. 2025, pp. 1–8.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Plataforma web mulsemedia

Fernando Boronat, Cristian Sáez, Lluc Simó y Javier Pastor
Departamento Comunicaciones,
Universitat Politècncia de València – Campus de Gandia
IIM R&D Group – Calle Paraninfo, 1, 46730, Grao de Gandia (Valencia-España)
fboronat@dcom.upv.es, {csaez1, lsimmar, fjpastor}@upv.es

Abstract. Los sistemas multimedia tradicionales incluyen contenidos audiovisuales que estimulan la vista y el oído. Sin embargo, estimular los otros sentidos proporciona experiencias más realistas e inmersivas, mejorando la calidad de la experiencia (QoE) de consumo de contenidos. Existen soluciones y estándares (ej. *MPEG-V*) que permiten integrar dispositivos generadores de efectos sensoriales en sistemas multimedia, como aromas, frío/calor, vibraciones, etc. La mayoría no están disponibles o han sido diseñadas y evaluadas en escenarios particulares. Además, en muchas se han utilizado dispositivos generadores desarrollados ad-hoc (no comerciales). En este artículo se propone una plataforma mulsemedia basada en web y en el uso de dispositivos IoT y generadores comerciales, para la distribución y consumo/simulación en 3D de contenidos mulsemedia, compatible con *MPEG-V*. Se presenta prototipo de implementación cuyo correcto funcionamiento ha sido validado de forma preliminar y se ha corroborado que la inclusión de efectos mejora la inmersión, el realismo y la QoE de consumo.

***Palabras Clave*- Efectos Sensoriales, Experiencia Sensorial, *MPEG-V*, Mulsemedia, Simulación Mulsemedia.**

I. Introducción

Los sistemas multimedia normalmente incluyen contenidos audiovisuales (AV) que estimulan los sentidos de la vista y el oído. Sin embargo, comunidad científica e industria llevan años trabajando en el desarrollo de sistemas multimedia que incluyen metadatos de efectos sensoriales (SEM) asociados a los contenidos AV, siendo capaces de generar dichos efectos estimulando todos los sentidos de los usuarios. Algunos ejemplos de efectos son aromas (olfato), sabores (gusto), vibraciones, presión, efectos de viento (tacto), iluminación especial, temperatura, humedad, humo, vaporizaciones, etc. Se persigue, de esta forma, conseguir experiencias de usuario más realistas e inmersivas, mejorando la calidad de las mismas (QoE) [1]. Actualmente, está creciendo el número de empresas comercializando generadores de efectos sensoriales y en algunos parques temáticos y cines 4D ya se ofrecen experiencias multisensoriales. El concepto ***MulSeMedia*** (*Multiple Sensorial Media* [2]) está tomando mucha fuerza en multitud de áreas, tales como telecomunicaciones, educación, publicidad, entretenimiento, turismo, etc. En [1] se describen estructuras o sistemas que permiten integrar contenido AV y generación de efectos sensoriales [1] . La investigación en este campo ha avanzado significativamente, incluso apareciendo estándares relacionados (ej., *MPEG-V* [3]) y arquitecturas IoT [4] que permiten dicha integración. En [5] se presenta un estudio exhaustivo sobre sistemas mulsemedia, destacando la poca exploración del campo mulsemedia, los desafíos y su aplicabilidad en las áreas mencionadas. En [6] se recopilan muchos de los avances más destacados en el área. En [7], se presentan recomendaciones para el diseño y la evaluación de sistemas mulsemedia. En [8] se presentan los antecedentes, la historia y las tecnologías esenciales de codificación, decodificación y comunicaciones que sustentan este campo, centrándose en aplicaciones de realidad extendida (XR) y las comunicaciones holográficas. Por otro lado, en [9], [10] y [11] se presentan revisiones de dispositivos generadores de efectos, y guías para construir entornos mulsemedia, en escenarios de escritorio y en escenarios inmersivos (RV o vídeos 360).

La mayoría de las soluciones software/hardware (SW/HW) encontradas para la inclusión de efectos sensoriales están orientadas comercialmente para ser utilizadas en escenarios locales y sólo son efectivas para distancias cortas. Además, a pesar de la existencia del estándar *MPEG-V*, aún quedan muchos retos por investigar [12]: cómo se señalizan y cómo se integran los efectos sensoriales en los flujos multimedia en la cadena de distribución; los resultados de las combinaciones posibles entre varios efectos sensoriales; la sincronización con contenidos AV; el número apropiado de efectos sensoriales a generar y sincronizar; analizar suficientemente el impacto de factores relevantes, como la intensidad, persistencia, grado y retardo de percepción; posibles efectos enmascaradores, etc. Por otra parte, la inclusión de contenido mulsemedia junto con contenidos omnidireccionales (ej., vídeo 360) presenta el reto adicional de la dependencia del campo de visionado. Para poder investigar en todos estos temas es necesario disponer de escenarios o entornos mulsemedia reales adecuados en los que realizar los experimentos pertinentes. En la Fig. 1 se muestra un ejemplo de la infraestructura necesaria para su implementación. Sin embargo, montar un escenario completo, con dispositivos reales colocados en diversas posiciones alrededor del usuario, sólo para comprobar el correcto funcionamiento de las experiencias mulsemedia diseñadas puede resultar costoso y poco flexible. Para ello, un método mucho más rápido y económico es el uso de simuladores mulsemedia. Sin embargo, un simulador no permite

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

comprobar la calidad, las características o la idoneidad de los efectos con respecto al contenido AV consumido.

En este artículo se presenta, como *principal contribución*, una plataforma modular, basada en tecnología web y dispositivos IoT, para la distribución y consumo (real o simulado) de contenidos mulsemedia, compatible con el estándar *MPEG-V*, siguiendo la estructura de la Fig. 1. La idea subyacente es la de poder utilizarla en el futuro para seguir investigando en los retos pendientes en el área mulsemedia. Se presenta su arquitectura, sus módulos y los principales procesos involucrados en ella. Para validar la propuesta, se ha implementado un prototipo de escenario mulsemedia basado en la misma, que ha sido validado inicialmente por 10 usuarios, corroborando, entre otras cosas, su correcto funcionamiento y que la inclusión de efectos sensoriales mejora la inmersión, el realismo y, por tanto, la QoE. Las principales contribuciones de este artículo se pueden resumir en las siguientes: 1) Plataforma mulsemedia para entorno local, con módulos reusables, basada en el uso tecnología web, dispositivos IoT de bajo coste y dispositivos comerciales generadores de efectos; 2) Aplicación web de reproducción y simulación 3D mulsemedia, compatible con el estándar *MPEG-V*; 3) Módulo de configuración vía web de preferencias mulsemedia del usuario; 4) Módulo de control de la generación de efectos sensoriales; 5) Protocolos de comunicaciones entre la aplicación y el módulo de control y entre este y los dispositivos de control de la generación de efectos; y 6) Prototipo de escenario mulsemedia basado en la propuesta.

La estructura del artículo es la siguiente: en la sección 2 se presenta un resumen de trabajos relacionados ; en la sección 3 se describe la plataforma desarrollada, su arquitectura y sus componentes; en la sección 4 se presenta el prototipo de escenario implementado; y, finalmente, se presentan unas breves conclusiones y las referencias bibliográficas utilizadas.

II. Trabajos Relacionados

Para poder crear experiencias mulsemedia, es necesario haber generado previamente metadatos SEM y asociarlos al contenido AV de alguna forma. De esta manera, posteriormente los contenidos AV junto con dichos metadatos pueden ser distribuidos y utilizados en sistemas mulsemedia que puedan interpretarlos. Además, entre los componentes de los sistemas mulsemedia hace falta un intercambio de información que debe seguir un formato y unos protocolos preestablecidos. El estándar *MPEG-V* proporciona un posible formato, pero los procedimientos de intercambio están muy abiertos. En esta sección, en primer lugar, se presenta brevemente lo que incluyen las partes del estándar que afectan directamente a la propuesta presentada. A continuación, se presentan una serie de herramientas de autoría y de simulación de presentaciones mulsemedia, y, finalmente, se comentan brevemente otras plataformas mulsemedia relacionadas.

II.A. El estándar MPEG-V: Partes 2, 3 y 5

El estándar *MPEG-V* proporciona herramientas para describir mundos reales y virtuales. Especifica una serie de formatos, basados en XML, para intercambiar datos entre ambos mundos. Se centra en la comunicación y representación

Fig. 1. Ejemplo de entorno local mulsemedia

de objetos del mundo virtual a objetos del mundo real y viceversa. Entre otras cosas, ofrece una gran cantidad de herramientas para describir y representar efectos y dispositivos sensoriales. Consta de 7 partes, de las cuales las relacionadas con la propuesta son las partes 2, 3 y 5.

La parte 2 define el esquema XML para proporcionar información sobre características de los sensores y generadores de efectos o actuadores, junto con preferencias de usuario en cuanto a su activación. Especifica el Lenguaje de Descripción de Información de Control, basado en esquemas XML, para describir la estructura básica de la información de control. Para ello se emplean varios vocabularios. El Vocabulario de Descripción de Capacidades de Dispositivos especifica esquemas XML para describir las capacidades de actuadores (ej, ventiladores, calefactores, vibradores o generadores de aromas). El Vocabulario de Descripción de Capacidades de Sensores permite describir las capacidades de sensores (ej., de luz, humedad o temperatura). El Vocabulario de Preferencias de Efectos Sensoriales permite describir las preferencias de los usuarios con respecto a efectos específicos. El Vocabulario de Preferencias de Adaptación de Sensores permite describir las preferencias de cada usuario sobre sensores específicos.

La parte 3 del estándar define la descripción de los efectos sensoriales, la sintaxis y la semántica de estos, proporcionando las estructuras básicas de información para poder generar los metadatos SEM. Dichas estructuras consisten en bloques de construcción y atributos comunes (ej., tipo de efecto, duración, intensidad, etc.) que se utilizan en todos los efectos especificados. También define un Lenguaje de Descripción de Efectos Sensoriales, basado en XML, para describir dichas estructuras. Los efectos se deben describir utilizando el denominado Vocabulario de Efectos Sensoriales, que permite especificar todos los efectos sensoriales en detalle, sus atributos generales (información de activación, duración, prioridad y posición) y, según el efecto, sus atributos específicos (ej., color y aromas concretos, para efectos de luces y olfativos, respectivamente). Se deben utilizar ambos, de forman conjunta, para crear metadatos SEM. Para describir los comandos del generador de un efecto (actuador), y la información detectada, *MPEG-V* proporciona el Lenguaje de Descripción de Interfaz de Interacción. Además, también define el Lenguaje de Descripción de Información de Control para ejecutar los comandos de los generadores de efectos o actuadores, de activación[1], según las preferencias del usuario.

[1] En adelante, se utiliza la palabra 'activación' para referirse tanto a operaciones de activación como de desactivación de efectos.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Este lenguaje también permite la especificación de las capacidades de los actuadores y de los sensores. En esta parte del estándar *MPEG-V*, todos los efectos tienen el atributo común *location*, que especifica la región desde la que se debe percibir el efecto. Se consideran 3 planos (frontal, medio y trasero), 3 alturas en cada uno (baja, media y alta) y 5 posiciones en cada altura. Ello supone un total de 45 posiciones desde las que se puede activar un efecto, representadas concatenando 3 palabras (ej., '*right:bottom:front*' indica zona frontal, suelo y a la derecha).

Por último, la parte 5 de *MPEG-V* define el esquema XML para información obtenida de sensores y los tipos de datos utilizados en los comandos de activación de efectos enviados a los dispositivos generadores de efectos. Especifica la sintaxis y la semántica de los formatos de datos para dispositivos de interacción. Incluye el IIDL para describir las estructuras básicas de los mensajes intercambiados . Como parte del IIDL, se proporcionan dos vocabularios: el Vocabulario de Comandos de Dispositivo para especificar mensajes de control de dispositivos generadores y el Vocabulario de Información obtenida de Sensores, que especifica interfaces para aceptar información proporcionada por sensores.

II.B. Generación de metadatos de efectos sensoriales (SEM)

En [13] se presenta una revisión y comparación de herramientas de autoría mulsemedia, que permiten la generación de metadatos SEM, y varias propuestas para la representación de efectos y sus características. Entre ellas cabe destacar *SEVino2* [1] , *STEVE* [14], *RoSE Studio* [20] , *SMURF* [15], *H-Studio* [16] o *Real 4D Studio* [17]. Todas permiten definir, de forma gráfica, los metadatos (con número limitado de tipos de efectos) y guardarlos en ficheros.

II.C. Simuladores de efectos mulsemedia

En *SEVino2* [1] se incluye la herramienta *Sensory Effect Simulator* (*SESim*) para comprobar la correcta activación de hasta 7 tipos de efectos sensoriales. Alrededor del reproductor de vídeo se colocan cuadros que representan los generadores de efectos y alguna característica adicional (posición). Se resaltan en rojo los efectos activados y datos adicionales (ej., el color de las luces o intensidad de los ventiladores) y muestra un panel de texto con los registros de activación. No permite parar la simulación en un instante dado ni saltar a un punto del contenido AV. *STEVE* [14] permite comprobar la activación de hasta 9 efectos sensoriales. Sobre el propio reproductor de contenidos se solapan iconos de los efectos que se van activando. No muestra desde dónde se está generando cada efecto. Otro simulador, también compatible con *MPEG-V* es el *Sensible Media Simulator* [18], pero está diseñado para simular efectos sensoriales en un automóvil. En [19] se presenta un simulador mulsemedia en 3D, de solo 3 tipos de efectos, creado con el SW *Maxon Cinema 4D*. Los actuadores están distribuidos en las paredes y el techo de una habitación 3D, representados por puntos negros con líneas negras en forma de cono apuntando al centro de la habitación. El color del cono se ilumina (blanco para efecto de luz, rojo para calor y verde para viento frío) con 3 tonalidades (3 niveles de intensidad) para indicar que está activo. Es compatible con *MPEG-V*, pero también contempla posicionamiento en coordenadas esféricas. En [17] se presenta una herramienta comercial propietaria de autoría de contenidos mulsemedia, *Real4DAStudio*, que permite realizar una simulación 3D con hasta 9 tipos efectos. No se ofrece comercialmente.

II.D. Plataformas Mulsemedia e IoT

En [21] se analizan tres desafíos técnicos de los sistemas mulsemedia y se ofrece una orientación práctica sobre cómo abordar la diversidad de HW y SW al integrar componentes en dichos sistemas. Dichos desafíos son: la multifuncionalidad que las soluciones mulsemedia deben proporcionar a las aplicaciones multimedia heterogéneas para su compatibilidad y la reutilización de los componentes; la sincronización; y la manejabilidad y configurabilidad considerando arquitecturas complejas compuestas por entidades heterogéneas. En [22], se presenta un conjunto de herramientas de código abierto que proporcionan un marco completo para la generación y consumo de efectos sensoriales de extremo a extremo. Entre ellas, está el reproductor *Sensory Effect Media Player*, que incluye un módulo de gestión de efectos y soporta varios dispositivos comerciales específicos de generación de efectos sensoriales (ventiladores, luces, aromas y vibración). En [21] se presenta la plataforma *PlaySEM*, que incluye herramientas SW para trabajar con aplicaciones mulsemedia y HW heterogéneo. Está compuesta por 3 componentes SW: la herramienta principal *PlaySEM Sensory Effects Renderer* (SER), responsable de producir los efectos sensoriales en el entorno del usuario convirtiendo los metadatos SEM *MPEG-V* en comandos; la herramienta *Sensory Effects (SE) Video Player,* responsable de reproducir el contenido AV y proporcionar los metadatos a la herramienta *SER*; y la herramienta *PlaySEM Device Control (DC)*, que recibe comandos de la herramienta *SER* para controlar los dispositivos generadores de efectos. Propone la separación de responsabilidades en los diferentes componentes del sistema que pueden funcionar por separado, de modo que se pueden reutilizar. Utiliza componentes de código abierto, *Cling UPnP Core* para comunicaciones entre herramientas y *VLJC*[2] para reproducir vídeos. En [23]-[24] se presentan plataformas mulsemedia para sistemas de cine en casa, basadas en *MPEG-V*. La de [23] incluye una unidad de control central que recibe información sensorial de un servidor de contenidos, la procesa y envía señales a las unidades de control de dispositivos generadores de efectos. En [24] se explica cómo utilizar *MPEG-V* y se propone el uso del protocolo *Simple Service Discovery Protocol* (*SSDP*) modificado para el descubrimiento de dispositivos, pero no se implementa ni evalúa la propuesta. En [25] se presenta una solución no compatible *MPEG-V*, extremo a extremo, integrando efectos sensoriales y elementos interactivos en sistema de TV híbrida.

El nuevo paradigma de la IoT, la aparición de multitud de dispositivos inteligentes y sensores conectados a la misma junto con el desarrollo de nuevas tecnologías de comunicación móvil de corto alcance, con mayor eficiencia energética, van a jugar un papel clave en el desarrollo de sistemas y plataformas mulsemedia en el futuro. En particular, existen varios puntos en común entre IoT y los términos definidos en *MPEG-V* (ej., sensores, actuadores, motores inteligentes, adaptación). Existen varios trabajos en los que se emplea el

[2] Java Framework for the VLC Media Player https://github.com/caprica/vlcj (últ. acc.: septiembre 2025)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

concepto de IoT en escenarios mulsemedia [26] [27]. En [26] se explica la relación entre IoT, multimedia y *MPEG-V*, y, además, se presentan escenarios de uso, la sintaxis XML utilizada, y varios ejemplos sobre su utilización. En [27] se propone una arquitectura basada en el uso de una plataforma IoT en la nube, para posibilitar la creación de escenarios mulsemedia de entretenimiento en el hogar. A diferencia de las plataformas estudiadas, en este artículo se propone una plataforma mulsemedia novel, modular, abierta, basada en tecnologías web y dispositivos IoT, que permite la gestión y el control de dispositivos comerciales de generación de efectos y sensores, y que se explica a continuación.

III. Plataforma Mulsemedia Propuesta

En esta Sección, se describe la arquitectura basada en capas de la plataforma desarrollada, así como una visión general de la misma.

III.A. Arquitectura de la plataforma

La arquitectura de la plataforma desarrollada, al igual que en [21] y [28], está dividida en varias capas independientes (Fig. 2), permitiendo incorporar nuevas tecnologías en alguna de ellas sin modificar las capas adyacentes. Tiene 5 capas:

a) Capa Física. Incluye objetos del mundo real (*real world objects* o *RWO*) que pueden ser sensores o actuadores. Se comunican con la capa superior mediante tecnologías cableadas o inalámbricas y protocolos de datos estándar.

b) Capa de Virtualización. Para desacoplar el HW de la representación del SW en la plataforma, se introduce el concepto de objeto virtual (*virtual object* o VO) como contraparte digital de cualquier *RWO*, de modo que cada objeto en la Capa Física se representa mediante una virtualización. El VO representa el *RWO* en términos de descripción semántica y funcionalidades. Cuenta con dos interfaces que permiten una doble comunicación: por una el VO se comunica vía IP con la capa superior; y, por la otra, se conecta con el *RWO*. En la implementación presentada más adelante, esta capa se integra en el SW instalado en cada dispositivo microcontrolador que controla a un *RWO*.

c) Capa de Agregación y Control. Se encarga de la agregación y control de datos provenientes de una o más VO para garantizar un alto nivel de reutilización. Por un lado, se encarga de obtener y procesar datos de los VO. Por otro lado, es la encargada de procesar los mensajes recibidos de la capa de aplicación y enviar los correspondientes mensajes de activación a los VO. En la implementación presentada más adelante, esta capa está implementada en el SW que se ejecuta en el dispositivo denominado *Controlador Mulsemedia* (app MC).

d) Capa de Aplicación. En esta capa están las aplicaciones de usuario, responsables de la reproducción de contenidos AV y procesamiento de sus metadatos SEM, según la información recibida de la capa anterior. En la implementación presentada más adelante, esta capa está implementada en la aplicación que se ejecuta en el dispositivo principal (app MS).

e) Capa de Usuario. En esta capa está el usuario final que, a través de la interface gráfica de la aplicación, puede, por un lado, configurar su perfil de usuario y preferencias sobre los efectos sensoriales a incluir en la experiencia y sus características, y, por otro lado, seleccionar el contenido AV principal e interactuar con la aplicación durante la sesión.

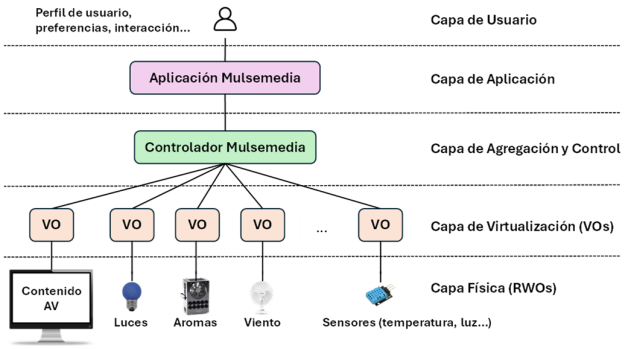


Fig. 2. Arquitectura de la plataforma propuesta

III.B. Descripción general de la plataforma

En esta sección se presenta la plataforma propuesta (Fig. 3) siguiendo la arquitectura anterior y la estructura de la Fig. 1. Está formada por varios componentes multifuncionales reusables [21]. Para la comunicación entre ellos, se utilizan protocolos de comunicaciones típicos ligeros y fiables. Por un lado, incluye un *servidor web* donde se alojará la aplicación web de reproducción (y simulación) y un *servidor de contenidos* AV junto con sus metadatos SEM. También incluye un *dispositivo principal* en el que, mediante la aplicación web (*aplicación Main Screen* o *app MS*) se pueden reproducir los contenidos AV y comprobar, mediante simulación gráfica en 3D, la activación de los efectos sensoriales en los instantes y posiciones adecuados durante la reproducción. Por otro lado, se dispone de *dispositivos sensores* y actuadores o *dispositivos generadores de efectos sensoriales reales* cuya activación se controla mediante un *microcontrolador o dispositivo IoT*. Además, la plataforma incluye un dispositivo intermedio, denominado *Controlador Mulsemedia*, ejecutando una aplicación denominada *app MC*, que hace las funciones de pasarela entre el dispositivo principal y los microcontroladores. Las comunicaciones entre el dispositivo principal y el *Controlador Mulsemedia* se realizan a través de un servidor de comunicaciones (*WebSocket*) mientras que las comunicaciones entre el *Controlador Mulsemedia* y los microcontroladores se realizan a través de HTTP (API REST). La app MC se encarga de descubrir la existencia de dispositivos generadores de efectos en el entorno local, así como de detectar si se producen cambios en su estado (apagar, perder conexión de red, etc.) para notificárselo a la app MS. Además, la app MC debe procesar e interpretar los mensajes recibidos de la app MS, compatibles con *MPEG-V*, de activación de efectos y, a su vez, crear y enviar los mensajes con los comandos apropiados a los correspondientes microcontroladores.

Para garantizar una buena QoE durante el uso de la plataforma es importante que los efectos se generen en los momentos oportunos para que los usuarios los perciban de forma sincronizada con el contenido AV. Para ello las referencias de reloj de las dos aplicaciones debe estar sincronizadas (por ejemplo, mediante protocolos como NTP).

A continuación, se presenta la implementación de un prototipo de sistema mulsemedia siguiendo la propuesta de plataforma descrita en esta Sección.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

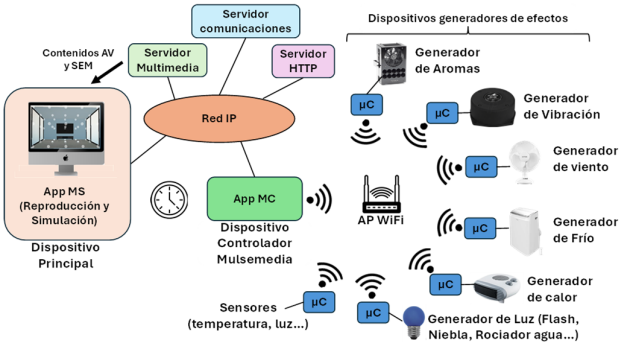


Fig. 3. Descripción general de la plataforma propuesta

IV. Prototipo Implementado

IV.A. Aplicación de reproducción y simulación mulsemedia

Como app MS se ha implementado una aplicación web utilizando tecnología HTML5 y librerías para trabajar en 3D en la web. Para su correcto funcionamiento necesita que se hayan generado previamente los metadatos SEM *MPEG-V* correspondientes a los contenidos AV a visualizar en la misma (p. ej., con alguna de las herramientas indicadas en la Sección II.B). Permite al usuario elegir un contenido (MP4 o MPEG-DASH) y reproducirlo, bien a pantalla completa o bien en un simulador 3D, activando los efectos indicados en los metadatos SEM.

La Fig. 4 muestra la arquitectura de la aplicación desarrollada, de la que cabe destacar 3 módulos: El *Gestor de Efectos Sensoriales* recibe del módulo *analizador XML* (*parser XML*) estructuras de datos internas con información de los efectos sensoriales y también recibe información del reproductor de contenido AV sobre su punto de reproducción. Será el encargado de comprobar si, en el instante de reproducción actual, es necesaria la activación de algún efecto en una determinada ubicación del escenario y, en ese caso enviar a través del módulo de comunicaciones el correspondiente mensaje, compatible con *MPEG-V*, al dispositivo *Controlador Mulsemedia* vía *WebSocket*. También le pasará la correspondiente indicación de activación de efectos al *Módulo de Simulación 3D*. El *Módulo de preferencias de usuario* permite al usuario elegir qué efectos, de entre los disponibles, desea que se activen o desactiven durante la sesión, descartándose los no deseados, así como configurar sus principales características (p.ej., sus intensidades máximas o mínimas deseadas). El *Módulo de simulación 3D* se encarga de representar, en un espacio 3D rectangular (creado con *Blender* y *three.js*) interactivo (admite rotación, ampliación, etc.), los diferentes efectos que se van activando durante la reproducción. Dicho espacio 3D contiene una pantalla virtual en su interior donde se renderiza el contenido AV y elementos 3D (esferas) que cambian de color representando la activación de los efectos considerados (en colores diferentes) en cada una de las 45 posiciones definidas en *MPEG-V*. Al hacer click sobre ellas se indican los efectos activos en su posición[3]. La Fig. 5 muestra la interface gráfica de la aplicación web desarrollada, indicando sus partes principales. La app MS se describe con más detalle en [29].

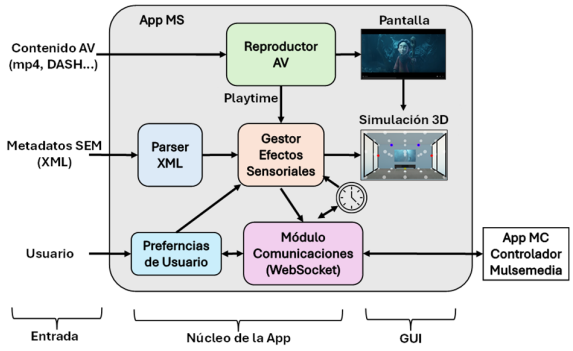


Fig. 4. Arquitectura de la app MS implementada

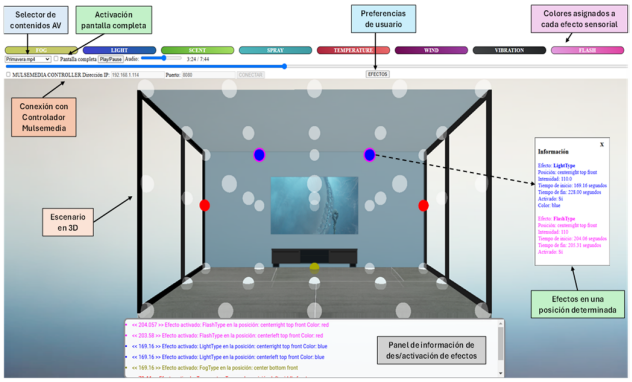


Fig. 5. Interface gráfica de la app MS implementada

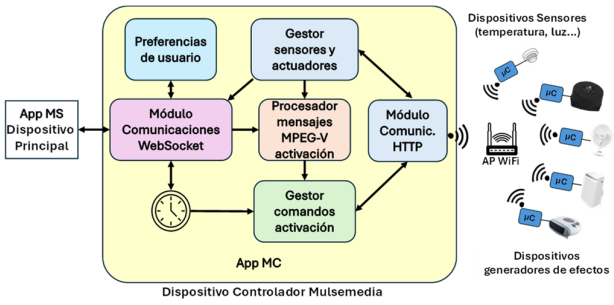


Fig. 6. Arquitectura de la app MC implementada

El funcionamiento de la app MS se basa en los siguientes pasos: 1) Conexión *WebSocket* con dispositivo *Controlador Mulsemedia*; 2) Selección de contenido AV; 3) Obtención de las características de los generadores de efectos; 4) Obtención de preferencias del usuario; 5) Reproducción o visualización del contenido AV (pantalla completa o en modo simulación); y 6) activación de efectos durante la misma.

IV.B. Aplicación del Controlador Mulsemedia

La app MC desarrollada se basa en el uso de *Node.js* y el lenguaje *JavaScript*. Tiene la arquitectura modular que se muestra en la Fig. 6. El *Módulo de comunicaciones por WebSocket* está encargado de las comunicaciones con la App MS a través del servidor de *WebSocket*. El *Módulo de preferencias del usuario* se encarga de almacenar el perfil y las preferencias que el usuario ha configurado en la app MS y de enviárselas cuando ésta las necesite. El *Módulo gestor de sensores y actuadores* se encarga de descubrir a los sensores y dispositivos generadores de efectos sensoriales y cambios en la red local. Obtendrá sus características (*capabilities*) y se las

[3] En https://youtu.be/TVm9APxoE3o se muestra un vídeo de la app MS.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

pasará a través del *módulo de comunicaciones* vía *WebSocket* a la app MS mediante mensajes *MPEG-V*. En el caso de los sensores le enviará a la app MS mensajes *MPEG-V* con los valores medidos. El *Módulo procesador de mensajes MPEG-V de activación de efectos* se encarga de procesar (*parser XML*) mensajes *MPEG-V* que le llegan de la app MS. El *Módulo de gestión de comandos de activación de efectos reales* recibe información de los efectos a activar, crea comandos en formato JSON y los envía vía API REST a los microcontroladores de los dispositivos generadores.

Cuando se inicia la app MC, esta abre una conexión con el servidor de *WebSocket* que proporciona comunicación con la app MS. Además, también inicia un proceso de descubrimiento de los dispositivos mulsemedia existentes en el entorno local, tanto generadores de efectos como sensores. Así, se obtienen los efectos disponibles y sus características: el tipo, la posición desde la que se activan y las características técnicas de su dispositivo generador. Cuando la app MS consigue, por primera vez, conexión vía *WebSocket*, con la app MC, le solicita dicha información y las preferencias del usuario. La app MC se lo envía en mensajes compatibles *MPEG-V*. Posteriormente, sólo enviará dicha información si se detectan cambios. De este modo, durante la reproducción, la app MS sólo envía a la app MC mensajes de activación de efectos que estén en el entorno en ese momento, en los instantes adecuados y según las preferencias del usuario. Cuando la app MC recibe uno de ellos, lo procesa y lo traduce al correspondiente comando de activación que enviará a través de la red local utilizando la API REST al microcontrolador correspondiente. Los microcontroladores utilizados (Arduinos, Raspberry Pis…) disponen del SW adecuado para proporcionar a la app MC información sobre los dispositivos de generación de efectos o sensores que controlan y de recibir y procesar los comandos de la app MC para la activación de dichos efectos en los instantes indicados. En el caso de controlar sensores, son capaces de recibir los datos del sensor y enviárselos, bien de forma periódica o bien bajo petición expresa, a la app MC.

IV.C. Comunicación entre las apps MS y MC

La comunicación entre las apps MS y MC se realiza a través del servidor *WebSocket,* tal y como muestra la Fig. 7. Para establecer la conexión entre ambas se ha implementado un intercambio de mensajes de 3 vías.

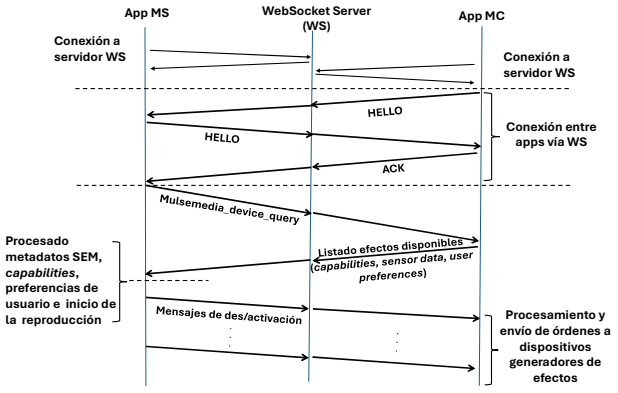


Fig. 7. Proceso de comunicación entre las apps MS y MC

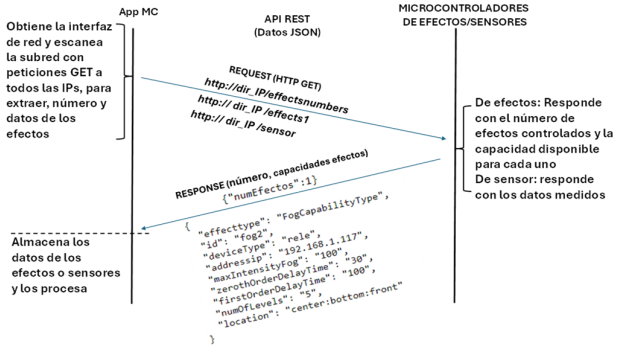


Fig. 8. Comunicación entre la app MC y los dispositivos microcontroladores

```
<cidl:ControlInfo
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="urn:mpeg:mpeg-v:2012:01-CIDL-NS CIDL.xsd"
    xmlns:cidl="urn:mpeg:mpegv:2012:01-CIDL-NS"
    xmlns:dcdv="urn:mpeg:mpeg-v:2012:01-DCDV-NS"
    xmlns:sapv="urn:mpeg:mpeg-v:2012:01-SAPV-NS"
    xmlns:scdv="urn:mpeg:mpeg-v:2012:01-SCDV-NS"
    xmlns:sepv="urn:mpeg:mpeg-v:2012:01-SEPV-NS"
    xmlns:mpegvct="urn:mpeg:mpeg-v:2012:01-CT-NS">

    <cidl:SensoryDeviceCapabilityList>
        <cidl:SensoryDeviceCapability xsi:type="SprayerCapabilityType"
            id="Sprayer2"
            sprayingType="urn:mpeg:mpeg-v:01-SI-SprayingTypeCS-NS:water"
            zerothOrderDelayTime= "5"
            firstOrderDelayTime= "5"
            maxIntensity="8"
            numOfLevels="3"
            locator="urn:mpeg:mpeg-v:01-SI-LocationCS-NS:centerleft:top:midway"/>
```

Fig. 9. Ejemplo mensaje *MPEG-V* con *capabilities* del efecto espray

IV.D. Descubrimiento de dispositivos

Al iniciarse, la app MC averigua la dirección de la subred IP a la que está conectada y envía una petición HTTP GET, con ciertos parámetros, a todas sus direcciones IP. Los dispositivos microcontroladores incluyen un servidor HTTP que implementa una API REST y, al recibir dicha petición, contestan con un mensaje HTTP 200 OK incluyendo datos en formato JSON del dispositivo que controlan, bien propios del efecto que controlan y sus características (*capabilities*) o bien del sensor que controlan. La Fig. 8 muestra el proceso (muestra datos JSON de respuesta de un microcontrolador de un efecto de niebla). A partir de las respuestas recibidas, la app MC crea mensajes XML (Fig. 9) compatibles *MPEG-V* con las características (*capabilities*) de cada uno de los sensores y de los efectos y sus dispositivos generadores en el escenario, y los envía a la app MS vía *WebSocket* cuando esta lo pida (Fig. 7). Este proceso se repite periódicamente pero sólo se envían de nuevo mensajes a la app MS cuando haya cambios.

IV.E. Sincronización

Se ha implementado un esquema de control de sincronización de tipo Maestro/Esclavo, tomando el reloj de la app MS como referencia, y la utilización del protocolo de *Round-Trip Synchronization* [30]. De forma periódica la app MC calcula el desfase entre su reloj y el de la app MS y corrige su propia referencia de reloj para sincronizarse con la app MS. La app MS envía los mensajes de activación a la app MC con suficiente antelación, teniendo en cuenta las características (*capabilities*) de los dispositivos generadores presentes en el escenario, con marcas de tiempo indicando el instante en el que se deben activar dichos dispositivos, para que el usuario perciba los efectos en el momento adecuado.

IV.F. Gestión de la activación de efectos sensoriales reales

Cuando la app MC recibe mensajes XML *MPEG-V* de activación de efectos de la app MS, los convierte en comandos de activación, en formato JSON, que se envían en mensajes

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

HTTP POST (haciendo uso de la API REST) en los instantes adecuados a los microcontroladores. La Fig. 10 muestra un ejemplo de mensaje de la app MS de activación de un efecto de frío y su traducción a comando JSON.

IV.G. Entorno o escenario mulsemedia real

Para poder evaluar la viabilidad de la propuesta, se ha implementado un prototipo que consiste en una recreación de un salón-comedor típico en el que un usuario podría tener una experiencia de consumo mulsemedia inmersiva. Se trata de un espacio de unos 3x3x3m, con un TV, un sofá enfrente de él a unos 2,5m del mismo, alrededor del cual se han colocado dispositivos generadores de efectos situados en diferentes posiciones (Fig. 11). Las características de los dispositivos se resumen en la Tabla I. Se han utilizado 2 miniPCs, uno como dispositivo principal (ejecutando la app MS) conectado al TV, y el otro como dispositivo *Controlador Mulsemedia* (ejecutando la app MC), servidor web, multimedia y de *WebSocket* (basados en Node.js). La Fig. 12 muestra el esquema general. Los microcontroladores utilizados son el ESP8266 con módulo de relés (efectos luz, flash niebla, espray y frío), ESP32 con módulo *dimmer* (efectos de viento y calor) o módulo amplificador y dos altavoces de bajas frecuencias (vibración). El generador de aromas Olorama™ dispone de una única interface Ethernet configurada por DHCP y un protocolo de configuración propietario basado mensajes de texto vía UDP/IP. En este caso, se ha utilizado una Raspberry Pi con el código adecuado, para, por un lado, asignar configuración IP (vía servidor local DHCP) al dispositivo, y, por otro, comunicarse con la app MC y con el dispositivo. Además, se ha incluido un sensor de la temperatura de la sala, controlado por un ESP32, para que la app MS la tenga en cuenta a la hora de activar efectos de temperatura decidiendo si enviar ordenes de activación de efecto de frío o de calor.

```
<?xml version="1.0"?>
<iidl:InteractionInfo xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:mpegvct="urn:mpeg:mpeg-v:2012:01-CT-NS"
xmlns:iidl="urn:mpeg:mpeg-v:2012:01-IIDL-NS"
xmlns:dcv="urn:mpeg:mpeg-v:2012:01-DCV-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2012:01-DCV-NS
http://standards.iso.org/ittf/PubliclyAvailableStandards/MPEG
V_schema_files/MPEG-V-DCV.xsd">
  <iidl:DeviceCommandList>
    <iidl:DeviceCommand xsi:type="dcv:CoolingType"
    deviceIdRef="cooling2" activate="true" intensity="62">
      <iidl:TimeStamp xsi:type="mpegvct:AbsoluteTimeType"
       absTime="1746436081218"/>
    </iidl:DeviceCommand>
```

```
{
  effectType: 'CoolingType',
  state: 'true',
  deviceIdRef: 'cooling2',
  intensity: '62'
}
```

Fig. 10. Conversión mensaje activación *MPEG-V* a comando JSON

Tabla I
CARACTERÍSTICAS DE LOS DISPOSITIVOS DEL ESCENARIO MULSEMEDIA

Dispositivo	Marca y modelo	Efecto
1 TV	Philips MT5802K17, con sistema de iluminación inteligente	Visual y auditivo
Bombillas	LED de diferentes colores	Visual
2 luces estroboscópicas	HQ-Power	Visual
2 ventiladores	Cata TB-40 y Raydan 20090	Viento
2 calefactores	Taurus, modelo Tropicano 2400	Calor
2 AACC portátil	Mundoclima, MUPO-12-H4	Frío
1 Sofá	2 plazas	Vibración
2 Generadores aromas	Olorama ™	Olfativo
1 sensor temperatura	AZDelivery DHT11 Breakout	-
2 MiniPCs	Intel NUC11TNHi5	-

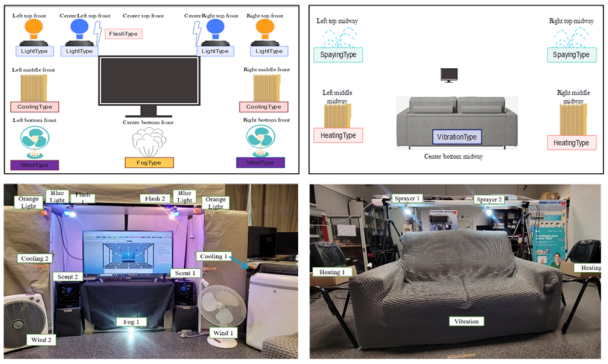


Fig. 11. Entorno mulsemedia

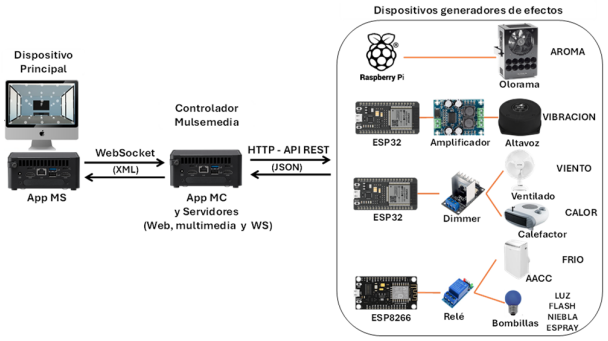


Fig. 12. Esquema general de la implementación realizada

IV.H. Pruebas de validación realizadas[4]

Con el único fin de validar el correcto funcionamiento del prototipo implementado y de las herramientas y aplicaciones desarrolladas, se realizó una evaluación muy preliminar, con 10 usuarios sin experiencia previa en el área mulsemedia (estudiantes). Como contenido se escogió el video '*Primavera*'[5], cortometraje de animación libre con una duración de 7 minutos y 44 segundos. Mediante *SEVino2*, y modificaciones manuales posteriores, se crearon metadatos SEM de todos los efectos y posiciones mostrados en la Fig. 11, que se activan en diferentes instantes específicos del vídeo. Los usuarios visualizaron la versión en MP4, con video 1920x1080, bitrate medio de 1452Kbits/s y 24 fps; y audio estéreo a 44,1 KHz. Primero visualizaron el vídeo sin efectos y después repitieron la sesión con efectos. Todos ellos afirmaron que habían percibido los efectos con la intensidad adecuada y de forma sincronizada con el contenido AV, y que la inclusión de los mismos mejoró la calidad de su experiencia de consumo. El efecto que más gustó fue el de aromas y el que menos el de luces (las bombillas no estaban debidamente camufladas ni colocadas en la posición más adecuada). En https://youtu.be/ae5KAQEftI0 se puede ver un prototipo muy preliminar de la plataforma en funcionamiento.

V. CONCLUSIONES

En el artículo se ha propuesto una plataforma mulsemedia con una arquitectura modular para ser implementada en entornos mulsemedia locales, basada el uso de dispositivos IoT de bajo coste para el control de dispositivos comerciales generadores de efectos reales existentes en dichos entornos. Incluye varias partes SW y HW reusables: una herramienta de

[4] Realizar una evaluación completa, tanto objetiva como subjetiva, de la plataforma está fuera del alcance de este trabajo y se plantea como trabajo futuro.

[5] https://www.youtube.com/watch?v=2JyGBG9dUmw (últ. acc.: septiembre 2025)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

reproducción mulsemedia y de simulación en 3D que incluye la gestión de las preferencias de los usuarios; una herramienta de gestión y control de la generación de efectos reales; dispositivos microcontroladores con el SW apropiado; dispositivos generadores de efectos reales; y dispositivos sensores. Es importante señalar que la arquitectura definida puede adaptarse y escalarse para integrar cualquier tipo adicional de dispositivos generadores de efectos, independientemente de su tipo y protocolo de comunicación. A modo de validación se ha implementado un prototipo escenario con dispositivos generadores de efectos comerciales, sin necesidad de desarrollar HW específico de generación de efectos. Se ha verificado su correcto funcionamiento además de corroborar las ventajas y aceptación por parte de los usuarios de la inclusión de efectos sensoriales en las experiencias multimedia, en cuanto a inmersividad y mejora de la QoE durante las sesiones. En el futuro se espera poder integrar mayor variedad de dispositivos y efectos y realizar evaluaciones objetivas (rendimiento, escalabilidad…) y subjetivas (QoE) más exhaustivas con grupos más numerosos y heterogéneos de usuarios.

AGRADECIMIENTOS

Esta publicación es parte de los proyectos de I+D+i con referencia PID2021-126645OB-I00, financiado por MICIU/AEI/10.13039/501100011033/ y "FEDER Una manera de hacer Europa", "FEDER/UE"; y ref. CIAICO/2022/025, financiado por la Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital de la Generalitat Valenciana (DOGV 8919/05.10.2020).

REFERENCIAS

[1] M. Waltl, B. Rainer, C. Timmerer, and H. Hellwagner, "A toolset for the authoring, simulation, and rendering of sensory experiences," in *MM 2012 - Proceedings of the 20th ACM International Conference on Multimedia*, 2012, pp. 1469–1472. doi: 10.1145/2393347.2396522.

[2] G. Ghinea, C. Timmerer, W. Lin, and S. R. Gulliver, "Mulsemedia: State of the art, perspectives, and challenges," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 11, Oct. 2014, doi: 10.1145/2617994.

[3] "ISO/IEC 23005. 2011. Information technology - Media context and control (MPEG-V)," 2011.

[4] L. Jalal, M. Anedda, V. Popescu, and M. Murroni, "Internet of Things for Enabling Multi Sensorial TV in Smart Home," in *2018 IEEE Broadcast Symposium, BTS 2018*, IEEE, Oct. 2018, pp. 1–5. doi: 10.1109/BTS.2018.8550959.

[5] A. Covaci, L. Zou, I. Tal, G. M. Muntean, and G. Ghinea, "Is multimedia multisensorial? - A review of mulsemedia systems," Aug. 01, 2018, *Association for Computing Machinery*. doi: 10.1145/3233774.

[6] C. Velasco and M. Obrist, "Multisensory Experiences: Where the senses meet techonology," *Oxford*, p. 112, 2020, doi: 10.1093/oso/9780198849629.001.0001.

[7] N. Murray, O. A. Ademoye, G. Ghinea, and G.-M. Muntean, "A Tutorial for Olfaction-Based Multisensorial Media Application Design and Evaluation," *ACM Comput Surv*, vol. 50, no. 5, pp. 1–30, Sep. 2017, doi: 10.1145/3108243.

[8] I. F. Akyildiz, H. Guo, R. Dai, and W. Gerstacker, "Mulsemedia Communication Research Challenges for Metaverse in 6G Wireless Systems," *ITU Journal on Future and Evolving Technologies*, vol. 4, no. 4, Dec. 2023.

[9] E. B. Saleme, A. Covaci, G. Mesfin, C. A. S. Santos, and G. Ghinea, "Mulsemedia DIY: A survey of devices and a tutorial for building your own mulsemedia environment," Jun. 01, 2019, *Association for Computing Machinery*. doi: 10.1145/3319853.

[10] J. Tewell and N. Ranasinghe, "A Review of Olfactory Display Designs for Virtual Reality Environments," *ACM Comput Surv*, vol. 56, no. 11, Jun. 2024, doi: 10.1145/3665243.

[11] A. C. da Silveira, E. C. Rodrigues, E. B. Saleme, A. Covaci, G. Ghinea, and C. A. S. Santos, "Thermal and wind devices for multisensory human-computer interaction: an overview," *Multimed Tools Appl*, vol. 82, no. 22, pp. 34485–34512, Sep. 2023, doi: 10.1007/S11042-023-14672-Y/TABLES/1.

[12] C. H. Muntean, I. Tal, D. Bogusevschi, M. Bratu, T. Bi, and G. M. Muntean, "Mulseplayer: A Multi-Sensorial Media Content Delivery Solution to Enhance End-User Quality of Experience," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, 2024, doi: 10.1109/BMSB62888.2024.10608351.

[13] D. P. De Mattos, D. C. Muchaluat-Saade, and G. Ghinea, "Beyond Multimedia Authoring: On the Need for Mulsemedia Authoring Tools," *ACM Comput Surv*, pp. 1–31, 2021.

[14] D. P. de Mattos and D. C. Muchaluat-Saade, "STEVE: a Hypermedia Authoring Tool based on the Simple Interactive Multimedia Model," in *DocEng '18: Proceedings of the ACM Symposium on Document Engineering 2018*, Halifax NS Canada, Aug. 2018, pp. 1–10.

[15] S-K. Kim, "Authoring multisensorial content," *Signal Process Image Commun*, vol. 28, no. 2, pp. 162–167, 2013.

[16] F. Danieau, J. Bernon, J. Fleureau, P. Guillotel, N. Mollet, and M. Christie, "H-Studio: An Authoring Tool for Adding Haptic and Motion Effects to Audiovisual Content," in *Proceedings of the adjunct publication of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 83–84.

[17] S. H. Shin, K. S. Ha, H. O. Yun, and Y. S. Nam, "Realistic media authoring tool based on MPEG-V international standard," in *International Conference on Ubiquitous and Future Networks, ICUFN*, IEEE, Jul. 2016, pp. 730–732. doi: 10.1109/ICUFN.2016.7537133.

[18] S-K. Kim, Y-S. Joo, and Y. Lee, "Sensible Media Simulation in an Automobile Application and Human Responses to Sensory Effects," *ETRI Journal*, vol. 35, no. 6, pp. 1001–1010, 2013.

[19] M. Josué *et al.*, "Modeling sensory effects as first-class entities in multimedia applications," *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, pp. 225–236, 2018, doi: 10.1145/3204949.3204967.

[20] D. P. De Mattos, D. C. Muchaluat-Saade, and G. Ghinea, "An Approach for Authoring Mulsemedia Documents Based on Events," *2020 International Conference on Computing, Networking and Communications, ICNC 2020*, pp. 273–277, Feb. 2020, doi: 10.1109/ICNC47757.2020.9049485.

[21] E. B. Saleme, C. A. S. Santos, and G. Ghinea, "Coping with the Challenges of Delivering Multiple Sensorial Media," *IEEE Multimedia*, pp. 1–1, 2018. doi: 10.1109/MMUL.2018.2873565.

[22] M. Waltl, B. Rainer, and H. Hellwagner, "An end-to-end tool chain for Sensory Experience based on MPEG-V," *Signal Process Image Commun*, vol. 28, no. 2, pp. 136–150, Feb. 2013, doi: 10.1016/J.IMAGE.2012.10.009.

[23] S. Bartocci, S. Betti, G. Marcone, M. Tabacchiera, F. Zanuccoli, and A. Chiari, "A novel multimedia-multisensorial 4D platform," *2015 AEIT International Annual Conference, AEIT 2015*, 2015, doi: 10.1109/AEIT.2015.7415215.

[24] K. Yoon, "End-to-end framework for 4-D broadcasting based on MPEG-V standard," *Signal Process Image Commun*, pp. 127–135, 2013.

[25] F. P. Luque, I. Galloso, C. Feijoo, C. A. Martín, and G. Cisneros, "Integration of Multisensorial Stimuli and Multimodal Interaction in a Hybrid 3DTV System," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 11, no. 1s, Oct. 2014, doi: 10.1145/2617992.

[26] S-K. Kim and M. Preda, "MPEG-V International Standard and the Internet of Things," in *International Conference on Software Intelligence Technologies and Applications & International Conference on Frontiers of Internet of Things 2014*, Hsinchu, 2014, pp. 166–173.

[27] L. Jalal, R. Puddu, M. Martini, V. Popescu, and M. Murroni, "A QoE Model for Mulsemedia TV in a Smart Home Environment," *IEEE Transactions on Broadcasting*, vol. 69, no. 1, pp. 179–190, Mar. 2023, doi: 10.1109/TBC.2022.3217676.

[28] L. Jalal, M. Anedda, V. Popescu, and M. Murroni, "QoE Assessment for IoT-Based Multi Sensorial Media Broadcasting," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 552–560, Jun. 2018, doi: 10.1109/TBC.2018.2823914.

[29] F. Boronat, E. Villashagñay, Ll. Simó, and J. González, "Simulador web mulsemedia en 3D," in *Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025)*, Cáceres (España), Nov. 2025, pp. 1–8.

[30] H. Melvin, J. Shannon, and K. Stanton, "Time, Frequency and Phase Synchronisation for Multimedia—Basics, Issues, Developments and Opportunities," in *MediaSync*, Cham: Springer International Publishing, 2018, pp. 105–146. doi: 10.1007/978-3-319-65840-7_4.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Platform for Gathering QoS-related Metrics for Training AI-based Media Synchronization Models

Juan González, Fernando Boronat.
Departamento de Comunicaciones,
Universitat Politècnica de València – Campus de Gandia
IIM R&D Group – Calle Paraninfo, 1, 46730, Grao de Gandia (Valencia-España)
juagons4@upv.edu.es, fboronat@dcom.upv.es

Abstract. This paper presents a novel web-based platform for gathering QoS-related metrics for training future AI-based Media Synchronization models. It allows two users to view the same media stream simultaneously, as well as to interact with each other through video and voice channels while those metrics are recorded. Existing network datasets are not suited for training AI models for media synchronization. The platform obtains metrics from the users' network connections and devices, allowing them to monitor the entire media session. The goal of this platform is to gather enough relevant metrics for training future AI models that will forecast and proactively adjust QoS parameters that help to improve the synchronization of the playout processes of the involved media streams in the session. Network simulation has been implemented for testing the platform with adverse conditions that could be found on the Internet.

***Palabras Clave*- Multimedia, Real-time, P2P, AI, Networking, IDMS**

I. Introduction

Shared media consumption experiences like Social-TV, synchronous e-learning and real-time online collaboration are becoming commonplace. The current research on network conditions forecasting for Quality of Service (QoS) management is mainly focused on routing, network security, and resource allocation, giving little attention to other mechanisms, like Inter-Destination Media Synchronization (IDMS) [1], which is critical for shared media applications. All the current IDMS solutions correct playout asynchronies after they are detected (reactive techniques). New proactive techniques that predict and avoid the appearance of such asynchronies should be further developed. The use of Artificial Intelligence (AI)-based methods would help to develop those techniques.

AI-based methods are currently being used for QoS predictions to overcome the limitations of reacting to events when they appear by anticipating them [2-5]. However, as far as the authors know, there are no solutions that can be applied for IDMS in shared media applications. Forecasting methods include Regression, Machine Learning (ML), and Moving Averages, among others. Recurrent Neural Networks (RNN) are a type of ML ideally preferred for modelling complex relationships and learning new trends over time thanks to Reinforced Learning (RL). Nevertheless, RNN require large datasets for the initial training.

Despite the existence of extensive datasets of networked applications [6-10], they are not appropriate for training models for shared media applications. The AI-based models need to know about media-related parameters (e.g., playout time or video resolution), along with network QoS parameters (e.g., end-to-end latency or packet loss), so they can identify complex relationships between the shared media and other QoS parameters. This way, mechanisms like IDMS solutions can be managed more precisely.

This article presents a web-based platform that gathers useful real-time metrics, storing them for subsequently training AI-based models to be developed and trained in the future. Because shared media applications require the synchronized playout of real-time media streams, the goal of the platform is to provide a testbed to develop and implement AI-based IDMS solutions once enough data has been gathered for training the developed models. Moreover, the platform simulates useful network QoS parameters, to be able to obtain enough data that reflect how different impairments affect the rest of parameters.

The paper contributes to the field with an accessible platform to gather metrics, with new media parameters. Additionally, the platform offers the following benefits:

- Customized QoS parameters: The streaming of media, communication and session control can be customized and optimized.
- Real-time monitoring: Metrics gathered are stored in a dataset with all the information from devices in a session (time series observations).
- Multiplatform: Compatible web standards have been used. End users only need a browser to use the platform, without any additional software tools
- Privacy: No identifiable information is captured.
- Extensible: It is a modular platform that can be adapted to multiple applications and with new metrics to monitor.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

The paper is structured as follows. The platform design is described in Section II. Then, in Section III, the data that can be gathered are explained. The paper finishes with the conclusions and future work, in Section IV.

II. Platform Design

As previously indicated, web technologies are employed for the purpose of being private, accessible and multiplatform, which also makes the development more agile and enables extending the platform.

The platform is divided into two parts: the Server, and the Client parts, which are made modular. Fig. 1 shows an implementation with one server and two clients, and how the Clients and Server devices are connected and how they communicate.

The Server part is built with Node.js [11] and consists of the following modules, each providing a different service to the clients:

- HTTP service: Offers the web files to the clients. It makes use of secure HTTP (HTTPS) for making the connections private.
- WebSocket service: Used for Server-Client communication and for receiving the Clients' metrics reports. It is also used by both Clients for exchanging control information during the establishment of a P2P communication channel between them.
- Global clock service: Based on NTP, it also acts as a time server for all the clients to provide them with a common global time reference. Clients will communicate with this module to ensure all follow the same global clock at the same time. This is important to be able to perform time analysis of the related monitored metrics and ensure the AI forecasts are correctly aligned to the time.
- Media streaming service: It provides video content to the clients by using the HTTP module.
- Metrics monitoring service: It stores in a database all the data included in the reports received from all the Clients. As mentioned before, the data will be used for training AI-based models and by those models to forecast the QoS parameters for dynamically managing the IDMS.
- AI-based Synchronization service: Not implemented yet. AI-based synchronization solutions will be included in this module that will forecast the evolution of relevant QoS parameters and subsequently will transmit synchronization control information to the Clients to adapt their playout processes to avoid the appearance of asynchrony between the clients' playout processes. Thanks to the clock sync, all data gathered is properly timestamped, and forecasts will work correctly, ensuring the AI Media Sync represent the appropriate time.

The web pages, which are served to the Clients, are built entirely with HTML, CSS, and JS. Fig. 2 shows the web page that the users view after they establish a shared media session with other users.

The users can select a video to watch at the same time, interact with each other by using the video/audioconference tool, and control the playback of the shared video. All these actions are communicated through the stablished Peer-to-Peer (P2P) channel [12]. So, when a user changes the video source or interacts with the playback controls, those actions are reflected in the other user's end. The Client part is comprised of the following modules:

- Connection module: In charge of creating and managing Client-Server and P2P communication. Also used by the Clients for sending metric reports to the server and for the clock sync.
- Media module: It manages the media streams of the users (webcams and microphones) and the shared media player (e. g., loading the video source, using playback controls).

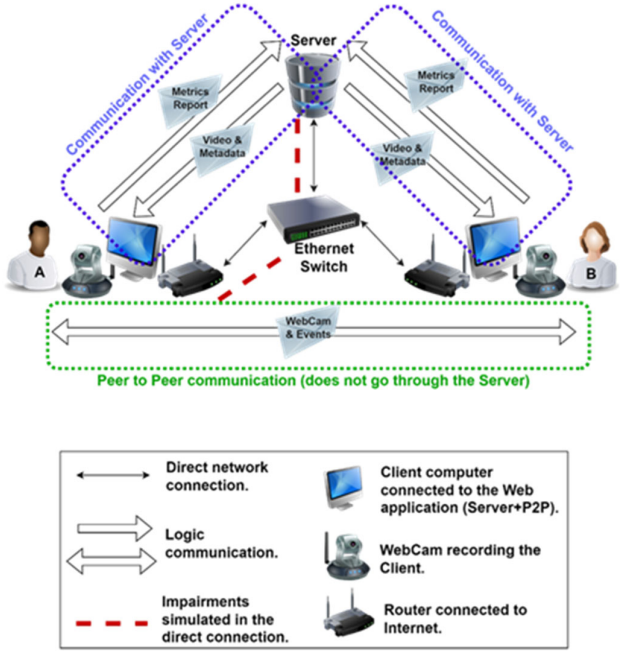


Fig. 1. General scheme of connections and communication of devices

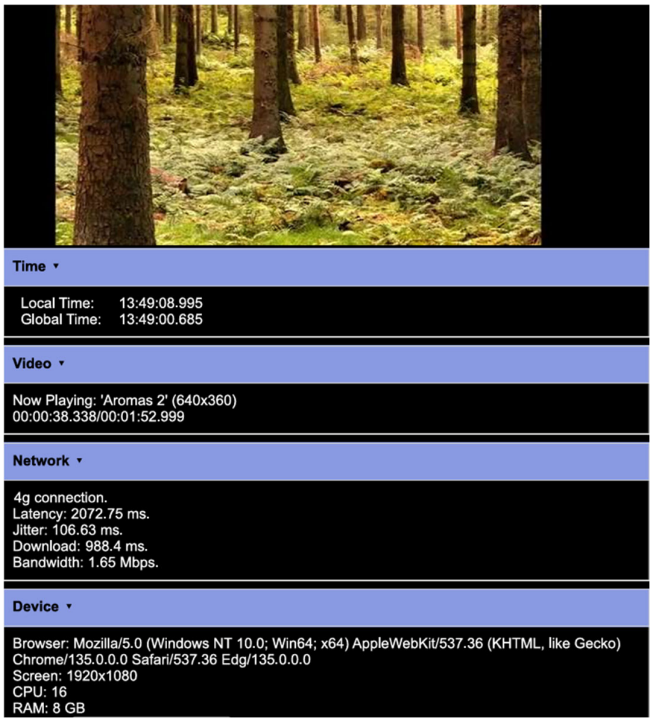


Fig. 2. Site design. Includes the video player with controls, some of the monitored metrics being displayed

- Monitor module: For configuring and monitoring all the events and states that are captured or measured during the session.
- Controller module: It combines the logic of the modules and the graphical interface of the site (e.g., running logic when a button is pressed).

A. HTTP-based Adaptive Streaming of the Shared Media

The shared media is transmitted from the Server by using HTTP-based streaming. The video files have been previously segmented into fragments with multiple qualities, providing an index file. This way, the videos are streamed progressively, and the video quality adapts to the available bandwidth.

For the playback on the Client web browser, the Video.js player [13] library has been used, which supports MP4, HLS, DASH, among other formats. Therefore, the clients can play videos from the platform's media server and from other domains (external servers).

The employed media client monitors metrics from its connection and from (media) data and playout states. This part is further explained in the third section of this paper.

B. WebRTC Peer-to-Peer Communication

For establishing a P2P audio and video connection between users, browsers support the WebRTC API [14], which does not require any plugin, only a signalling server (the platform's WebSocket server in this case) to support its initial signalling process so the clients (peers) can discover each other and establish the P2P communication.

For starting the P2P connection, the clients communicate with the signaling server with WebSockets [15], then use the MediaDevices [16] and RTCPeerConnection [17] APIs to get the audio/video streams from one peer's device and transmit them in real-time to the other peer. To simplify the design, the platform uses the PeerJS library [18] that wraps the WebSocket and WebRTC together.

WebRTC also offers the getStats method [19], to obtain metrics about the P2P connections between clients. This part is further explained in the third section of this paper.

C. Network Simulation

For simulating the adverse network conditions, the clumsy tool [20] for packet interception has been utilized in each device. It can hold, drop, and alter network packets to simulate different network conditions regarding network delay, limited bandwidth, packet loss, packet duplication, packet reordering, throttling and other network QoS parameters. In the platform, the simulation has been programmed to dynamically change these parameters, so it can simulate varying network conditions changing over time, as could happen with a real Internet connection.

III. Communication Monitoring

There are four main groups in which the collected metrics have been divided: Network, Peer, Device and Player. There is an additional Simulation group, with the metrics that represent the network parameters that are simulated by the simulator tool. The main four groups, in contrast, are obtained from the actual measure process (e.g., pinging to estimate the network delay).

Outside of testing and research scenarios, it is not usual to have access to the precise state of the network (e.g., Internet topology). Therefore, the parameters from the simulator will probably not be available in real applications. Nonetheless, it is important to have them, so that AI-models can also be trained to estimate them.

The Network group contains the metrics of the client's network connection (i.e., Internet). The Peer group contains the metrics regarding the WebRTC (P2P) connection. The Device group contains the metrics about the client's device, as well as the browser it is using. The Player group contains the metrics related to the video player (Video.js).

Because the WebRTC native getStats method returns the different metrics separately (e.g., codecs, channels, streams, etc.), there is a need for relating and grouping them, so the data can be clearly understood and in their context. For example, all the streams of the same connection can be associated together by using the proper identifiers. The WebRTCStats library [21] employed already leverages this task.

A. Gathered Metrics

The Metrics Monitoring module offers the ability to configure different capture rates for all the metrics in each group. For example, the metrics in the Player group may change faster than the ones in the Network group. A configuration file is used to individually enable or disable the metrics to capture. Additionally, the frequency at which the metrics reports are sent to the server by the clients can be changed as well. The following is a list of the measures taken with the aim of gathering enough and proper data for developing and training future AI-based models that help with the predictive feature of new proactive IDMS solutions, along with the units used to measure them:

1. **Network parameters:**
   a) Estimated Latency (milliseconds, ms).
   b) Estimated Upload Bandwidth (Megabytes, MB).
   c) Estimated Download Bandwidth (MB).
2. **Peer parameters:**
   a) Estimated Round Trip Time (ms).
   b) Transmission throughput (kilobits/second, kbps).
   c) Reception throughput (kbps).
   d) Packets per second transmitted (packets/second, pps).
   e) Packets per second received (pps).
   f) Current camera capture time (timestamp).
   g) Current camera playout time (timestamp).
3. **Device parameters:**
   a) User Identifier in Session (Text)[1].
   b) Number of CPU cores[1].
   c) Base speed of CPU (GHz)[1].
   d) Available RAM (GB)[1].
   e) Used RAM (%).
   f) Machine GPU Model (Text)[1].
   g) Drive Type (either HDD or SSD)[1].
   h) Connection Type (2g, 3g, 4g, wifi, ethernet).
   i) Battery Level Remaining (%).
   j) Estimated Browser Computing (Million operations/s).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

4. **Player parameters:**
   a) Current video playout time (ms).
   b) Current video duration (ms)[2].
   c) Content in buffer (ms).
   d) Content successfully played (ms).
   e) Number of frames dropped or corrupted (frames).
   f) Number of total frames available (frames).
   g) Video resolution (Width×Height, in pixels)[2].
   h) Media Readiness State (Possible values in [22]).
   i) Media Fetching State (Possible values in [23]).
   j) Video Playout Rate (Frames/second, fps).
   k) Current video bitrate (kbps)[2].
   l) Video Changed[3].
   m) Video Played[3].
   n) Video Paused[3].
   o) Video Seek to Another Time[3].
   p) Video Playout Reached its End[3].
   q) Playout Delay between Users (ms).
5. **Simulation parameters:**
   a) Network delay (ms).
   b) Bandwidth Available (MB).
   c) Packet Drop Rate (%).
   d) Packet Duplication Rate (%).
   e) Packet Reordering Rate (%).
   f) Packet Throttling Rate (%).
   g) Packet Throttling Duration (ms).
   h) Packet Tampering Rate (%).
   i) Simulation on Transmission (Boolean).
   j) Simulation on Reception (Boolean).

Some parameters stay unchanged during the same session (marked with 1). Others change when the source of the shared media changes (2). And some others indicate discrete events (3), and are related to which user triggered them, when caused by one.

These parameters are identified and stored for each device in the session. For instance, the obtained dataset includes a list of rows with the timestamps of the measures and the metrics from the clients and server obtained in that same instant. This way, the current video playout points of the different users can be contrasted to find the delays and be analysed together compared with the parameters of the rest of devices over time. More information about all the available parameters can be found in the respective official documentation [13],[16],[19].

B. Reactive Synchronization

For the task of ensuring that the playout points of all clients are synchronized in real time, the platform's server can send a message to the clients with playout reference information, and the clients can use that information to update their playback point and rate accordingly. This way, when asynchrony situations have been forecasted, the playout can be adjusted dynamically and in advance instead of reactively. So, metrics before and after IDMS is implemented can be gathered for comparison.

The proposed compensation algorithm for synchronizing delayed playout between destinations consists of a smooth playout adjustment (i.e., AMP or Adaptive Media Playout) that consists of varying the playout rate (i.e., slowing down or speeding up the rate of advanced or lagged players, respectively) when playout differences between users' players exceed a configured threshold, so they eventually become synchronized [1],[24]. AMP adjustments are less noticeable by users than aggressive adjustments, such as skips or pauses in the playout process.

Although the playout rate can be modified, there may be instances when the delay is too high, either requiring a higher playback rate change that would be noticeable, or an aggressive measure (skip or pause). To avoid those cases, forecasting future events (e.g., network conditions or playout state changes) enables the dynamic playout correction to work in advance, preventing those situations from arising, which is the main objective for future research, after gathering metrics and training AI-based synchronization models.

C. Evaluation

The platform has been tested to confirm its correct performance. Metrics have been gathered for posterior steps of analysing and training AI-based models of time-series forecast.

Fig. 3 illustrates the scheme of simulating different network conditions on a single device (Client A) for only the outgoing traffic. More (or all) devices can also be running the simulator. The simulations in a device will affect its direct communication.

Currently, a few samples during a single session have been taken from three different desktop computers, but more metrics are being generated for having a larger dataset, so more information and insights will be obtained. Table I contains only a set of metrics and samples for the purpose of illustration.

Network conditions (high latency, low latency, random packet losses, and so on) were simulated in the different devices, so that the gathered metrics represent different internet scenarios of network congestion. This way, the models should work in forecasting the media async related to the different states of a network.

AI models will be trained using the collected data. The models need to perform fast in real-time applications and Moving Averages (MA) have been used in the past [4],[5]. However, MA works well with few variables and linear relationship. Reinforced Learning (RL) AI techniques like Long Short-Term Memory (LSTM) work well for forecasting multiple related variables, even when relationships are not strong (e.g., low correlation) [3]. LSTM techniques require a lot of samples and high computing power for training them. Support Vector Machine (SVM) is a supervised AI technique that can separate the samples into classes based on their multiple values [2], improving the forecasts performance.

Moreover, forecasting playout async allows two methodologies: binary forecast (whether future playout will surpass a predefined tolerable threshold) and regressive forecast (predicting the value, in milliseconds, of async). Besides the forecasts, the models can also output confidence values (the probability the forecasts are right, or expectable deviation), and other explainable insights (to understand the basis of how the AI models determined their specific forecasts). All these aspects will be studied in future work.

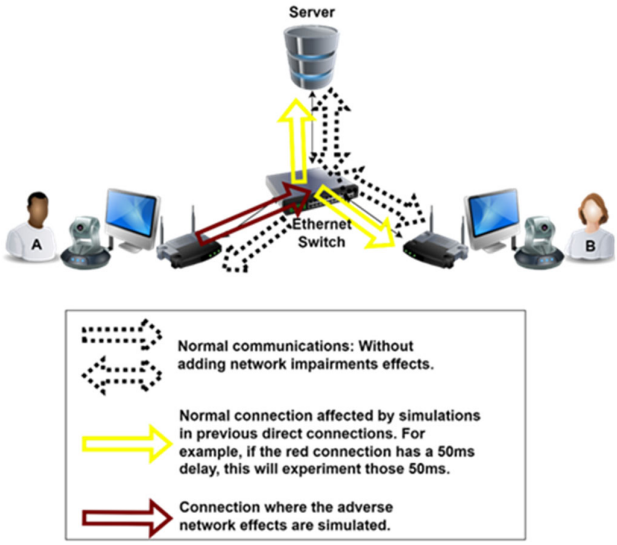


Fig. 3. Simulation example. Only one client simulating outgoing effects

Table I
EXAMPLE OF SOME FEW COLLECTED METRICS

Timestamp	Playout Time (ms)	Seek Event	Buffered Content (ms)	Latency (ms)
23:10.513	66430		29593	223
23:15.006	32343	User B	5990	226
23:19.507	36423		39990	227
23:24.005	41012		58583	226
23:28.505	45603		58583	226
23:33.005	72534	User A	6982	222
23:37.509	52968		18096	232
23:42.007	57559		40124	226

D. Data analysis

The preliminary collected data has been analysed to get some insights. Data gathered were cleaned and analysed.

Table II shows the variables mostly correlated with the playout asynchronies between users (i.e., with media sync). This could be used to select the variables for training AI-based models for estimating when media sync techniques should be triggered. Examples of such techniques are correcting user's playout (to avoid increasing the asynchrony) or modifying QoS parameters (like video quality) to prioritize the media playout and prevent future asynchrony situations.

Table II
CORRELATION OF VARIABLES WITH THE PLAYOUT DELAY

Variable	Correlation
Content in buffer	19.12%
Estimated latency	9.66%
Estimated download bandwidth	8.63%
Network latency variation	7.34%
Download delay	5.18%
Frames dropped or corrupted	2.42%
Used RAM	2.07%

Data showed some expected results, for example that network latency and video download are related to the appearance of playout async. However, non-linear relationship, and temporal relationships could be exploited by potential AI models, like Deep Learning models, which would adapt to the complex data and give more accurate results that non-AI solutions would not.

Another observation is that the reduced precision of measurements, compared to a native application, could introduce noise to the models, impairing their accuracy. To prepare for this, the data obtained until the moment does not show any indication that those imprecisions would impact the quality of the data, as the magnitudes (e.g., tens or hundreds of milliseconds) are found to be lower than the measurement resolutions (e.g., clock drift).

Further, very interesting insights are expected when more data are gathered and studied in the future.

IV. CONCLUSION AND FUTURE WORK

The presented platform is an effort to have a common space for generating datasets and for implementing and testing AI-based techniques for shared multimedia applications with the objective of developing and testing novel proactive synchronization solutions. The platform includes Client-Server and P2P communications, so techniques that synchronize both types of stream channels can be implemented and tested later.

Testing the platform with non-optimal network conditions is important for discovering performance issues. As the datasets contain samples over time, they can be used for time series forecasting, enabling them to develop proactive solutions rather than reactive solutions.

The platform in its current state has some limitations that will be overcome in the future. The act of taking measures, for instance, when monitoring, may impact on the observed results. Therefore, the gathering of metrics should be ensured to add low overhead. Finding strong correlations and which metrics are more important, so the other ones are not monitored, will help too. Once a completer and more definitive version of the platform has been completed, the source code will be made publicly available.

As future work, when a completer and more useful dataset is generated, it will be used to train an RNN model for being used in developing proactive IDMS solutions to be integrated and tested in the platform and to compare their performance with current well-known reactive IDMS solutions.

ACKNOWLEDGEMENTS

This work was supported, in part, by the following grants: ACIF/2021/192 from Programa I+D+i de la Generalitat Valenciana (GVA) and the European Social Fund (ESF); CIAICO/2022/025, from Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital of the GVA (DOGV 8919/05.10.2020); and grant PID2021-126645OB-I00, funded by MICIU/AEI/10.13039/501100011033/ and by "ERDF A way of making Europe".

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

REFERENCES

[1] M. Montagud and F. Boronat, "Enhanced adaptive RTCP-based Inter-Destination Multimedia Synchronization approach for distributed applications," in Computer Networks, vol. 56, no. 12, pp. 2912-2933, 2012, doi: 10.1016/j.comnet.2012.05.003.

[2] Z. Liu, Q. Z. Sheng, X. Xu, D. Chu and W. E. Zhang, "Context-Aware and Adaptive QoS Prediction for Mobile Edge Computing Services," in IEEE Transactions on Services Computing, vol. 15, no. 1, pp. 400-413, 1 Jan.-Feb. 2022, doi: 10.1109/TSC.2019.2944596.

[3] X. Wang, et al., "Deep Learning-Based Network Traffic Prediction for Secure Backbone Networks in Internet of Vehicles," in ACM Transactions on Internet Technology, vol. 22, no. 4, Article 87, November 2022, doi: 10.1145/3433548.

[4] A. Bernal, M. Richart, M. Ruiz, A. Castro, and L. Velasco, "Near real-time estimation of end-to-end performance in converged fixed-mobile networks," in Computer Communications, vol. 150, pp. 393-404, 2020, doi: 10.1016/j.comcom.2019.11.052.

[5] D. H. L. Oliveira, T. P. de Araujo and R. L. Gomes, "An Adaptive Forecasting Model for Slice Allocation in Softwarized Networks," in IEEE Transactions on Network and Service Management, vol. 18, no. 1, pp. 94-103, March 2021, doi: 10.1109/TNSM.2021.3055174.

[6] "LAN Network Stability," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/garystafford/ping-data. [Accessed: may 2025].

[7] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, 2015, pp. 1-6, doi: 10.1109/MilCIS.2015.7348942.

[8] R. Zhu, B. Liu, D. Niu, Z. Li, and H. V. Zhao, "Network Latency Estimation for Personal Devices: A Matrix Completion Approach," in IEEE/ACM Transactions on Networking, vol. 25, pp. 724-737, April 2017, doi: 10.1109/TNET.2016.2612695.

[9] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," in Future Generation Computer Systems, vol. 100, pp. 779-796, 2019, doi: 10.1016/j.future.2019.05.041.

[10] N. Moustafa, "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets," in Sustainable Cities and Society, vol. 72, 2021. doi: 10.1016/j.scs.2021.102994.

[11] "Node.js JavaScript runtime environment," GitHub, [Online]. Available: https://github.com/nodejs/node/. [Accessed: may 2025].

[12] "WebRTC RTCDataChannel Interface," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/RTCDataChannel/. [Accessed: may 2025].

[13] "Video.js web video player," [Online]. Available: http://videojs.com/. [Accessed: may 2025].

[14] "WebRTC: Real-Time Communication in Browsers," W3C, [Online]. Available: https://www.w3.org/TR/webrtc/. [Accessed: may 2025].

[15] "The WebSocket API," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API/. [Accessed: may 2025].

[16] "MediaDevices," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/MediaDevices/getUserMedia/. [Accessed: may 2025].

[17] "WebRTC RTCPeerConnection Interface," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/RTCPeerConnection/. [Accessed: may 2025].

[18] "PeerJS Server," GitHub, [Online]. Available: https://github.com/peers/peerjs-server/. [Accessed: may 2025].

[19] "WebRTC's Statistics API," W3C, [Online]. Available: https://w3c.github.io/webrtc-stats/. [Accessed: may 2025].

[20] "clumsy 0.3," GitHub, [Online]. Available: https://github.com/jagt/clumsy/. [Accessed: may 2025].

[21] "WebRTCStats Library," GitHub, [Online]. Available: https://github.com/peermetrics/webrtc-stats/. [Accessed: may 2025].

[22] "readyState property," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/HTMLMediaElement/readyState/. [Accessed: may 2025].

[23] "networkState property," Mozilla, [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/HTMLMediaElement/networkState/. [Accessed: may 2025].

[24] M. Montagud, F. Boronat, B. Roig, and A. Sapena, "How to perform AMP? Cubic adjustments for improving the QoE," in Computer Communications, vol. 103, pp. 61-73, 2017, doi: 10.1016/j.comcom.2017.01.017.

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Analysing the Effectiveness of CRM Processes in Acceleration Services: The Business process of Research Infrastructure Catalogue in Unite! Agora platform

Muhammad Anas Alghamian, Farnaz Haji Mohammadali, Jesus Alcober, Juan Lopez Rubio and Toni Oller
Network Engineering,
Universitat Politècnica de Cataluny
Castelldefels, Spain.
muhammad.anas.alghamian@upc.edu, farnaz.haji.mohammadali@upc.edu, jesus.alcober@upc.edu, juan.lopez-rubio@upc.edu and antoni.oller@upc.edu

This paper explores the vital role of Customer Relationship Management (CRM) processes in supporting acceleration services within Higher Education Institutions (HEIs). It focuses specifically on the "Research Infrastructure" service offered through the Unite! Agora platform and examines its impact on collaboration across university alliances. By analysing the CRM processes that support these services, the study highlights key benefits such as improved service management, enhanced institutional collaboration, and strengthened technological capacity within HEIs.

The paper also outlines the methodologies and best practices required for successful implementation of these services. It emphasizes the use of ITIL v4 as a guiding framework to increase process maturity, align with industry standards, and enable strategic decision-making and continuous improvement within the Agora platform.

The main contributions of this paper are twofold. First, it validates the effectiveness of ITIL v4 as a comprehensive framework for structuring, managing, and improving service request processes, highlighting its adaptability to academic and research-focused contexts. Second, it offers a practical model for managing research infrastructure services in alignment with ITIL v4's principles and workflows, providing actionable insights for HEIs to foster collaboration, enhance access to research infrastructure, and advance technological capabilities through a structured approach that supports continuous improvement.

***Keywords*- Business process, Agora, Acceleration services.**

I. INTRODUCTION

The University Partnership for the Acceleration of European Universities (aUPaEU[1]) project main objective is to address challenges in the research and knowledge sharing in HEIs by creating acceleration services for university alliances in Europe. As part of its strategy, aUPaEU[2] , has developed an innovative platform that brings together key services, such as sharing information, training and coaching services, and practical tools. This platform acts as a coaching service designed for different stakeholders, including scientists, industry experts, and social organisations. It allows stakeholders to choose from a range of acceleration services available at the Agora.

An acceleration service is a structured offering, typically delivered through a digital platform, designed to support HEIs across Europe in advancing their institutional transformation, particularly in terms of digitalization, cross-institutional collaboration, and alignment with European Research and Education policy goals. These services play a vital role in enabling HEIs to adapt to evolving challenges in the education sector and enhance their impact on knowledge creation and innovation. Furthermore, they align with key European priorities, including the European Research Area (ERA) and the European Education Area (EEA).

One of the main services on this platform is the "Research Infrastructure catalogue " which provides important resources and support for researchers and scientists. Our study will closely examine the CRM processes involved in these

[1] https://cordis.europa.eu/project/id/101095314/reporting last accessed 03 May 2025

[2] https://aupaeu.widening.eu/ last accessed 03 May 2025

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

services. We aim to identify and use the best management practices to improve service delivery, optimize resource utilization and deliver stronger support for researchers through the platform..

The Research Infrastructure Catalogue is a centralized digital inventory of laboratories, facilities, scientific equipment, and related resources available across the member universities. And it's considered an acceleration service because it streamlines and enhances access to valuable scientific resources, helping HEIs and researchers advance more quickly in their work. By making high-quality infrastructures discoverable and accessible through a central digital platform, the service reduces barriers to collaboration and innovation. Researchers no longer need to spend time searching for facilities or negotiating access, as the catalogue provides clear, organized information and direct contact options. The openness of these infrastructures to external users further promotes cross-institutional partnerships, aligning with European goals. Additionally, the integration of tools like automated contact forms and CRM tracking ensures a smooth and efficient process from inquiry to follow-up. In this way, the catalogue not only connects users to physical infrastructures but also accelerates research, innovation, and institutional transformation across the university alliance.

In many industries CRM projects often fail because people do not fully understand what a CRM is, how it works, and the technology behind it[1]. It is important to understand the key parts of CRM which are : people, technology, and processes[2]. It's important to know how CRM processes work and helps identify the changes needed and ensures they fit well with other business activities[4], Also important to see the benefits of CRM and measure how it affects the organisation's performance.

In this paper, we will closely look at the services already in place within the Unite! Agora[3]. We will explain how these services work and identify the best ways to manage them. Then, we will carefully review these services using the suggested management framework.

We are focusing on the Research Infrastructure service within Unite! Agora. We will explain how this service works, how it is currently managed, and the specific methods used. We will also discuss the best management practices that have been put in place to ensure it runs smoothly.

After detailing the service and how it is managed, we will assess how effective these practices are using the proposed management framework. This will help us understand the impact and efficiency of the services in Unite! Agora. While our focus is on Unite! Agora, we will also consider the larger aUPaEU initiative, which supports open science and the development of the Acceleration services of Agora.

Through this study, we aim to provide insights into how aUPaEU is driving research and innovation in European universities. Unite! The Agora is the digital collaboration platform of Unite!, serving as a central space where partner institutions share services, infrastructures, and tools developed by aUPaEU Project for the University Alliance for Innovation and Technology (Unite![4]), which connects nine partner universities across Europe. Its purpose is to foster collaboration in areas such as research, education, and innovation by providing tools and resources for knowledge sharing and networking. The Agora platform focuses on enabling easy communication and cooperation among member institutions, researchers, and students. It supports activities like sharing research infrastructures, promoting joint initiatives, and offering acceleration services to enhance institutional collaboration and innovation.

ITIL v4 best practices have been used in this article, ITIL v4 is the world's most widely adopted framework for IT service management (ITSM). Launched in February 2019, ITIL v4 is the result of extensive global collaboration among IT and service management experts, including practitioners, trainers, consultants, and vendors. The framework originated in 1989 through the UK's Central Computer and Telecommunications Agency (CCTA), which aimed to standardise IT services and practices. ITIL v4 is built on four dimensions (Organizations and People, Information and Technology, Partners and Suppliers, Value Streams and Processes) and is guided by seven principles:

- Focus on Value
- Start Where You Are
- Progress Iteratively with Feedback
- Collaborate and Promote Visibility
- Think and Work Holistically
- Keep It Simple and Practical
- Optimise and Automate ITIL v4 also includes 34 practices organised into three domains: General Management, Service Management, and Technology Management (AXELOS[5], 2019).

While this study initially frames its investigation in the context of CRM processes in acceleration services, it specifically focuses on the Research Infrastructure service as a representative CRM-related process within the Unite! Agora platform. The evaluation applies ITIL v4's Service Request Management framework to analyze how effectively this individual service is managed, with the intention of providing a model that can later be extended to other CRM processes within the platform.

A. ITIL v4 Best Practices and Their Relevance to the Evaluation

This study is grounded in the principles and practices of ITIL v4, the latest iteration of the globally recognized framework for ITSM. ITIL v4 introduces a holistic approach to service management, built around the Service Value System (SVS) and four dimensions Organizations and People, Information and Technology, Partners and Suppliers, and Value Streams and Processes. It emphasizes flexibility, collaboration, and continuous improvement through its seven guiding principles, including "Focus on Value," "Collaborate and Promote Visibility," and "Progress Iteratively with Feedback."

The Service Request Management practice from ITIL v4 is of particular relevance to this paper. This practice is designed to support the effective handling of user requests for services, information, or standard changes. It is not concerned with unplanned service interruptions (incidents) but focuses

[3] https://agora.unite-university.eu/ last accessed 03 May 2025

[4] https://www.unite-university.eu/ last accessed 03 May 2025

[5] https://www.axelos.com/certifications/itil-service-managementlast accessed 03 May 2025

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

on repeatable requests that can be handled with predefined workflows. These include tasks such as accessing infrastructure, requesting documentation, or booking facilities—all of which align with the services offered through the Unite! Agora platform.

By adopting this specific practice, the evaluation in this study assesses how well the current processes within the "Research Infrastructure" service conform to ITIL v4 standards. The practice ensures that requests are captured, categorized, fulfilled, and closed in a structured and efficient manner, supported by defined roles and metrics such as resolution time, first-time resolution rate, customer satisfaction, and process compliance.

The use of ITIL v4 not only provides a structured methodology for assessing service effectiveness, but also supports strategic objectives such as process maturity improvement, standardization, stakeholder alignment, and evidence-based continuous improvement. Embedding ITIL v4 into the evaluation framework ensures that the assessment reflects internationally recognized best practices, offering valuable insights for the enhancement of CRM-related processes within acceleration services for Higher Education Institutions..

B. Overview of Customer Relationship Management

Strong customer service on modern CRM systems treats each customer individually and unique, offering personalised service and strong support. Technology helps understand customer needs and respond quickly [2].

According to previous research, organisational resources have the greatest positive effect on most CRM stages, focusing on the need for a company structure that supports competitive advantage. Human resources are particularly important in the early and middle stages, highlighting the benefit of developing employee skills[5].

CRM performance has influence on knowledge management capabilities across customer life cycle management phases. Studies show each knowledge management capability impacts different customer life cycle management phases, affecting customer satisfaction and loyalty, providing valuable insights for improving CRM[5].

Customer loyalty is a key goal in CRM. However, acting early when customers first show signs of losing interest is more effective[7]. Customer satisfaction acts as a bridge: it partly connects vision to customer retention and fully links organisational collaboration to customer retention[5].

In order to succeed, organizations should focus on leadership development,understanding CRM's benefits, forming cross functional teams, implementing change management, and providing comprehensive training. Additionally, appointing a dedicated CRM leader is crucial[2].

Beyond general CRM adoption, studies in the higher education context emphasize the importance of IT service management maturity for institutional effectiveness. Rubio (2022) provides a quantitative analysis of IT management practices in educational institutions, showing that ITIL-aligned processes remain in early stages of adoption. This highlights the need for structured maturity models to guide universities toward improved service quality and standardization [3].

C. Overview of the Research Infrastructures catalogue

One of the current services in Unite! Agora is the Research Infrastructures[6] which provides access to advanced scientific resources and facilities. The **Research Infrastructures Catalogue** in Unite! Agora is a digital tool designed to provide centralized, structured access to scientific facilities and resources across partner institutions within the alliance. It serves as a strategic acceleration service aimed at enhancing research capabilities and fostering collaboration among universities. The catalogue includes a curated list of infrastructures that are open to both internal and external users, supporting interdisciplinary research and innovation. These infrastructures span a wide range of scientific domains from acoustics and electronics to additive manufacturing and neuroimaging making them valuable assets for researchers, educators, and innovators across Europe. By improving visibility and accessibility, the catalogue aligns with broader European priorities .

The platform is designed for ease of use, offering intuitive navigation and detailed entries for each infrastructure. Each listing includes information about the facility's capabilities, scientific fields, associated keywords, and the hosting institution. Users can filter and search through the catalogue to find infrastructures that match their research needs. A direct **contact button** on each entry opens a form, allowing researchers to send inquiries effortlessly. Behind the scenes, a **CRM system** tracks these interactions to ensure smooth follow-up and measure user engagement. This integrated digital approach reduces administrative burden, accelerates the matchmaking process between researchers and facilities, and ensures that no opportunity for collaboration or discovery is missed. Overall, the catalogue not only connects people with physical infrastructure but also acts as a service that accelerates research, fosters innovation, and supports the digital transformation of higher education institutions across the Unite! Alliance.

This paper is structured in a way that Section 1 provides an overview of ITIL v4 best practices and explains their relevance to the evaluation of service processes. Section 2 reviews the fundamental concepts of CRM and their application in acceleration services. Section 3 outlines the Research Infrastructure service within the Unite! Agora platform. Section 4 details the methodology used for evaluating CRM processes based on ITIL v4. Section 5 presents and discusses the findings of the evaluation. Finally, Section 6 concludes the paper by summarizing key insights and implications for future improvements.

II. METHODOLOGY

The methodology for this paper involves analyzing the "Research Infrastructure" service within the Unite! Agora platform by mapping its workflow from initiation to closure, detailing stakeholder roles, and documenting current processes. The study exclusively adopts ITIL v4 best practices, with a particular focus on Service Request Management.

[6] https://agora.unite-university.eu/resources/infrastructure last accessed 03 May 2025

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

The Service Request Management practice in ITIL v4 ensures that requests from users such as information inquiries, resource access, or standard changes are handled effectively and efficiently. The methodology evaluates the efficiency and impact of these services through Key Performance Indicators (KPIs), including request acknowledgment time, resolution time, customer satisfaction, and process compliance rates.

User feedback is analyzed to foster continuous improvement, and pre- and post implementation metrics are compared to assess the effectiveness of ITIL v4 practices. The process lifecycle for Service Request Management is structured around ITIL v4's principles and workflows, ensuring alignment with the framework's guidance for streamlined request handling and value delivery.

III. Discussion

This study focuses on the Infrastructure Service within the Unite! Agora platform and the key methods and best practices for its successful implementation. By evaluating the effectiveness and impact of ITIL v4 practices, the study provides valuable insights to support strategic decision-making and promote continuous improvement within the platform.

Building on previous research into improving process maturity and alignment with best practices, this study adopts ITIL v4 as the exclusive framework for guiding Service Management. ITIL v4's Service Request Management practice serves as the foundation for mapping and optimizing Service Request workflows. This practice ensures that service requests are handled efficiently, from initiation to closure, while maintaining alignment with business objectives.

The process lifecycle is mapped based on ITIL v4's comprehensive guidance, which emphasizes value delivery, stakeholder collaboration, and iterative improvements. The mapping highlights the structured approach of ITIL v4, ensuring that all stages of the service request lifecycle are systematically managed and controlled.

By leveraging ITIL v4 principles and practices, the study aims to enhance process efficiency, improve service quality, and foster operational excellence. This exclusive focus on ITIL v4 provides a streamlined methodology for achieving Service Request objectives effectively while ensuring alignment with internationally recognized best practices.

The purpose of this evaluation is not only to map the existing service process and align it with ITIL v4 best practices but also to identify areas for improvement within the current service execution model. The analysis aims to highlight potential inefficiencies, gaps in role allocation, and delays in key performance stages such as request acknowledgement, assignment, and closure. While this study does not yet incorporate empirical execution data from event logs, it establishes a baseline model for how such data could be integrated in future iterations of the evaluation. This includes the use of process mining techniques or performance dashboards to monitor compliance with KPIs over time. The recommendations derived are based on theoretical alignment with ITIL v4 practices and gaps identified through process inspection.

Stage	**Description**	**Role**
Initiation	The Researcher identifies a need and initiates a service request through the predefined channel.	Researcher
Fulfilment	The Institutional Coordinator logs the service request into the system, assigning a unique identifier for tracking.	Institutional Coordinator
	The Institutional Coordinator categorises and prioritises the service request based on its type, urgency, and impact.	Institutional Coordinator
	The Institutional Coordinator assigns the request to the appropriate Infrastructure Lead or team based on expertise and availability.	Institutional Coordinator
	The Infrastructure Lead works on resolving the request, providing regular updates to the Researcher and Institutional Coordinator.	Infrastructure Lead
Validation and Closure	Once the request is resolved, the Institutional Coordinator closes the request in the system and notifies the Researcher.	Institutional Coordinator
	Feedback from the Researcher is analysed by the Institutional Coordinator to identify areas for improvement in the service request process.	Researcher and Institutional Coordinator

Table I. Process Stages and description

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

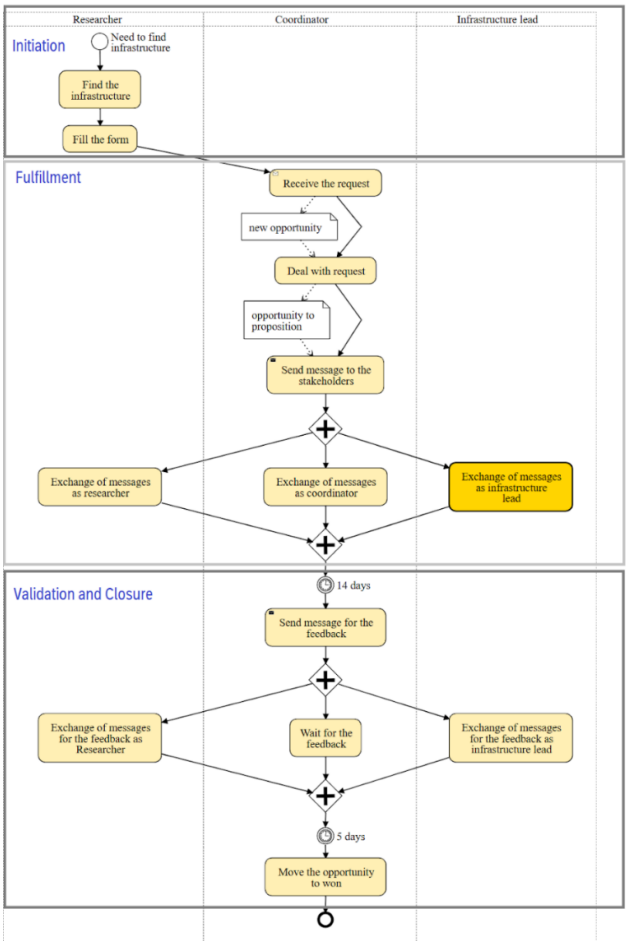


Fig. 1. sketch miner diagram of business process of research infrastructure

Figure 1 Business process workflow for handling research infrastructure requests in the Unite! Agora platform. The diagram shows the three main stages Initiation, Fulfilment, and Validation & Closure mapped against ITIL v4 Service Request Management practices and role responsibilities (Researcher, Institutional Coordinator, and Infrastructure Lead). The diagram of three-stage process for handling research infrastructure requests:

1. Initiation: The researcher identifies a need, locates the infrastructure, and submits a request via a form.
2. Fulfillment: The coordinator reviews the request, evaluates opportunities, and communicates with stakeholders (researcher, coordinator, and infrastructure lead) to address it.
3. Validation and Closure: After 14 days, feedback is requested, and all roles exchange messages to finalise the process.

A. Process Definition

The process which is used here, is ITIL v4 Service request management which is a comprehensive practice that focuses on efficiently managing and fulfilling service requests from users or customers. It involves several key components that work together to provide a smooth and user-centric experience. It encompasses the end-to-end process of managing service requests, from initiation to fulfilment, with the goal of meeting user needs and delivering value.

In this initial evaluation, the KPIs presented are illustrative targets aligned with ITIL v4 expectations rather than actual system data. Future work will incorporate real service metrics and process mining techniques to validate these targets empirically and enhance the credibility of the evaluation.. For instance, while the target resolution time for high-priority requests is set at 8 hours, infrastructure availability and researcher response delays suggest that this target may need adjustment or further support mechanisms.

Table 2. Process Metrics

Key Performance Indicators (KPIs)	**Description**	**Target**	**Role**
Request Acknowledgement Time	The time taken to acknowledge a service request from the moment it is submitted.	95% of requests acknowledged within 15 minutes.	Institutional Coordinator
Request Categorization and Prioritization Time	The time taken to categorise and prioritise a service request after it has been logged.	90% of requests categorised and prioritised within 1 hour.	Institutional Coordinator
Request Assignment Time	The time taken to assign a service request to the appropriate team or individual after categorization	85% of requests are assigned within 2 hours.	Institutional Coordinator
Initial Response Time	The time taken to provide an initial response to the requester after the request has been assigned.	80% of requests receive an initial response within 4 hours.	Infrastructure Lead
Request Resolution Time	The time taken to resolve a service request from the moment it is logged.	High Priority: 90% resolved within 8 hours. Medium Priority: 85% resolved within 24 hours. Low Priority: 80% resolved within 3 business days.	Infrastructure Lead
Request Closure Time	The time taken to close a service request after it has been resolved.	95% of requests closed within 2 hours of resolution.	Institutional Coordinator
First-Time Resolution Rate	The percentage of service requests resolved on the first attempt.	75% of requests are resolved on the first attempt.	Infrastructure Lead

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

	tempt without requiring further follow-up.		
Customer Satisfaction Score	The average satisfaction rating provided by requesters after their service request has been resolved	Achieve an average satisfaction score of 4.5 out of 5.	Institutional Coordinator
Request Backlog	The number of unresolved service requests that are pending beyond the SLA targets.	Less than 10% of total requests in backlog	Institutional Coordinator
Process Compliance Rate	The percentage of service requests processed in compliance with the defined process steps and SLAs.	95% compliance with process steps and SLAs.	Institutional Coordinator
Request Trends	Analysis of trends in the volume and types of service requests over time.	Identify and address any recurring trends monthly.	Institutional Coordinator

B. Roles and Responsibilities

The current role distribution suggests a centralized dependency on the Institutional Coordinator for multiple process stages, from request logging to feedback analysis. This centralization may pose a bottleneck risk. To mitigate this, a role reallocation strategy or workflow automation could be recommended, especially for high-volume request periods. Additionally, a layered accountability model, where validation and follow-up are shared between roles, would improve scalability and compliance. And here we will add some responsibilities based on the identified roles, which are:

- **Researcher**: Searches for the required infrastructure.
- **Institutional Coordinator**: Manages services.
- **Contact of Infrastructure Lead**: Resource provider (HEIs or HEI alliances).

Table 3. Roles and Responsibilities

Activity/Task	**Researcher**	**Institutional Coordinator**	**Infrastructure Lead**
Initiate Service Request	A/R		
Log Service Request	A/R	I	
Categorise and Prioritise Service Request		A/R	
Assign Service Request		A/R	
Fulfil Service Request			A/R
Provide Updates to Requester		A	R
Resolve Service Request			R
Close Service Request	I	R	
Seek Customer Feedback	I	A/R	
Evaluate and Improve Process		A/R	R

IV. Conclusion and Recommendations

The evaluation of the Research Infrastructure service using ITIL v4's Service Request Management framework provides a structured understanding of the process and identifies areas where performance could be enhanced. The current study is primarily conceptual, presenting a baseline model for evaluating CRM-related processes using ITIL v4 practices. While actual performance measurements before and after ITIL implementation are not yet included, the framework outlined here provides a structured foundation for empirical validation. Future work will incorporate real service data and comparative analysis to demonstrate the impact of ITIL adoption on performance outcomes. Future iterations of this work should include event log analysis and real-time performance data to verify KPI achievement and support evidence-based process optimization. Based on the current evaluation, recommendations have been proposed to enhance process clarity, redistribute roles, and align performance expectations with operational capacity. Here are the Recommendations:

- Redistribute Role Load: Delegate some responsibilities from the Institutional Coordinator to avoid bottlenecks in high-volume periods.
- Implement Categorization Automation: Introduce rule-based triage tools to speed up request classification.
- Monitor Real Performance: Plan the implementation of a dashboard for KPI tracking and compliance audits.
- Feedback Integration Loop: Use real-time satisfaction metrics instead of post-closure only.

Acknowledgement

We gratefully acknowledge that this project has received funding from the European Union's Horizon Europe programme under grant agreement No :101095314 (aUPaEU). We sincerely appreciate the opportunities and resources available to enable us to conduct our research and achieve our objectives.

Refrences

[1] S. Mithas, M. S. Krishnan, and C. Fornell, 'Why Do Customer Relationship Management Applications Affect Customer Satisfaction?', *J. Mark.*, vol. 69, no. 4, pp. 201–209, 2005.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[2] V. Guerola-Navarro, H. Gil-Gomez, R. Oltra-Badenes, and P. Soto-Acosta, 'Customer relationship management and its impact on entrepreneurial marketing: a literature review', *Int. Entrep. Manag. J.*, vol. 20, no. 2, pp. 507–547, Jun. 2024, doi: 10.1007/s11365-022-00800-x.

[3] J. L. Rubio, "A Quantitative Analysis of Information Systems Management in the Educational Industry," Information, vol. 3, no. 2, pp. 237–251, 2022, doi: 10.3390/info3020013.

[4] R. U. Khan, Y. Salamzadeh, Q. Iqbal, and S. Yang, 'The Impact of Customer Relationship Management and Company Reputation on Customer Loyalty: The Mediating Role of Customer Satisfaction', Jan. 2022.

[5] S. Das and M. Mishra, 'The Impact of Customer Relationship Management (CRM) Practices on Customer Satisfaction', *Springer Books*, pp. 43–54, 2019.

[6] L. Kuo-Wei, Lanting, Maria Corazon L., and Rojdamrongratana, Maneesap, 'Managing customer life cycle through knowledge management capability : a contextual role of information technology - EconBiz'. Accessed: Dec. 03, 2024. [Online]. Available: https://www.econbiz.de/Record/managing-customer-life-cycle-through-knowledge-management-capability-a-contextual-role-of-information-technology-lee-kuo-wei/10011816994

[7] P. Frow, 'A Strategic Framework for Customer Relationship Management', *J. Mark.*, Jan. 2005, Accessed: Dec. 03, 2024. [Online]. Available: https://www.academia.edu/5398630/A_Strategic_Framework_for_Customer_Relationship_Management_Author_s_Adrian

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Evaluación y Mejora de los Sistemas de Detección de Intrusiones en IoMT: La Importancia de los Entornos Específicos

Jordi Doménech[1,2], Olga León[1], Muhammad Shuaib Siddiqui[2], Josep Pegueroles[1]
[1]Departamento de Ingeniería Telemática, Universitat Politècnica de Catalunya, Barcelona, 08034
[2] Cybersecurity & Blockchain Research Group, i2CAT, Barcelona, 08034
{jordi.domenech.fons, olga.leon, josep.pegueroles}@upc.edu[1], {jordi.domenech, shuaib.siddiqui}@i2cat.net[2]

El avance del Internet de las Cosas Médicas está mejorando la atención sanitaria, pero también aumenta los riesgos de ciberseguridad. Este estudio compara el rendimiento de modelos de detección de intrusiones entrenados con datos de entornos genéricos IoT, frente a datos específicos de entornos IoMT. Además, se proponen mejoras en el preprocesamiento de datos que logran aumentar significativamente la precisión de los modelos. Estos hallazgos subrayan la necesidad de utilizar datasets específicos para proteger estos entornos.

***Palabras Clave*—IoMT, IDS, ML, Cambio de Dominio en Datos, Optimización de Conjuntos de Datos.**

I. RESUMEN DE TRABAJO YA PUBLICADO

El Internet de las Cosas Médicas (IoMT) está transformando la atención sanitaria mediante el uso de dispositivos conectados para monitorizar pacientes, facilitar diagnósticos y mejorar la eficiencia. No obstante, esta innovación conlleva nuevos desafíos de ciberseguridad. En 2023, los ataques a entornos sanitarios aumentaron un 55% según ENISA, siendo los dispositivos IoMT objetivos prioritarios. Los Sistemas de Detección de Intrusiones (IDS) basados en Inteligencia Artificial (IA) y Aprendizaje Automático (ML) se han convertido en una solución prometedora, aunque la mayoría de estudios actuales utilizan conjuntos de datos genéricos del IoT, lo cual compromete la aplicabilidad en contextos médicos específicos.

En [1] se presenta el trabajo de investigación realizado que aborda esta brecha mediante una comparación sistemática entre dos conjuntos de datos: CICIoT2023 (IoT general) y CICIoMT2024 (IoMT). Se utilizan cuatro modelos de ML ampliamente reconocidos: Decision Tree (DT), Random Forest (RF), XGBoost (XGB) y Feedforward Neural Network (FNN). Se plantean tres contextos de evaluación: (1) Contexto de línea base: entrenamiento y prueba en el mismo conjunto de datos, (2) Contexto de generalización: entrenamiento en IoT general y prueba en IoMT, y (3) Contexto de extensión: se entrena con IoT general más un 0,1% de IoMT. Los resultados muestran caídas del 66,87% en la F1-score cuando se entrena con un conjunto y se prueba con el otro.

Además, en [1] se proponen mejoras en el preprocesamiento del conjunto IoMT para mejorar la detección de ataques en contextos sanitarios reales, solucionando los siguientes retos: (1) Tamaños de ventana inconsistentes: algunos ataques se agrupan en ventanas de 100 paquetes, mientras que otros usan 10. Se propone una ventana uniforme para todos los ataques, (2) Ausencia de conjunto de validación: solo se emplean conjuntos de entrenamiento y prueba. Se propone una división adecuada en entrenamiento/validación/prueba, (3) Dependencias temporales: se debe garantizar un preprocesamiento adecuado que elimine la correlación temporal mediante barajado de datos, y (4) Desbalanceo de clases: algunos ataques dominan el dataset. Se sugiere usar técnicas de sobremostreo o submuestreo. El mejor resultado fue obtenido por RF con SMOTE, logrando una precisión del 99.85% y F1-score del 97%, mejorando un 26% respecto a trabajos previos.

Este estudio valida la necesidad de datasets específicos IoMT para desarrollar IDS robustos: el uso exclusivo de datasets IoT conlleva importantes pérdidas de rendimiento. También demuestra cómo técnicas de optimización de preprocesamiento mejoran notablemente los resultados.

AGRADECIMIENTOS

Esta publicación y sus resultados de investigación cuentan con el apoyo del programa predoctoral AGAUR-FI Ajuts Joan Oró (2024 FI-1 00643) Ajuts Joan Oró (Generalitat de Catalunya) y de la Cátedra de Ciberseguridad CARISMATICA financiada por la UE e INCIBE.

REFERENCIAS

[1] J. Doménech, O. León, M. S. Siddiqui, J. Pegueroles, "Evaluating and Enhancing Intrusion Detection Systems in IoMT: The Importance of Domain-Specific Datasets", *Internet of Things*, Volume 32, 2025, 101631, ISSN 2542-6605, https://doi.org/10.1016/j.iot.2025.101631.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

MuHoW: Distributed protocol for resource sharing in collaborative edge-computing

Isaias Martinez-Yelmo, Joaquin Alvarez-Horcajo, Elisa Rojas, Juan A. Carral, Victoria Noci-Luna
Universidad de Alcalá, Departamento de Automática.
Pza. San Diego, s/n, 28801 Alcalá de Henares (Madrid), España
{isaias.martinezy, j.alvarez, elisa.rojas, juanantonio.carral, victoria.noci}@uah.es

Multi-Hop Wireless Resource Sharing Protocol (MuHoW) is a lightweight protocol tailored for multi-hop wireless networks. MuHoW enables resource sharing within collaborative edge-computing networks by facilitating the discovery of wireless neighbours and establishing a confluence tree directed toward the edge/fog infrastructure. This tree serves as a conduit for aggregating essential resource-sharing information, ensuring the establishment of a seamless collaborative edge/fog environment. The empirical findings highlight the scalability of MuHoW, due to the linear growth of control messages with network size and its high efficiency even in lossy environments.

***Key words*—Edge Computing,Resource Sharing,Routing**

I. Introduction

The main contributions of MuHoW [1] are as follows: (1) Definition of a lightweight and distributed protocol for resource sharing designed for low-capacity and low-energy devices. (2) Full support of multi-hop edge computing scenarios in which the sensor/nodes may be more than one-hop away from a root node (gateway). (3) A performance evaluation considering wireless scenarios with/without background traffic and losses.

II. MuHoW protocol

MuHoW protocol creates a local map that allows each node to discover, share, and request resources in a distributed manner. This local map only contains information about neighbours; hence, it is independent of the network size. It stores routing labels associated to each neighbour; so, multi-hop paths are directly inferred from them.

The MuHoW nodes rely on two periodic procedures (based on hierarchical labelling) to discover and identify its neighbours and to build a confluence tree rooted at the node designated as the root node (gateway) that will set the direction of the resource demand/offer message exchange that allows to build the target local map.

A. Neighbour Discovery

This procedure is based on the periodic broadcast of `Hello` messages. Neighbours are detected with the source Media Access Control (MAC) addresses from the `Hello` messages received, and each node constructs its table of neighbours identified by those MAC addresses. Section 3.2.1 from [1] contains further information.

B. Confluence Tree Construction

A confluence tree rooted at the *Root* node is built from it by broadcasting `SetTree` messages from each node to its neighbours. In this way, each node acquires a hierarchical routing label and autonomously derives its branch/leaf role in the tree and a parent/child association to its neighbours. See Section 3.2.2 from [1] for further information.

C. Resource Information Sharing

The aggregation of shareable resources takes place towards the *Root* node, in the opposite direction to how the confluence tree is built. Asynchronously, *Leaf* nodes begin the propagation of resource information (demands/offers) by sending `Load` messages to their *Parent* nodes, which consolidate the demands/offers from their children. Afterwards, they send the result with their own `Load` messages to their parents, until reaching the *Root* node. Finally, the *Root* node obtains the balance of resources in the tree. If necessary, additional resources are requested to the edge infrastructure. Section 3.2.3 from [1] contains further information.

III. Conclusions

The results from [1] demonstrate that MuHoW is scalable, since the number of required messages and the convergence time grow linearly with the number of nodes, and its functionality is maintained even in lossy environments.

Acknowledgements

This work has been partially funded by the project TUCAN6-CM (TEC-2024/COM-460) from Comunidad de Madrid and the grants from the Universidad de Alcalá "Ayudas para la Recualificación del Sistema Universitario" and promotion and advancement of research and technology transfer (CM/DEMG/2024-038).

Referencias

[1] J. Alvarez-Horcajo, I. Martinez-Yelmo, E. Rojas, J. A. Carral, and V. Noci-Luna, "Muhow: Distributed protocol for resource sharing in collaborative edge-computing networks," *Computer Networks*, vol. 242, p. 110243, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128624000756

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Beyond 5G Satellite Transmission and Indoor UAVs in Dams Monitoring: An Opportunistic Data Collection Strategy for Isolated Rural Areas

Manuel Jesús-Azabal, José Gómez-delaHiz, José García-Alonso, Jaime Galán-Jiménez
Departamento de Sistemas Informáticos y Telemáticos,
Universidad de Extremadura
Avenida de la Universidad, S/N, 10003, Cáceres
{manuel, jagomezdh, jgaralo, jaime}@unex.es

***Palabras Clave*—B5G, UAVs, LEO Satellite, opportunistic networks, Internet of things**

I. SUMMARY

In regions where traditional communication infrastructure is lacking, the maintenance of critical hydraulic structures such as dams becomes particularly challenging[1]. Manual auscultation—routine structural inspections—inside dam galleries is time consuming and drive to delays in hazard detection, especially in isolated rural areas. To address this issue, a fully autonomous monitoring architecture is proposed, integrating indoor Unmanned Aerial Vehicles (UAVs), Bluetooth Low Energy (BLE), Beyond 5G (B5G) satellite connectivity, and edge-based Machine Learning (ML) analysis. This system enables opportunistic data collection and real-time emergency alerting, even in environments without Internet access.

The architecture operates through three main phases: sensor data capture, satellite-based data forwarding, and cloud-based storage and analysis. During the capture phase, UAVs autonomously navigate the dam's interior using a pre-programmed flight path, collecting measurements opportunistically from BLE sensors that monitor variables such as water flow, temperature, and structural displacement. The UAVs also record video footage for visual assessment. Onboard processing allows data to be locally stored using lightweight database solutions and containerized services.

Following data acquisition, the UAVs transmit collected information via B5G to a Low Earth Orbit (LEO) satellite. The satellite processes the data using pre-trained ML models. Unsupervised algorithms such as K-means are employed to detect statistical anomalies in the sensor data, while convolutional neural networks (YOLO) analyze video streams to identify early signs of structural degradation, including cracks and humidity marks. Based on the analysis, the system distinguishes between minor irregularities, which prompt routine inspections, and critical anomalies, which trigger immediate emergency alerts. The final phase sends the analyzed data to a central server that stores past inspection records and updates ML models. A client-facing application interfaces with the database to visualize trends and support technical decision-making. This feedback loop enhances the precision of the models over time and improves overall infrastructure monitoring effectiveness.

To evaluate technical viability and Quality of Service (QoS), a testbed replicating a opportunistic communication was implemented. BLE sensors using ESP32 microcontrollers generated simulated data, while a manually piloted UAV equipped with a Raspberry Pi and 5G antenna performed the inspection. Tests conducted at different UAV speeds (0.5, 1.0, and 1.5 m/s) revealed that lower speeds ensure higher data delivery success, achieving 100% delivery at 0.5 m/s. Higher speeds reduced latency but also resulted in a modest loss of transmitted data. These results confirm that UAV speed significantly influences the balance between data completeness and communication delay.

[1]This work has been published in IT Professional, IEEE 2024. Impact factor: 2.2, Q2 JCR. DOI: https://doi.org/10.1109/MITP.2024.3433399. This work was supported in part by MCIN/AEI/10.13039/501100011033 and by the European Union "Next GenerationEU/PRTR," by the Ministry of Science, Innovation, and Universities (projects TED2021-130913B-I00, PDC2022-133465-I00), European Union's Digital Europe Program (DIGITAL) under Grant 101083667-TECH4EFFICIENCYEDIH. Co-funded by the Ministry of Industry, Trade, and Tourism of the Government of Spain, within the framework of the Recovery and Resilience Mechanism. However, the views and opinions expressed are solely those of the author(s) and do not necessarily reflect those of the European Union or the Government of Spain. Also, Science and Digital Agenda of the Regional Government of Extremadura (GR21133) and the European Regional Development Fund.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Proposing AI-based Network Forecasting in Proactive Sync for Shared Media

Juan González Salinas, Fernando Boronat Segui.
Departamento de Comunicaciones,
Universitat Politècncia de València – Campus de Gandia
IIM R&D Group – Calle Paraninfo, 1, 46730, Grao de Gandia (Valencia-España)
juagons4@upv.edu.es, fboronat@dcom.upv.es

Abstract. Advancements in multimedia streaming, extended reality, and telecommunications allow geographically distant users to share media experiences as if they were together. These technologies enable rapid data exchange and seamless shared media consumption, but also present challenges for immersive real-time interaction. A key challenge is ensuring synchronized media playout across all devices in shared sessions to maintain a high-quality user experience. Current synchronization methods are mostly reactive, correcting asynchronies after they occur, which can disrupt user experience. To overcome this, the authors propose preventive synchronization using network forecasting. This paper also explores how AI-based techniques could be applied to improve media synchronization techniques, by optimizing the network usage and preventing asynchronous playout before they happen.

***Palabras Clave*- Artificial Intelligence (AI). Multimedia. Media Synchronization. Computer-Mediated communications (CMC). Quality of Service (QoS). Quality of Experience (QoE).**

I. Introduction

Nowadays, people consume increasingly immersive content (video, audio, text, omnidirectional audio/video, extended reality (XR), scents, haptic, metadata…), using heterogeneous broadcast and broadband delivery technologies (e.g., DVB-based, IP-based streaming, IPTV, 4G/5G mobile networks…) and consumer devices (TV, PC, tablet, smartphone, head-mounted display…). These technologies can be combined (as in, e.g., Hybrid Broadcast Broadband TV), but this diversity adds complexity to the multimedia systems, increasing the potential points of failure that must be addressed.

Remotely shared media experiences, such as Social TV and the metaverse, are gaining traction, requiring synchronized media playout across local and distributed devices to ensure high-quality user experience (QoE). Modern multimedia applications support real-time interactions and introduce advanced media formats, including ultra-high-definition video, 360° video, multi-sensory effects (mulsemedia), XR, and holograms [1]. To maintain the QoE in more complex distributed multimedia systems, more advanced media synchronization (abbreviated as sync, hereafter) techniques are needed [1], [3]. Traditional techniques barely considered emerging technologies and media, and although networks are becoming faster, they are not always 100% reliable, therefore, modern techniques are still needed [2]-[5].

Several works have identified processes that contribute to end-to-end (E2E) latency in multimedia communication [13], from which network delay is often the most significant. Media playout async creates inconsistencies, where users experience different temporal or spatial relationships across devices [13]. For example, distributed viewers may see different frames at the same time, or elements in a 3D environment in incorrect positions, worsening the QoE [13].

The sync between distributed related media playout processes (called Inter-destination Media Synchronization, IDMS) was first proposed in 1994 [14]. It ensures that distributed users perceive media streams (video, audio, text, metadata) and interactions (messages, voice, commands) in the correct order despite network impairments. However, shortcomings in the management and results of the sync have been identified [13], [15].

Although IDMS solutions exist that make corrections less noticeable [13], they correct the async after detection. Moreover, async above a tolerable threshold (e.g., 150-200ms for VoIP calls), require reactive corrections. As an alternative, preventive (or proactive) techniques work in advance.

The Artificial Intelligence (AI) field is continuously expanding, helping in inference, data analysis and synthesis, forecasting, and decision making [2], [6]. Regarding media sync, AI can optimize network usage, predict network congestion, classify content, compress data, and adjust media delivery [7]-[13].

In this paper, the authors propose the use of AI to forecast the media playout asynchrony (abbreviated as async, hereafter) states before they happen and to avoid them (i.e., proactive sync techniques). More specifically, the main contributions in this paper are:

- Highlighting the current limits of traditional media sync.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- Showing how AI-based techniques could benefit shared media consumption experiences.
- Proposing the use of AI-based network forecasting for developing new preventive media sync solutions.

The contents of this paper are distributed as follows. First, in Section II, the current state of media sync and AI is summarized. Then, in Section III, AI-based forecasting for media sync is explored, comparing and discussing how it would benefit the media sync. Afterwards, the gaps in the media sync field that may benefit from the use of AI-based forecasting are identified in Section IV, providing a framework on how to implement the forecasting techniques to fill those gaps. Finally, some conclusions and insights into the future development of media sync techniques, including AI-based forecasting techniques, are summarized in Section V.

II. Related Work

Figure 1 represents most of the elements involved in an advanced shared media session, although simpler systems would not include everything. End users, from different geographical regions, through distinct connections, and using varying number and types of devices, experience different Points of View (POV) of the same media and interact in different ways, which makes the system complex.

The authors of this paper have reviewed the existing traditional sync techniques in [15], summarized in Table I.

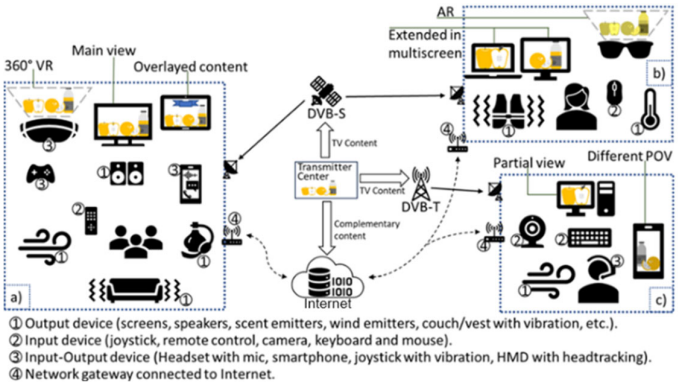


Fig. 1. A complex interactive multimedia system with most of the elements it can contain.

Table I
Comparison of Traditional Sync Techniques

Sync Technique	Limitations
Bucket Sync (BS)	Requires low number of users
Local Lag (LL)	Requires low network latency
Time Warp Sync (TWS)	Not suited for low-end devices
Dynamic Local Lag (DLL)	Requires low network latency
Breathing Time Buckets (BTB)	Requires low number of users
Adaptive Δ-causality	Requires low network latency
Lockstep Sync (LS)	Not for real-time applications
Asynchronous Sync (AS)	Hard to keep global consistency
Adaptive Event Sync (AES)	Hard to keep global consistency
Breathing Time Warp (BTW)	Not suited for low-end devices
Trailing State Sync (TSS)	Hard to implement, requires more computing, still has inconsistencies
Event Correlation Sync (ECS)	Hard to implement, requires more computing, still has inconsistencies
Optimistic Obsolescence-based Sync (OOS)	Hard to implement, requires more computing, still has inconsistencies

All the existing techniques but LS are reactive ones as they correct the async states once they occur. Some of the reactive techniques, like TSS, ECS, and OOS, try to minimize the times when async states happen, allowing a certain degree of inconsistencies that does not impair the QoE. On the contrary, LS is a proactive sync technique, as it avoids asynchronous states, but it does so by locking further interaction until all the participants have synchronized the last content, thus a low-latency client in a session would slow the rest of the participants.

Another issue with traditional sync techniques is the scalability. As the amount of media data transferred in the network continues increasing, with new types of media content and better quality (e.g., bigger image resolution) being available, media systems become more complex to manage. AI can automate processes and reduce computing and data size requirements [7], [16], [17] allowing a higher number of simultaneous connections, streams and devices.

The works of [9], [10] apply AI for time sync, i.e., for synchronizing the clocks of multiple independent devices. For the goal of media sync, ensuring proper time sync is also crucial, but media sync additionally requires dealing with the media playout. To date, AI has been applied to network and media-related challenges, although not directly for media sync. Ideally, AI in those fields could improve the Quality of Service (QoS) or directly manage the sync. As the sync techniques employ network communication, use computing resources, and access the media (e.g., when generating timing metadata based on video frames), any way of improving those requirements can increase the performance of the sync and scalability of the media systems. Regarding network and media applications, AI has already been used for performance optimization, traffic classification, modelling, and state forecasting.

Regarding network performance optimization, AI- and non-AI-based solutions have been developed to efficiently manage media stream delivery, organizing the media streaming and playout and reducing bandwidth usage [18], [19]. Advanced AI techniques, like Machine Learning (ML), can be used to cache data and analyse patterns in media consumption so data can be compressed or filtered without harshly affecting the final quality [20]-[23]. For instance, ML can model user behaviour when watching a 360° video, so that only the parts that the user is likely paying attention to are transmitted at higher quality.

Traffic analysis is employed to classify the flows, nodes, connections, and contents of the network. The current goals of AI for traffic classification are detecting network attacks and other malicious traffic, providing network security, enabling firewall filtering, QoS management and security monitoring [24]. Nonetheless, AI could also help analyze complex communication patterns and dynamically reorganize the routing to ensure the network requirements of shared media applications. This would help building more robust media systems, as media data and sync control packets would be better filtered and prioritized [24].

Analogously, modelling the parameters and changes in the state of the network can be used for detecting bottlenecks [25]. In shared media, the sync techniques could be (re)adapted or interchanged to take advantage of the available network resources and save computing time when estimating parameters like network bandwidth and latency [25].

Lastly, network QoS parameters can be optimized before bottlenecks happen by using previously known network states to estimate future ones [11], [12]. For example, network patterns may repeat periodically and geographically, as users connect from different places at different times. So, using AI to forecast when and how network parameters change allows for fine-tuning the QoS [11], [23]. Currently, AI-based forecasting is not specifically applied to media sync.

Table II summarizes the ways in which the authors consider that the application of AI-based techniques can contribute to improving the QoS in a way that could help to improve the media sync. Although forecasting has been used in networks in the past, it is yet to be applied in the prediction of future media playout asynchronies.

III. AI-BASED NETWORK CONDITIONS FORECASTING

New sync techniques should be predictive and prevent asynchronies that can be perceived by the users. Predicting when asynchronies are going to happen allows taking measures in time, removing the need for reactive sync techniques. The authors could not find any research on this topic, but dead reckoning techniques, which were used in the past in the context of media sync [15]. Nevertheless, although dead reckoning is predictive and optimizes network usage, it cannot anticipate asynchronies [15].

Table II
APPLICATIONS OF AI IN NETWORKS THAT BENEFIT THE MEDIA SYNCHRONIZATION

Networking fields	Applications of AI-based techniques	Benefits provided to the media sync
Performance optimization	Filtering and compression	More available bandwidth (e.g., for sync control messages)
	Caching and distribution	Faster delivery of media streams
Traffic classification	Packet classification	Reduce errors in media sync
	Flow detection	Faster delivery of media streams
	Packet Filtering	More available bandwidth (e.g., for sync control messages)
	Detection of network changes	Dynamically (re)configure the sync solution
Modelling	Saving computing time and resources	More computing time and resources available for the sync solution
	Detection of network changes	Dynamically (re)configure the sync solution
State forecasting	Detection of future events	Proactively (re)configure the sync solution

In shared media sessions, media contents are generally captured or generated, stored, modified, transmitted, and played. AI-based techniques could be employed in each of these steps to forecast problems before they happen. Media sync ensures the correct playout notwithstanding delays caused during those processes. The delay is not usually constant, and other problems, such as like data loss and corruption, can increase it (i.e., due to retransmission of the lost data or waiting for the next content to resume playout).

Table III summarizes the existing challenges regarding synchronized shared media playout.

Network conditions, such as bandwidth, packet loss and latency, have been studied in the context of shared media for media sync [14]. The media source (i.e., where media data are captured or generated) should consider the network conditions at the destinations (media playout). For example, if an on-demand streaming service is distributing videos to tablets with a mobile connection, it should support low bandwidth connections and low computing capabilities by, for instance, offering lower video qualities and faster decompression algorithms.

To correct the asynchronies caused by E2E problems, Adaptive Media Playout (AMP [13]) can be used. It smoothly modifies the shared media playout rates, so users do not notice the corrections (unlike with aggressive adjustments, like playout skips and pauses) [3], [13]. However, AMP cannot work when the playout delays become greater than a tolerable threshold. Thus, to solve it, AI-forecasting is needed.

To start applying AI for media sync, the authors first identify the components of the sync algorithms [13]: sync control data; media and sync control data buffers; transmission schedule; media capture; and playout.

- Sync control data: Algorithms use timing data to align streams. Media streams may be unaligned due to errors in the process of generating, storing or transmitting data.
- Media and sync control data buffers: Data is written and read at the sources (origin of the media), destinations (media players) and during the communication (e.g., through Internet nodes).
- Transmission schedule: Media and sync control data are transmitted at different rates. As mentioned before, network delay can have an important impact.
- Capture and playout: Aggressive techniques like pausing the capture or the playout should be prevented. Smooth techniques like dynamically adjusting the capture and playback rates improve buffer usage and QoE [13]. Processing the captured media to generate timing information could also be optimized.

Table III
CHALLENGES IN THE END-TO-END MEDIA DISTRIBUTION CHAIN

Processes	Challenges
Capture & Generation [3], [6], [13], [22]	Multiple sources & destinations
	Network conditions
	Different media types
Storage [19]	Time-alignment
	Scalability
	Data integrity
Modifications [19], [20]	Timing preservation
	Quality loss
	Resource and computing consumption
Transmission [2], [12], [17], [20]	Coordination of multiple streams
	Network conditions
Playout [3], [6], [17]	Network conditions
	Resource and computing consumption

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Ultimately, all the components must work together so the media sync solution works correctly. Monitoring the full state of the E2E communication would allow to manage the sync more efficiently and accurately. However, in most common cases, there is no control over all those components (for example, on the Internet).

As already discussed, AI-based forecasting consists in using AI models (e.g., ML) to predict future events. In this case, the focus of this paper is to use AI-based techniques to forecast asynchronies in the media playout so they can be prevented.

The process of generating AI models in the context of network forecasting has been researched in the past. This requires data collection, feature engineering, model training, evaluation, and continuous learning. The authors have considered those steps in the proposed framework, described later.

A. Data collection

The time-series data of the different network characteristics gathered over time are necessary for creating AI-based models to forecast the network behaviour [12]. Currently, studies have focused on either running network simulations or capturing data from very specific networks (like universities or research centres). Thus, more realistic datasets from the Internet and from actual users are required for creating more accurate network forecasting models.

Another characteristic is that current solutions for data collection have low-level access to the network architecture, using packet sniffing, capturing most of the possible data. However, shared media service providers may have less information available, as they will be using the Internet or Software Defined Networks (SDN) to offer their services. In conclusion, data gathered at the endpoints (Clients or Servers) of the E2E communications will be the most relevant for creating AI-based forecasting models.

B. Feature Engineering

Defining the most relevant characteristics of the network data will improve the performance and accuracy of the models. The most relevant network characteristics for media sync that have been identified in the past are latency, throughput and jitter. Nonetheless, media-related characteristics (like video bitrate, buffering, compression, etc.) will need to be studied in more detail to determine their impact on media sync.

AI-based techniques can calculate network characteristics from partial data (e.g., predicting the overall network latency from a few previous latency samples) and even estimate unknown data [11]. This offers a great advantage when forecasting network states with imprecise information (e.g., measuring the network with web browser code instead of low-level packet sniffing).

C. Model training

Previous works have forecasted the network with Moving Average (MA) models, Autoregressive (AR) models, Support Vector Machines (SVM), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) and Ensemble Learning (EL) [12]. Because network data behaves non-linearly, approaches like MA may not be suited. And, because forecasting may be required to be done in real-time, models that require high computation (such as RNN and LSTM) are also less suited. EL averages the results from multiple less-performant, but faster, models to improve the overall accuracy of the system to forecast the network [12]. Fuzzy Logic and Decision Trees (DT) are other, less explored, options in network forecasting. They offer a good balance between accuracy and inference time, so they should be considered when applying EL.

D. Model Evaluation

Studies on network forecasting have also defined metrics to employ for determining the quality of the AI-based models. These metrics include Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) [12]. Additionally, Normalized Mean Absolute Error (NMAE) has been studied to measure the spatial and temporal errors of network traffic prediction [11].

Because the main goal is to forecast adverse situations of the network that impair the shared media, AI-based techniques should focus on anomaly detection and decision making [12]. For example, predicting network load increases can be used to start mechanisms that help to reduce the network load [12] before it exceeds a specific threshold. Hence, the challenge resides in determining not only if the forecast was accurate but also if it allowed for taking effective measures and on time.

E. Continuous learning

The network behaviour can change over time, so the AI-based models should be able to adapt and retrain when there is a drift in the network conditions. The authors propose continuous learning to also adapt to other changes in media applications, like the number of simultaneous users or the types and quality of media.

IV. Framework for applying AI-based network state/conditions Forecasting in Media Synchronization techniques

Media sync is a complex problem, with decisions that should be made in real-time. So, creating AI-based techniques for media sync requires a strategy that deals with the E2E communication (capture, transmission…), the model creation (data collection, training…) and the components of media sync techniques (media data, sync control data…).

Considering what has been discussed in section III, the authors propose the following framework for applying AI-based techniques, including the possible decisions to be made based on what the AI-models forecast. This framework considers the existing research explained earlier in this paper and some experiments made by the authors. The framework is focused on the creation of media sync techniques that forecast playout asynchronies. Forecasting will allow the E2E media systems to proactively manage the media sync in shared media, adapting the system to changes in network conditions.

It is also important to note that, when adding AI-based techniques to the media shared systems, the added latency (due to the AI algorithms execution) should be minded. Although some systems may have a trade-off between the added delay and any other QoS quality (e.g., increased video resolution in exchange for higher network bandwidth usage),

employing AI can also spend less time than traditional mechanisms [22].

Analysing the users' behaviour and the current network state can be used to create AI-based techniques that predict future users' actions and network states [5], [17], allowing to make the media sync-related actions (e.g., playout adjustments or event triggering) in advance, before the media playout processes get out-of-sync (preventive).

Figure 2 illustrates the architecture of the proposed framework including a media sync solution based on a centralized approach and the creation and application of AI-based trained models, showing media sources and destinations, and how data collection, modelling, etc. happen and interact. The following subsections present each part of the proposed framework, with algorithms and figures illustrating them.

A. Media source

Different kinds of media data sources exist. As presented in Figure 1, media providers can distribute the media through different means (e.g., Internet, satellite, etc.), and possibly simultaneously. Clients (destinations of the main sources) can also interact by generating and transmitting more media. For example, in Social-TV, users can share camera and audio with other participants, then being sources and destinations of media as well.

The proposed architecture in Figure 2, shows a single instance of source and destination, as well as a centralized sync node. However, the roles of source, sync node, metric storage, playout monitoring, etc., can be shared between the same or multiple devices. For example, this architecture could be adapted to include proxies for storing per-client metrics, the source node also running the forecasts, or even multiple independent nodes with access to the same database sending the forecasts to the sync node or the proxies for applying the sync corrections.

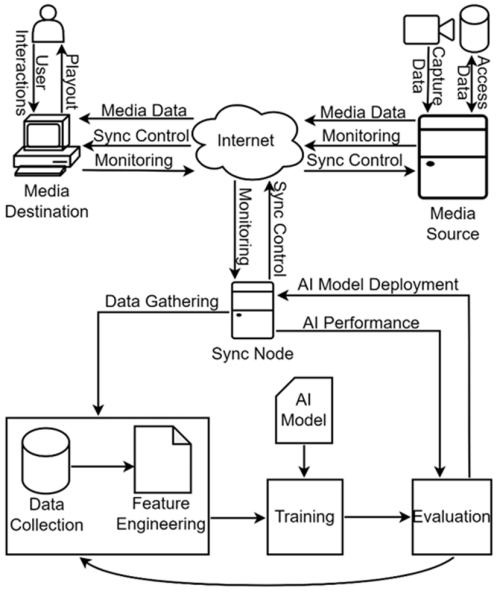


Fig. 2. Framework of AI integration in shared media systems.

At sources, the processes (summarized in Table III) of data capturing, storage, modification and transmission can happen. Algorithms 1 to 3 show the proposed ways to apply AI for the goal of forecasting network impairments and avoiding their impact on the correct media playout.

Obtaining associated sync information (e.g., timestamps, sequences, identifiers) is done in most sync techniques [3]. In the past, AI has already been used for generating timing data to align multiple streams, saving computing time compared to traditional mechanisms [6].

Forecasting network states for the media capture process allows modifying according to the QoS parameters regarding the capture, storage and transmission. To avoid bottlenecking the network or source computing devices, the rates at which data are read, captured and sent may be properly configured.

Data that are less necessary could be skipped or dropped to avoid wasting network bandwidth. Data can also be held back (paused) and transmitted later when enough bandwidth is available. Alternatively, reducing the number or quality of streams can optimize the employed bandwidth without failing to deliver media on time [3], [20]. Instead of skipping or pausing the media playout, forecasting when the available bandwidth will decrease can help optimize the transmitted streams (e.g., reducing video bitrate), thus avoiding situations where the destinations do not have enough media to playout.

Additionally, to save space, when storing, and bandwidth, when transmitting, media should be compressed. Although this process can add delay, AI-based compression can offer the same compression performance with a low required computing time [19]. Thanks to network forecasting, when to apply a specific level of compression can be anticipated (e.g., when near-future low network bandwidth is forecasted).

Data interleaving can ease the media sync but may be problematic regarding network bandwidth usage and the subsequent added latency that it takes to combine multiple streams at the source and the reverse process at the destination [3]. Forecasting when the network conditions will vary can help adapt the best transmission approach and configuration (e.g., interleaving, protocols…).

Algorithm 1. Media capture	
1	Read a media file or record from the capture device.
2	Read sync information or generate it using the media.
3	Data modification (Algorithm 2).
4	Data storage.
5	Send media data to destinations (Algorithm 3).
6	Measure and send parameters to sync service.

Algorithm 2. Modification	
1	Skip or pause individual or all streams.
2	Bandwidth optimization.
3	Data compression.

Algorithm 3. Transmission	
1	Determine the destinations.
2	Select transmission protocol.
3	Interleave streams together.
4	Send data through the network.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Media buffers underflow and overflow (where the media data are stored before being forwarded) should be avoided [3], [13]. Those situations can be prevented by controlling the generation, transmission or playout rates.

B. Media Destination

The clients in a shared media session receive the media from the sources, and other clients (when additionally acting as sources). Processes like modification (e.g., skipping, decompression, etc.) and storage can also happen at the destination to aid the media playout (as mentioned in Section III). Algorithm 4 shows how this process should proceed in the destinations.

When some data are missing, extrapolation could use previous data to generate them [3], [19]. Packet losses or network delay may require missing or late data to be generated in time at the destinations. For example, when transmitting video with multiple different resolution levels, AI-based super-resolution techniques can increase the quality using lower quality frames [19], [21]. As the extrapolation can take time, forecasting the network allows for realizing this action beforehand.

At the destination, the playout rate can also be dynamically adjusted [3], [13] to avoid asynchronies noticeable by users. Applying AMP when the forecasts indicate possible future playout async, or buffer underflow or overflow, could directly avoid those problems before they arise [13].

Otherwise, skipping can be used when media data are not necessary or when increasing the playout rate is not enough to restore the sync on time. With forecasting, this situation should be avoided by applying AMP earlier in time. Likewise, playout can be paused when reducing the playout rate is not enough, or when further media data are not being received. Forecasting will prevent pausing by maintaining QoS requirements and adjusting the playout rate.

C. Media Sync

In any case, the different media streams should be played synchronously at the destinations. Video players, speakers, multi-sensorial effects generators, and so on, should be in sync thanks to the media sync techniques. To sync the playout processes of different streams related to the same destination (e.g., audio-to-video sync), the destination can calculate and apply the corresponding sync adjustments when needed.

Usually, monitoring the playout of multiple devices and deciding how and when to make the sync adjustments happens outside of those devices. The most common approach is to have a centralized server or sync node (often the source of the main media, although not necessary). That server keeps track of a playout reference (i.e., the time the media playout should be at). Destinations periodically send their current playout state (as part of the monitored data), and then the server decides and notifies the destinations that need to correct their playout to be in sync with that reference. Alternatively, destinations can exchange their current playout points and decide how and when to adjust sync their media playout processes [15].

Table I showed the summary of different media sync techniques that can be applied in shared media scenarios [15]. Algorithm 5 expresses how to apply the AI-based forecasting models at either the sync server (centralized scheme) or at the destinations (distributed scheme).

Algorithm 4. Media Playout

```
1  If no new data is available to play:
2      Use previous data or extrapolate new data.
3  Data modification (Algorithm 2).
4  Store data (for possible future use).
5  Play new data if available.
6  Apply media sync mechanisms (e.g., AMP).
7  Measure and send parameters to sync service (Section IV.C).
```

Algorithm 5. AI-based Sync

```
1   Store new monitored data.
2   Use recent data with the AI models.
3   Read forecasts of the AI model for all participants.
4   For every media destination:
5       If predicted async playout:
6           Notify playout adjustments to apply.
7           Notify media session QoS adjustments.
8       If predicted bad network conditions:
9           Notify media session QoS adjustments.
10  For every media source:
11      If predicted bad network conditions:
12          Notify media session QoS adjustments.
```

The AI models will process the monitored data (see Table IV) after receiving them or at a schedule. Based on what the models forecast, the sync node will determine required adjustments to the playout (if needed) and transmit the proper messages to the playout devices. Tolerable async can vary depending on context and users. For example, async from 80ms to 160ms between video and audio playout (i.e., lip-sync) are noticeable but tolerable by users, while async exceeding 160ms are considered intolerable [13].

Besides notifying the media playout processes about the need of corrections to avoid future async, media sources could also be notified by the sync nodes for configuring their media streaming-related parameters as preventive measure [17], [20]. For example, if a future decrease in available bandwidth for a destination is forecasted, the media sources could apply bitrate reduction to the media streams sent to that destination.

D. Data Collection for AI Development

Sync control messages are generated by the nodes in charge of the media sync (e.g., the main media source, multiple media destinations or an independent server), represented with the Sync Node in Figure 2. The Sync Node is where the AI model is executed. It processes the monitored network and playout data and transmits sync control messages to notify the required corrections to the corresponding playout devices to avoid async. Feature engineering and model training are processes that happen outside the shared media consumption scenario. Preprocessing and employing the data to train AI-based models requires previously obtained data. After training, the models can be deployed and used in the media systems, although new data can still be collected for continuous learning purposes. Continuous learning can happen again outside the shared session, or in real-time to adapt fast to network changes.

As the network state may vary over time, the accuracy of the forecasts may start to decay. Monitoring the performance of the AI-models is important so that the AI-based techniques adapt and continue to improve.

Regarding data collection, Table IV summarizes most of the information that can be measured in the sources and destinations during a shared media session. These monitored parameters can be used for training AI models and measuring the sync performance.

Table IV
MONITORED PARAMETERS

Group	Parameter	Example Units
Network	Transmission delay.	Milliseconds (ms).
	Reception delay.	ms.
	Round-Trip time.	ms.
	Latency variance (Jitter).	ms.
	Maximum transmission bandwidth.	Mbps.
	Maximum reception bandwidth.	Mbps.
	Current bandwidth usage.	Mbps, %.
	Packet loss rate.	%.
	Data corruption rate.	%.
	Number of duplicated packets.	Packets, %.
	Number of unordered packets.	Packets, %.
	Number of current connections.	Connections, %.
	Maximum number of connections.	Connections.
	Network topology.	Mesh, Hybrid.
Connection	Supported communication protocols.	HTTP, TCP.
	Time taken to start connection.	Seconds, ms.
	Source.	IP, Port, identifier.
	Destination.	IP, Port, identifier.
	Transmission rate.	Packets or bits / s
	Reception rate.	Packets or bits / s
	Packet inter-arrival time.	ms.
	Times of packet transmission.	Timestamp.
	Times of packet reception.	Timestamp.
	Start and end of the communication.	Timestamp.
	Compression.	True/False, LZ77.
	Encryption.	True/False, RSA.
Device	Computing capability.	CPU speed.
	Current computing usage.	GHz, MOps, %.
	System memory (e.g., RAM).	GB, MB.
	Current memory usage.	GB, %.
	Storage.	HDD/SSD, TB.
	Connectivity.	4G, Ethernet.
	Battery level.	%.
	Battery charging.	True/False.
	Device type.	Phone, Laptop.
	Operating system.	Android, Mac.
	Screen resolution.	Width × Height.
	Camera resolution.	MP.
	Mulsemedia features.	Haptics, scents.
	Geographic location.	GPS coordinates.
Media	Media (e.g., video) duration.	Minutes, Hours.
	Video resolution.	Width × Height.
	Video format.	MP4, AVI.
	Audio format.	MP3.
	Framerate.	FPS.
	Bitrate level.	Kbps.
Playout	Current video playout time.	Seconds, ms.
	Content in buffer.	Frames, ms, %.
	Content successfully played.	Frames, %.
	Content dropped or corrupted.	Frames, %.
	Video playing/paused.	True/False.
	Video playout speed.	x1, x1.5, x0.75.
	Volume level.	dB, %.
AI Sync	Media sync technique employed.	AMP, Bucket Time.
	Forecast moment and prediction.	Async value, ms.
	Accuracy of the forecast.	NMAE, %.
	Performance of the preventive action.	RMSE, %.

E. Data Analysis

The proposed framework has been partially implemented and tested in a web-based shared media platform. The platform includes a video player, playing the same series of videos between multiple clients, with (aggressive) sync adjustments when playout async is detected by the server.

This platform included gathering of most of the parameters indicated in Table IV, which can be obtained from probing the network (e.g., ping) and measuring from the browser (e.g., video buffer load). The server in charge of monitoring those parameters receives them from all the connected clients and stores them in a database for posterior analysis. Furthermore, as currently no AI-based forecast model is implemented, the real-time monitored parameters are yet to be used to forecast the network or playout async. So, it only corrects the playout async when it is detected above a tolerable threshold.

To analyse the platform under different network conditions (which may impair the proper playout) the platform simulates different adverse network conditions (high network delay and jitter, low bandwidth, etc.). This way, playout async can be studied in the context of common network drawbacks.

Five (5) personal computers (PC) were connected over the period of a month, with randomly varying network states being simulated (lag, packet loss, bandwidth decay, jitter). Most of the metrics were captured at a rate of 300ms. Those metrics that were not captured at 300ms, were metrics like the number of CPU cores and such, that are bound to not change drastically over time.

The data captured in the database (1.5GB) was exported for processing and gathering insights. After cleaning and converting the data, and interpolating missing values, the resulting 4.5GB of data were further filtered to clean outliers and balancing (partially removing redundancies) to 2.8GB of final data for our analysis. The data were then checked for correlations (presented in Table V) and for the distributions of each variable (how the values behave, e.g., if they change fast, or remain constant).

The main aim of the collected dataset is to train AI models (to be developed in the future) that will forecast playout async so adjustments can be made beforehand (e.g., change playout rate, compression, filtering, etc.), aiming to prevent the occurrence of predicted async. Those models will be developed, trained and tested in the platform by using the gathered data until they have good performance in playout async forecasting.

When they are available in the future, the currently implemented reactive sync in the platform will be replaced by the AI-based sync solutions using those novel AI-based models, passing to the model the parameters that are monitored in real-time and executing the proper decisions (e.g., changing playback rate).

After running a few tests on the developed platform, some insights were obtained. With the collected data, the authors confirmed the already mentioned relationship between the network connection impairments and the media playout async. For example, reduced network bandwidth reduces video buffer load, probably leading to buffer underflow, with the consequent pauses in the video playout, very annoying to the users. Additionally, initial analysis showed a lot of correlation between some parameters which is useful for determining which parameters to measure (e.g., the more accessible). For example, if network bandwidth cannot be measured, measuring buffer load rate could be used instead.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Table V
PARAMETERS MOST CORRELATED WITH MEDIA PLAYOUT ASYNC

Parameter	Correlation (%)
Content in buffer	19.12
Round trip time	9.66
Download bandwidth usage	8.63
Latency varyance (Jitter)	7.34
Reception delay	5.18
Content dropped or corrupted	2.42
Current memory usage	2.07

Table V summarizes the parameters obtained during the media session that are most correlated with the media playout async. These will be the main employed for training AI-models in the future.

V. CONCLUSIONS AND FUTURE RESEARCH

Networked multimedia systems enable social interaction and sync communications with different devices, content types, and delivery technologies. An important challenge is ensuring synchronized shared consumption and interactions. Existing sync techniques focus on traditional media and are reactive.

Preventive sync can anticipate and avoid async playout, improving QoE. Despite their potential, AI solutions remain underexplored. With increasing data and computing power, more advanced AI forecasting techniques can be integrated into multimedia systems.

This paper proposes combining traditional sync techniques with AI models to enhance existing solutions, such as compression for bandwidth optimization. A platform based on the proposed framework is currently being implemented for testing the forecast solutions. Future work will define AI models, develop sync techniques within this framework, and evaluate their performance. Real-world assessments will demonstrate the framework's effectiveness, further improving media sync.

ACKNOWLEDGEMENTS

This work was supported, in part, by the following grants: ACIF/2021/192 from Programa I+D+i de la Generalitat Valenciana (GVA) and the European Social Fund (ESF); CIAICO/2022/025, from Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital of the GVA (DOGV 8919/05.10.2020); and grant PID2021-126645OB-I00, funded by MICIU/AEI/10.13039/501100011033/ and by "ERDF A way of making Europe".

REFERENCES

[1] J.-M. Chung, Emerging Metaverse XR and Video Multimedia Technologies. Berkeley, CA: Apress, 2023.
[2] M. W. Akhtar, S. A. Hassan, R. Ghaffar, H. Jung, S. Garg, and M. S. Hossain, "The shift to 6G communications: vision and requirements," Human-centric Comput. Inf. Sci., vol. 10, no. 1, 2020.
[3] G. Manfredi, et al., "Live Streaming Synchronization Using Event-Triggered Consensus Control," IEEE Transactions on Control of Network Systems, vol. 12, no. 2, pp. 1381–1392, 2025.
[4] Y. Yang, "A New Network Traffic Prediction Approach in Software Defined Networks," Mob. Networks Appl., vol. 26, no. 2, pp. 681–690, 2021.
[5] M. Onur, et al., "A new hybrid approach combining GCN and LSTM for real-time anomaly detection from dynamic computer network data," Computer Networks, vol. 268, no. August, 111372, 2025.
[6] C. K. Sahu, C. Young, and R. Rai, "Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review," Int. J. Prod. Res., vol. 59, no. 16, pp. 4903–4959, 2021.
[7] S. Izadi, M. Ahmadi, and A. Rajabzadeh, "Network Traffic Classification Using Deep Learning Networks and Bayesian Data Fusion," J. Netw. Syst. Manag., vol. 30, no. 2, pp. 1–21, 2022.
[8] M. Moulay, R. G. Leiva, P. J. Rojo Maroni, F. Diez, V. Mancuso, and A. Fernández Anta, "Automated identification of network anomalies and their causes with interpretable machine learning: The CIAN methodology and TTrees implementation," Comput. Commun., vol. 191, no. May, pp. 327–348, 2022.
[9] M. K. Hasan, M. M. Ahmed, A. H. A. Hashim, A. Razzaque, S. Islam, and B. Pandey, "A Novel Artificial Intelligence Based Timing Synchronization Scheme for Smart Grid Applications," Wirel. Pers. Commun., vol. 114, no. 2, pp. 1067–1084, 2020.
[10] B. Hu, Z. Sun, and J. Liu, "Distributed time synchronization algorithm based on sequential belief propagation in wireless sensor networks," Comput. Commun., vol. 176, no. September 2020, pp. 119–127, 2021.
[11] X. Wang et al., "Deep Learning-Based Network Traffic Prediction for Secure Backbone Networks in Internet of Vehicles," ACM Trans. Internet Technol., vol. 22, no. 4, 2022.
[12] U. Haider, M. Waqas, M. Hanif, H. Alasmary, and S. M. Qaisar, "Network load prediction and anomaly detection using ensemble learning in 5G cellular networks," Comput. Commun., vol. 197, no. December 2021, pp. 141–150, 2023.
[13] Z. Huang, K. Nahrstedt, and R. Steinmetz, "Evolution of temporal multimedia synchronization principles," in MediaSync: Handbook on Multimedia Synchronization, M. Montagud, P. Cesar, F. Boronat, and J. Jansen, Eds., Cham: Springer International Publishing, 2018, pp. 33–71.
[14] J. Escobar, C. Partridge, and D. Deutsch, "Flow synchronization protocol," IEEE/ACM Trans. Netw., vol. 2, no. 2, pp. 111–121, Apr. 1994.
[15] J. González, et al., "Key Technologies for Networked Virtual Environments," Multimed. Tools Appl., no. Iim, 2023.
[16] B. Gezici and A. K. Tarhan, "Systematic literature review on software quality for AI-based software," Empir. Softw. Eng., vol. 27, no. 3, 2022.
[17] L. Wang, A. Durning, and D. T. Delaney, "ML-based Video Streaming QoE Modeling with E2E and Link Metrics," 2022 30th Int. Conf. Software, Telecommun. Comput. Networks, SoftCOM 2022, pp. 1–8, 2022.
[18] "ISO/IEC 23005-3:2019 Part 3: Sensory information." 2019. [Online]. Available: https://www.iso.org/obp/ui/
[19] H. J. Kim, J. W. Kang, and B. U. Lee, "360° Image Reference-Based Super-Resolution Using Latitude-Aware Convolution Learned from Synthetic to Real," IEEE Access, vol. 9, pp. 155924–155935, 2021.
[20] A. Bentaleb, et al., "Solutions, Challenges, and Opportunities in Volumetric Video Streaming: An Architectural Perspective," ACM Transactions on Multimedia Computing, Communications and Applications, vol. 21, no. 7, pp. 1–35, 2025.
[21] J. Jiang, C. Wang, X. Liu, and J. Ma, "Deep Learning-based Face Super-resolution: A Survey," ACM Comput. Surv., vol. 55, no. 1, pp. 1–36, 2021.
[22] J. K. P. Seng, K. L. Ang, E. Peter, and A. Mmonyi, "Artificial Intelligence (AI) and Machine Learning for Multimedia and Edge Information Processing," Electronics, vol. 11, no. 14, p. 2239, Jul. 2022.
[23] Z. Liu, Q. Z. Sheng, X. Xu, D. Chu, and W. E. Zhang, "Context-Aware and Adaptive QoS Prediction for Mobile Edge Computing Services," IEEE Trans. Serv. Comput., vol. 15, no. 1, pp. 400–413, 2022.
[24] J. Zhong and S. Chen, "Efficient multi-category packet classification using TCAM," Comput. Commun., vol. 169, no. January, pp. 1–10, 2021.
[25] R. Tripathi and K. Rajawat, "Adaptive Network Latency Prediction from Noisy Measurements," IEEE Trans. Netw. Serv. Manag., vol. 18, no. 1, pp. 807–821, 2021.

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

DE-RCM: Desanonimización Explicable de MACs Aleatorias en 802.11 WLANs

Juan Manuel Montes-Lopez, Pablo Serrano, Marco Gramaglia, Aruna Prem Bianzino
Departamento de Ingeniería Telemática
Universidad Carlos III de Madrid
Avda. Universidad, 30. 28911 Leganés.
{jumontes, pablo, mgramagl, abianzin}@it.uc3m.es

El uso de WiFi puede revelar patrones de comportamiento que comprometen la privacidad. Para abordar este problema se han propuesto esquemas de Direcciones MAC Aleatorias y Cambiantes (RCM, por sus siglas en inglés) para modificar periódicamente las direcciones MAC, con el objetivo de desacoplar la identidad del dispositivo de sus patrones de uso. En este artículo diseñamos DE-RCM, un esquema que desanonimiza de forma explicable MACs aleatorias a partir datos históricos del comportamiento de los dispositivos WiFi. DE-RCM se basa en *random forests* y ofrece una precisión del 82% con tan sólo dos semanas de datos para el conjunto de 100 dispositivos considerado.

***Palabras Clave*—WiFi, RCM, desanonimización**

I. INTRODUCCIÓN

La actividad inalámbrica de los dispositivos móviles deja un rastro de información que puede utilizarse para identificar inequívocamente a los usuarios. Por ejemplo, para el caso de redes móviles, el trabajo [1] muestra que cuatro puntos espacio-temporales son suficientes para identificar al 95% de los individuos, mientras que para el caso de WiFi, en [2] se analiza cómo la superposición de redes preferidas revela vínculos sociales entre individuos.

Para evitar la relación tan estrecha entre comportamiento espacio-temporal e identidad, recientemente se ha propuesto el uso de Direcciones MAC Aleatorias y Cambiantes (RCM, por sus siglas en inglés), cuya idea clave es modificar la dirección MAC empleada en WiFi para reducir la cantidad de información asociable a un dispositivo. Existe una amplia variedad de esquemas RCM [3]; a modo de ejemplo, los dispositivos Android permiten tanto un esquema *persistente*, donde el dispositivo genera una dirección aleatoria por cada red (evitando unir información de diferentes redes), como un esquema *no persistente*, donde la dirección se cambia cada vez que se asocia a una red o pasadas 24 horas (con el objetivo de anonimizar dentro de la propia WiFi).

En este artículo, analizamos si, incluso bajo esquemas no persistente, existen patrones de uso que puedan servir para identificar inequívocamente a un usuario dentro de la misma red WiFi (nótese que tanto [1], [2] analizaron escenarios mucho más amplios). Para ello, analizamos las trazas de uso real de un conjunto de 28 voluntarios en un edificio universitario. Analizamos diferentes características diarias de sus dispositivos e identificamos ciertas características que pueden servir para identificar parcialmente un subconjunto de dispositivos. Aprovechando estas observaciones, diseñamos DE-RCM: Desanonimización Explicable de Direcciones MAC Aleatorias, basado en *random forests*. Evaluamos su rendimiento emulando un esquema RCM no persistente con un período diario, alcanzando una precisión en la desanonimización de hasta el 80%, en función de la cantidad de información previa disponible sobre la actividad del dispositivo. Gracias al uso de *random forests* y su correspondiente explicabilidad, nuestros resultados pueden ayudar a diseñar mejores algoritmos RCM.

II. DATOS EMPLEADOS

Los datos se obtienen de la red `eduroam` del Campus de Leganés de la Universidad Carlos III de Madrid (UC3M), que consta de 278 puntos de acceso (APs) desplegados en siete edificios,[1] la mayoría de los cuales tienen al menos tres plantas, y cada AP proporciona cobertura para aproximadamente 100 m^2. Cada vez que un dispositivo se asocia o reasocia con un AP diferente, el AP notifica a un servidor RADIUS; además, este servidor actualiza el estado de cada dispositivo con una periodicidad de al menos 15 minutos. Además de la información espacio-temporal, a saber, marca de tiempo, identificador de usuario, direcciones MAC del dispositivo y del AP al que está conectado, el servidor RADIUS también recopila el número total de bytes transmitidos y recibidos.

[1] https://www.uc3m.es/life-on-campus/campuses-plans/leganes

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En este trabajo nos centramos en 28 voluntarios que aceptaron esta recopilación de datos, en su mayoría personal de investigación, lo que resultó en un total de 98 dispositivos diferentes durante un período de 15 meses. Obtenemos la información del servidor RADIUS a través de una API que protege las identidades de los usuarios y los dispositivos realizando un hash MD5 de los identificadores y las direcciones. Para simplificar, decidimos restringir nuestro análisis a un único edificio con 47 APs, ya que el 95% de los voluntarios pasaron la mayor parte de su tiempo asociados a alguno de estos APs.

III. ATRIBUTOS PARA LA DESANONIMIZACIÓN

En esta sección, analizamos diversos atributos de las trazas de los dispositivos que podrían servir para caracterizarlos inequívocamente.

A. Horas de llegada y salida

Primero analizamos tres características temporales: la hora de la primera y última conexión y el tiempo total de conexión. Estos aspectos son relevantes en nuestro entorno académico, donde los horarios y hábitos deberían reflejarse en las trazas obtenidas.

Nuestros resultados muestran que la mayoría de los dispositivos se conectan por primera vez alrededor de las 8 AM, explicable dado que la mayoría de los usuarios son investigadores. Se observa cierta variabilidad en las horas de conexión y desconexión entre las 6 AM y las 6 PM, lo que podría corresponder a los horarios menos regulares de algunos estudiantes. Por su parte, las horas de desconexión tienden a concentrarse alrededor de las 6 PM. Por último, en cuanto al tiempo total de conexión, lo más reseñable es que hay dispositivos que permanecen conectados durante la 24 horas, lo que sugiere que son equipos de escritorio o portátiles que permanecen en la red de forma continua.

En resumen, se confirma que la información temporal permite identificar parcialmente ciertos tipos de dispositivos, pero para una identificación más precisa se requiere información adicional.

B. Número de APs diferentes visitados

A continuación, analizamos el comportamiento de los dispositivos mientras están conectados a la red. En primer lugar, nos centramos en el grupo de dispositivos que permanecen siempre conectados, los cuales representan el 16% del total. Observamos que estos dispositivos se conectan constantemente a un único punto de acceso (AP), lo que sugiere que se trata de ordenadores portátiles o de escritorio sin conexión cableada.

Para el resto de dispositivos, calculamos el número de APs diferentes a los que se asocia (denotado como #APs) durante un período de tiempo de longitud T, que varía entre 1 día, 1 semana, 2 semanas y 1 mes. Más específicamente, calculamos para cada dispositivo el número medio de #APs para 250 períodos diferentes de longitud T[2]. Según los resultados: a) al considerar $T = 1$ día, #APs varía entre 1 y 10.3, con una media de 3.3 APs (nótese que nuestro análisis se limita a un solo edificio); b) al extender a $T = 1$ semana, los valores de #APs prácticamente se duplican, lo que indica una variación significativa al pasar de un patrón diario a uno semanal; c) para $T = 2$ dos semanas, el incremento de #APs de del 50%, mientras que d) para $T = 1$ mes, los resultados son prácticamente idénticos a los del caso anterior.

De acuerdo con estos resultados, se confirma que tanto un día o una semana pueden ser insuficientes para capturar la movilidad de un dispositivo, mientras que más de un mes puede resultar redundante.

C. Entropía

En la sección anterior se analizó el número de APs diferentes a los que se conectan los dispositivos, sin tener en cuenta el tiempo que pasa en cada AP. En esta sección se considera tanto el número de APs como el tiempo que pasa en cada uno mediante el cálculo de la *entropía H* de la traza, definida como [4] $H = \sum_{i=1}^{\#AP} p_i \log_2(p_i)$, donde p_i representa la cantidad relativa de tiempo que el dispositivo pasa en el AP i.[3]

Se representa la Función de Distribución Acumulada (CDF) de la entropía para cada dispositivo utilizando las mismas ventanas de tiempo consideradas anteriormente en la Fig. 1. Los principales resultados son: (1) independientemente T, casi el 20% de los dispositivos tienen una entropía cercana a cero, es decir, prácticamente no se mueven[4]; (2) la entropía crece con T, es decir, los dispositivos visitan más lugares diferentes a medida que se considera más tiempo; (3) hay poca diferencia en términos de entropía entre dos semanas y un mes, lo que sugiere que un valor umbral $T = 2$ semanas para capturar el patrón de movilidad de un dispositivo[5]; (4) además de los dispositivos con entropía cercana a cero, existe una notable variedad en términos de entropía entre los dispositivos, lo que confirma que esta métrica también podría servir para re-identificar dispositivos.

D. APs más frecuentes

Aquí analizamos los k APs a los que un dispositivo pasa más tiempo conectado, motivado por el trabajo previo de [2] que identificó que la lista de SSIDs preferidos de un dispositivo puede servir para revelar relaciones entre dispositivos. El primer desafío es el cálculo de k, ya que un valor muy pequeño no distinguirá entre dispositivos (p.ej.,

[2]Para minimizar posibles sesgos en nuestros resultados, seleccionamos aleatoriamente 250 períodos de T días de duración para cada dispositivo.

[3]Una entropía baja indica que se pasa la mayor parte del tiempo conectado a pocos APs, mientras que un valor alto indica que el tiempo de conexión está distribuido de manera más uniforme entre múltiples APs. Dado un valor de H, se puede estimar un *número efectivo* de ubicaciones distintas $L = 2^H$, que representa el número de $\#APs$ que, si se visitaran por igual, darían el mismo valor de H.

[4]Nótese que en la Sección A identificamos que el 16% de los dispositivos siempre están conectados a la red. Aquí hemos identificado que algunos dispositivos, *cuando están conectados*, prácticamente siempre están conectados al mismo AP.

[5]Por ejemplo, la entropía al considerar un día es $H = 0.40$, es decir, una movilidad efectiva a lo largo de $L = 2^{0.40} = 1.32$ APs, mientras que para los casos de una semana, dos semanas y un mes, la movilidad efectiva sería de 1.74, 1.96 y 1.99, respectivamente.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

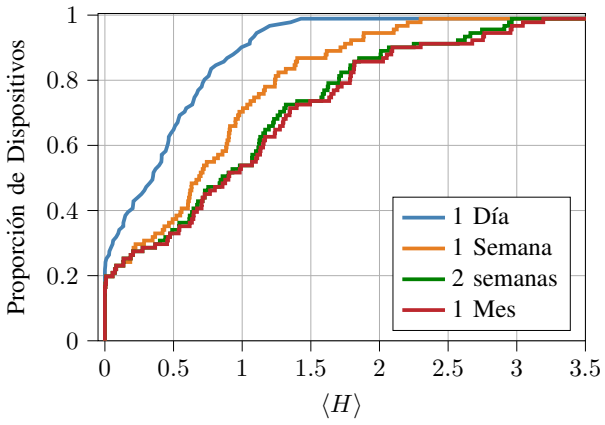


Fig. 1. ECDF de la Entropía H.

usuarios que comparten despacho), mientras que un valor excesivamente grande podría clasificar como diferentes dos días distintos del mismo usuario. Para determinar el valor adecuado de k, consideramos los seis APs más utilizados y calculamos el porcentaje promedio de tiempo que pasan conectados a ellos, lo que revela que la mayoría de los dispositivos pasan alrededor del 70% en el AP más frecuente y un 20% adicional en el segundo AP más común. En consecuencia, un dispositivo dedica el 90% de su tiempo de conexión en $k = 2$ APs, valor que usaremos para nuestro análisis posterior.

A continuación, analizamos para este valor de k cuántos dispositivos son *únicos*, es decir, ningún otro dispositivo tiene los mismos dos APs más frecuentes en el mismo día. Según nuestros resultados, el 88% de los dispositivos tienen una pareja de APs que los hacen únicos cada día. También medimos cuántas veces a lo largo de los días un dispositivo tiene el mismo par de APs más frecuentes: sólo el 33% de los dispositivos. De esta manera, aunque la tupla de $k = 2$ de los APs más comunes puede servir para caracterizar un dispositivo en un día determinado, esta característica es insuficiente para identificar unívocamente a un dispositivo a lo largo de varios días.

E. Tráfico descargado

Como última característica que pueda servir para identificar a los dispositivos analizamos el tráfico descargado por día (previa confirmación que el tráfico de subida es directamente proporcional al tráfico descargado). Según nuestros resultados preliminares, la distribución sigue una *ley de potencias*, que revela una cola pesada sugiriendo tanto la presencia de tráfico de *streaming* o actualizaciones importantes de software, como transferencias de datos de mucha menor envergadura (como navegación o actividades en segundo plano). De esta manera, el tráfico descargado también podría servir para identificar a ciertos usuarios. Además, observamos que, en promedio, los dispositivos consumen aproximadamente 350 MB de datos mientras están conectados a la red (ligeramente por debajo del consumo diario promedio en España de 400 MB[6]).

IV. DISEÑO DE DE-RCM

En esta sección se presenta DE-RCM, el algoritmo de desanonimización para reidentificar un dispositivo comparando sus características diarias actuales con las calculadas a partir de la actividad de los últimos T días. DE-RCM se basa en *Random Forests*, método basado a su vez en Árboles de Decisión que mejora el rendimiento y la robustez al promediar las predicciones de múltiples árboles entrenados en diferentes subconjuntos de datos. Para interpretar los resultados, utilizamos LIME [5], que aproxima las predicciones del modelo con un modelo interpretable alrededor de instancias específicas. En este trabajo, decidimos utilizar *random forests* con árboles poco profundos para la clasificación, debido a su mayor precisión y eficacia con datos limitados [6].

A. Generación de instancias

Calculamos el *perfil* diario de dispositivo basándonos en los atributos presentados. El perfil del dispositivo i en un día se define como $\mathbf{x}_i = (T_s,\ t_a,\ \#APs,\ H, [APs], D)$, donde T_s representa el tiempo total en la red, t_a la hora de llegada, $\#APs$ el número de puntos de acceso diferentes visitados, H la entropía, $[APs]$ los dos puntos de acceso más comunes, y D el tráfico total descargado. La desanonimización requiere determinar si dos perfiles pertenecen al mismo dispositivo o no, comparando sus características mediante un clasificador. Para generar instancias para el clasificador, dados dos perfiles $\mathbf{x}_i$ y $\mathbf{x}_j$ se calcula la tupla $(\mathbf{\Delta x}_{i,j}, \delta_{i,j})$, donde $\Delta \mathbf{x}_{i,j}$ representa la diferencia numérica entre los perfiles[7] comparados, mientras que $(\delta_{i,j})$ es una etiqueta binaria, con un valor de 1 si los perfiles corresponden al mismo dispositivo y 0 en caso contrario.

B. Proceso de inferencia

Una vez entrenado el modelo con información de días previos, comparamos el perfil de un dispositivo desconocido con los perfiles del conjunto de entrenamiento, calculando las diferencias correspondientes y asignando al dispositivo desconocido la identidad del dispositivo con menor diferencia. Matemáticamente, sea $\{d\}$ el conjunto de dispositivos conocidos y $\{\mathbf{x_t}\}_d$ como el conjunto de perfiles para un dispositivo específico d del conjunto de entrenamiento. Sea d_n el dispositivo desconocido y su vector de características $\mathbf{x_n}$. La desanonimización consiste en encontrar el dispositivo con la mayor probabilidad promedio de coincidir, esto es:

$$d_n = \arg\max_d \left(\frac{1}{T_d} \sum_{t=0}^{t=T_d} P(y = 1 | \Delta \mathbf{x_{t,n}}) \right) \qquad (1)$$

donde $\Delta \mathbf{x_{t,n}}$ es la diferencia entre el perfil del dispositivo d en el día t y el perfil del nuevo dispositivo d_n y $P\left(y = 1 | \Delta \mathbf{x_{t,n}}\right)$ es la probabilidad de que los perfiles correspondan al mismo dispositivo.

[6]https://bandaancha.eu/articulos/consumo-datos-moviles-linea-alcanza-12-11043

[7]Para $[APs]$, calculamos la superposición de la lista de APs, que varía de 0 (sin superposición) a 2 (misma lista)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

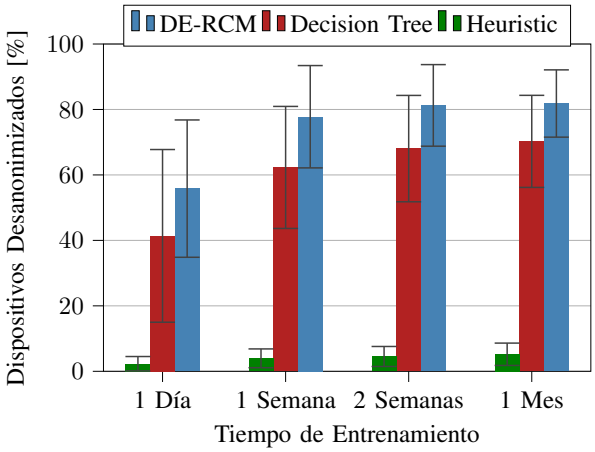


Fig. 2. Precisión para diferentes periodos de entrenamiento T.

V. ANÁLISIS DE PRESTACIONES

A continuacion evaluamos el rendimiento de DE-RCM, comparándolo con: *Decision Tree*, un modelo que clasifica dividiendo recursivamente en subconjuntos basados en los atributos, y *Heurístico*, un clasificador basado exclusivamente en los $k = 2$ APs más visitados.

A. Precisión

Primero, analizamos la precisión de DE-RCM para diferentes valores del período de entrenamiento T. Al igual que en los casos anteriores, para minimizar los sesgos, seleccionamos aleatoriamente 250 períodos de T días consecutivos, entrenamos el modelo en cada período y probamos su capacidad para identificar los perfiles de los dispositivos al día siguiente. La precisión se define como la proporción de dispositivos identificados correctamente sobre el número total de dispositivos en ese día.

En la Fig. 2 se presenta la precisión para DE-RCM y los otros métodos. Según los resultados, incluso con $T = 1$ día, DE-RCM alcanza un 55%, mucho más que una elección al azar (aprox. el 3%). La precisión mejora cuanto mayor es T, con más del 80% con 2 semanas. Como referencia, el rendimiento del heurístico es muy bajo, alcanzando como máximo una tasa de éxito del 5% (ligeramente superior al azar), mientras que el árbol de decisión alcanzar una precisión del 70%, es decir, un 10% peor que DE-RCM (y sin explicabilidad).

B. Explicabilidad

Aquí ilustramos la capacidad de DE-RCM para producir desanonimización explicable, aprovechando el peso relativo de los atributos LIME. Por simplicidad, nos centramos en $T = 2$ semanas. En la Fig. 3 representamos la explicación proporcionada por LIME para una identificación correcta del mismo dispositivo en diferentes días (figura de la izquierda), y una identificación correcta de diferentes dispositivos (figura de la derecha). Según estos resultados, la característica más influyente en el proceso de clasificación es la lista de los APs más comunes, $[APs]$. Más específicamente, la decisión de *mismo dispositivo* está

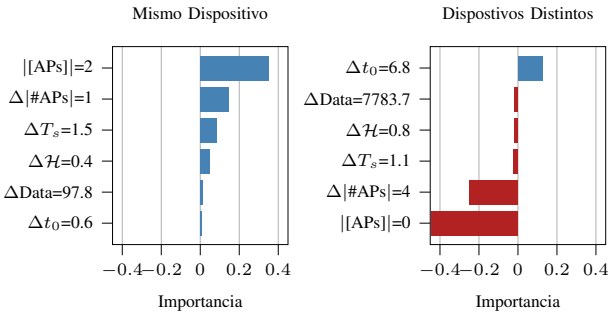


Fig. 3. Explicabilidad de las decisiones de DE-RCM.

motivada por tener una lista idéntica de los APs más comunes (fila superior) y valores similares para el resto de los parámetros (tiempo total, entropía, etc.), mientras que la decisión de *dispositivo diferente* está motivada por diferentes listas de los APs más comunes (fila inferior) y número de diferentes APs visitados, con valores similares para el resto de los parámetros.

VI. RESUMEN Y TRABAJO FUTURO

Hemos presentado un prometedor método para desanonimizar trazas que empleen RCM, validado con un conjunto reducido de usuarios. Para un trabajo futuro estamos ampliando el número de dispositivos estudiados, y empleando los resultados de explicabilidad para mejorar el diseño de los métodos RCM.

AGRADECIMIENTOS

Agradecemos a Rafael Calzada por su apoyo en el proceso de recolección y a Jose Furones por su guía en protección de datos. También agradecemos a Emmanuel Mera por los análisis iniciales. Este trabajo ha sido financiado por el Programa de Investigación e Innovación Horizon-JU-SNS-2022 de la Unión Europea (No. 101139198, iTrust6G), por el Ministerio de Asuntos Económicos y Transformación Digital de España y la Unión Europea-NextGenerationEU a través de los proyectos UNICO 5G I+D 6G-RIEMANN y 6G-SORUS, y por la Comunidad de Madrid (ORDEN 5696/2024) a través del proyecto TUCAN6-CM (TEC-2024/COM-460).

REFERENCIAS

[1] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," *Scientific Reports*, vol. 3, p. 1376, 2013.

[2] M. Cunche, M. A. Kaafar, and R. Boreli, "I know who you will meet this evening! Linking wireless devices using Wi-Fi probe requests," in *2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2012, pp. 1–9.

[3] J. Martin, T. Mayberry, C. Donahue, L. Foppe, L. Brown, C. Riggins, E. C. Rye, and D. Brown, "A study of mac address randomization in mobile devices and when it fails," 2017.

[4] X. Lu, E. Wetter, N. Bharti, A. J. Tatem, and L. Bengtsson, "Approaching the Limit of Predictability in Human Mobility," *Scientific Reports*, vol. 3, no. 1, p. 2923, Oct. 2013.

[5] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," Aug. 2016. [Online]. Available: http://arxiv.org/abs/1602.04938

[6] T. C. Au, "Random Forests, Decision Trees, and Categorical Predictors: The "Absent Levels" Problem," Oct. 2018, arXiv:1706.03492 [cs, stat]. [Online]. Available: http://arxiv.org/abs/1706.03492

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Control en bucle cerrado para mitigar alucinaciones en modelos de Inteligencia Artificial. Aplicación a pacientes virtuales.

Daniel García Torres[1], César Fernández Peris[2], María Asunción Vicente Ripoll[2], José Joaquín Mira Solves[1].
[1]Dpto. de Psicología de la Salud; [2]Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche.
Universidad Miguel Hernández de Elche
Av. de la Universidad s/n. 03202 Elche (Alicante).
daniel.garciat@umh.es, c.fernandez@umh.es, suni@umh.es, jose.mira@umh.es

Los modelos de inteligencia artificial generativa (IAG) facilitan el desarrollo de pacientes virtuales de gran realismo. En estas aplicaciones se debe garantizar alta fiabilidad (evitando errores o alucinaciones) y, al mismo tiempo, ofrecer respuestas naturales y creativas, para aumentar el realismo de las interacciones. Proponemos un mecanismo de control en bucle cerrado mediante una IAG secundaria para supervisar las respuestas de la IAG principal y reducir la temperatura de su modelo en caso necesario. Tras registrar 2549 interacciones estudiante/paciente con dos pacientes virtuales diferentes, se ha podido comprobar la viabilidad de la supervisión. El trabajo futuro se centrará en el diseño y dimensionamiento del sistema de control.

***Palabras Clave*- inteligencia artificial, supervisión, control en bucle cerrado, pacientes virtuales.**

I. Introducción

Los modelos de inteligencia artificial generativa (IAG) son capaces de producir contenido nuevo —como texto, imágenes o audio— a partir de instrucciones concretas, gracias a su entrenamiento con grandes volúmenes de datos. En el ámbito educativo sanitario, estos avances han revolucionado el desarrollo de pacientes virtuales conversacionales.

Un paciente virtual es una simulación informática que representa a un paciente real según un caso clínico predefinido. Este recurso permite que estudiantes de medicina, enfermería o psicología practiquen decisiones clínicas en un entorno seguro y sin riesgo para personas reales [1,2]. Su complejidad varía desde entornos textuales básicos hasta simulaciones multimedia altamente realistas [3], ofreciendo beneficios en razonamiento clínico, precisión diagnóstica y otras habilidades médicas [4]. Además, suponen una alternativa más eficiente frente a métodos costosos como el uso de actores o la participación en prácticas reales [5].

En concreto, los pacientes virtuales conversacionales, se centran exclusivamente en la simulación de consultas clínicas mediante texto, voz o vídeo. A diferencia de los actores humanos, estos sistemas permiten prácticas ilimitadas, en cualquier momento y lugar, con múltiples casos clínicos y bajo coste.

Con la aparición de modelos de IAG, como ChatGPT, ha sido posible generar pacientes virtuales realistas a partir de simples instrucciones o "prompts" [6,7]. Basta con describir una situación clínica para obtener una simulación coherente y personalizada. Un prompt tan simple como “adquiere el rol de una persona que acaba de sufrir un infarto y acude a urgencias” ya se comporta como un paciente virtual básico, que responde de modo muy parecido a un paciente real. Para un comportamiento más específico, basta con añadir detalles en el prompt: síntomas específicos para un caso clínico concreto, alergias a medicamentos, antecedentes personales y familiares, etc. En todos los casos, la IAG se comportará de acuerdo con la información que le proporcionemos, a la que sumará todo el conocimiento que ha obtenido durante su entrenamiento. No obstante, estos modelos pueden presentar errores, como mostrar síntomas incorrectos o sesgos derivados del entrenamiento, por ejemplo, representar síntomas masculinos en pacientes femeninas con infarto [8]. Aunque la IAG no diagnostica, es crucial vigilar estas inconsistencias para garantizar simulaciones clínicas precisas y educativas. Estos errores o respuestas incorrectas se conocen habitualmente como alucinaciones.

II. Temperatura y alucinaciones en IAGs

Todos los modelos de IAG permiten configurar su comportamiento para que sus respuestas sean más predecibles o bien más creativas. Los principales parámetros que permiten regular este comportamiento son la temperatura (valores bajos de temperatura producen resultados más determinísticos) y el *top-p* (*nucleus sampling*, valores bajos también producen resultados más probables y menos aleatorios). En general, no

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

se recomienda alterar ambos parámetros simultáneamente porque sus efectos están interrelacionados.

Una creatividad más elevada produce respuestas muy variadas para la IAG, con las que las conversaciones ganan en naturalidad, pero a cambio incrementa la posibilidad de que tales respuestas introduzcan alucinaciones o incorrecciones.

En ciertas aplicaciones, interesa que las conversaciones sean lo más naturales posible, pero al mismo tiempo se requiere reducir la probabilidad de aparición de alucinaciones al máximo. Como ejemplo de este tipo de aplicaciones, podemos mencionar las que ofrecen consejos de salud (en las que proporcionar información incorrecta puede tener consecuencias graves), o las educativas (en las que la información incorrecta puede dificultar el aprendizaje). Los pacientes virtuales se encuentran en este segundo grupo de aplicaciones: su objetivo es la formación del personal sanitario y, para ello, un paciente virtual debe comportarse del modo más parecido a un paciente real, sin inventar síntomas incorrectos que confundirían a los estudiantes.

Con el objetivo de reducir alucinaciones, pero mantener la variabilidad y naturalidad de las conversaciones en aplicaciones de IAG como las expuestas anteriormente, presentamos una propuesta de control en bucle cerrado mediante supervisión mutua entre diferentes modelos de IAG, que hemos aplicado al desarrollo de pacientes virtuales.

III. Arquitectura propuesta

El esquema estándar de un sistema de control en bucle cerrado es el mostrado en la Fig.1. Resumidamente, se lee de modo continuo la salida del sistema mediante un sensor y se compara con la consigna deseada. En función de la diferencia entre consigna y valor real de salida, el controlador actúa sobre el sistema.

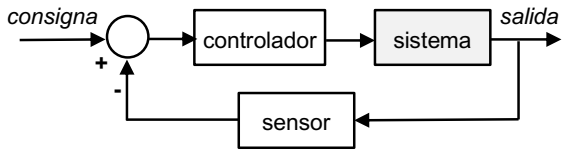


Fig. 1. Estructura estándar de control en bucle cerrado.

Proponemos una arquitectura basada en este concepto, en la que la consigna indica la probabilidad máxima de error admisible, el sensor estima el error que puede existir en cada respuesta de la IAG, y el controlador ajusta la temperatura para conseguir que la probabilidad de error se ajuste a la consigna.

La Fig. 2 detalla el esquema propuesto, sustituyendo el bloque "sistema" por la IAG1 (IAG principal, encargada de responder las preguntas del estudiante simulando ser un paciente); y el bloque "sensor" por la IAG2 (IAG alternativa, encargada de estimar la probabilidad de error en la respuesta de la IAG1). Cuando la probabilidad de error no cumple la consigna (es superior al umbral establecido), el controlador actúa reduciendo el parámetro temperatura en la IAG1 y pidiéndole una regeneración de su respuesta, de modo que las respuestas incorrectas no se llegan a mostrar al usuario. Este proceso se repite hasta que la respuesta cumple los requisitos en cuanto a probabilidad de error (o hasta que la temperatura de la IAG1 se ha reducido al mínimo posible). En nuestro prototipo, partimos de una temperatura estándar de 0.7 y la reducimos progresivamente hasta 0.1 en saltos de 0.2. El límite 0.1 se ha eligió para conservar una cierta aleatoriedad; y el incremento de 0.2 para que, en el peor de los casos, sólo se realicen cuatro llamadas a las IAGs (con temperaturas 0.7, 0.5, 0.3 y 0.1) antes de devolver una respuesta, evitando tiempos de espera excesivamente elevados.

Tanto la IAG1 como la IAG2 acceden al mismo contexto, que contiene tanto la descripción del paciente virtual (síntomas, historia clínica, etc.) como el historial de preguntas y respuestas anteriores de la conversación. La IAG2 utiliza un *prompt* diseñado específicamente para supervisión: "dado el contexto y la conversación anterior, estima la probabilidad de que esta respuesta sea errónea y devuelve un valor entre 0 y 100, donde 0 indica respuesta completamente correcta y 100 indica respuesta completamente errónea".

Para los experimentos de puesta a punto del sistema, registramos adicionalmente el razonamiento seguido por la IAG2 para estimar la probabilidad de error, con el objeto de realizar análisis posteriores.

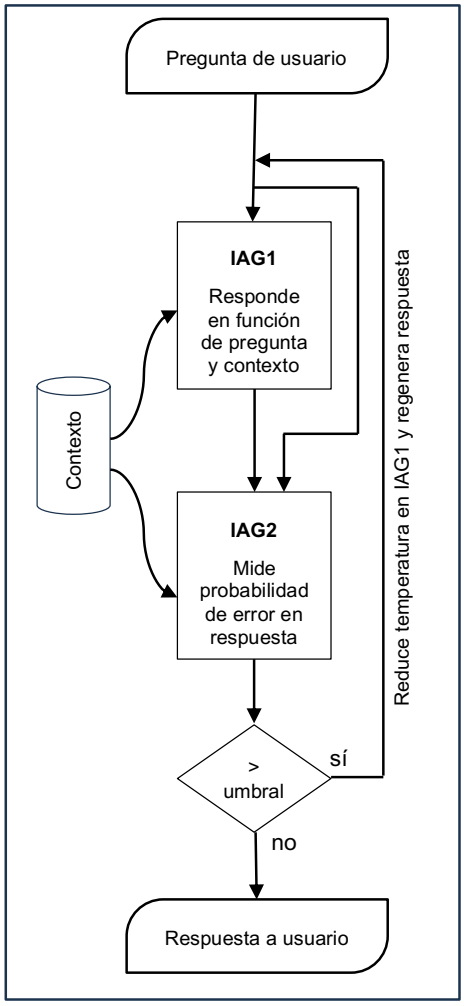


Fig. 2. Estructura de control propuesta.

IV. Datos de partida

Partimos de 15 pacientes virtuales desarrollados con anterioridad (sin control en bucle cerrado) con los que se han realizado prácticas de psicología y pediatría.

Cada paciente virtual representa una persona que acude a consulta y a la que cada estudiante debe entrevistar para poder

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

llegar a un diagnóstico. Algunos de los pacientes disponen de información adicional como analíticas, informes previos, etc. Tras terminar la entrevista, se presenta un cuestionario en el que hay que contestar un conjunto de preguntas que evalúan si la entrevista, el razonamiento, el diagnóstico y los tratamientos elegidos por los estudiantes han sido correctos, La Fig. 3 muestra capturas de pantalla de uno de los pacientes, en el que puede apreciarse la página de conversación y los botones que permiten acceder a información adicional o al cuestionario final.

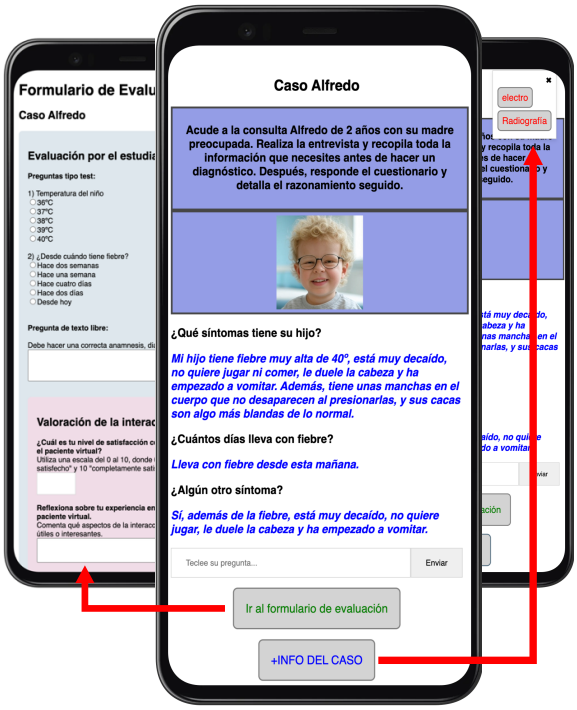


Fig. 3. Capturas de pantalla de uno de los pacientes virtuales.

En total, disponemos de 1970 prácticas realizadas con estos pacientes, de las que se han registrado las conversaciones completas (preguntas de estudiantes y respuestas de la IAG actuando como paciente virtual). Esas conversaciones se utilizarán como datos de partida para evaluar la viabilidad de un sistema de control como el que proponemos.

V. Resultados

Como primer paso, hemos analizado la fiabilidad del bloque "sensor" en nuestro esquema de control, esto es, la IAG2 encargada de estimar la probabilidad de error en las respuestas de la IAG1. Para ello, hemos recopilado las conversaciones realizadas durante las prácticas con los dos pacientes virtuales de pediatría (en total 2549 preguntas realizadas por los estudiantes junto con sus correspondientes respuestas) y hemos calculado la probabilidad de error para cada una de ellas con la IAG2. En las prácticas se utilizó como IAG1 el modelo *"GPT4o-mini"* de OpenAI (ChatGPT), dado que el modelo *"GTP4o"* planteó problemas en la utilización simultánea por muchos estudiantes en el aula. Para la supervisión se ha utilizado el modelo *"mistral_large_latest"* de Mistral, para evitar que las respuestas de una IAG fuesen evaluadas por ella misma. La temperatura utilizada durante las prácticas con la IAG1 se fijó en 0.7 (un compromiso entre variabilidad y fiabilidad ampliamente reconocido); mientras que, para la IAG2, dado que realiza tareas de supervisión, se fijó una temperatura más baja, concretamente 0.2. Todos los experimentos se repitieron 5 veces.

En la Fig. 4 se representan los resultados de la supervisión en cada una de las 5 repeticiones para las 2549 respuestas de la IAG1 analizadas por la IAG2 (2549*5 = 12745 resultados de supervisión). En la mayor parte de los casos, la probabilidad de error estimada es cero (6594 casos) o, como máximo, del 20%, pero existe un número considerable de respuestas que la IAG2 considera completamente erróneas; en total 855 análisis en los que la probabilidad de error estimada es del 100%.

Dado que se realizan 5 análisis diferentes para cada respuesta de la IAG1, es interesante analizar los valores promedio (Fig. 5) y, sobre todo, las desviaciones estándar (Fig. 6). En concreto, los valores de las desviaciones estándar nos indican una alta repetitividad en las medidas, con desviaciones estándar menores de 20 puntos en la mayor parte de los casos.

Por último, realizamos una comparación de las respuestas obtenidas en cada uno de los dos pacientes virtuales de pediatría por separado, tal y como se representa en la Fig.7.

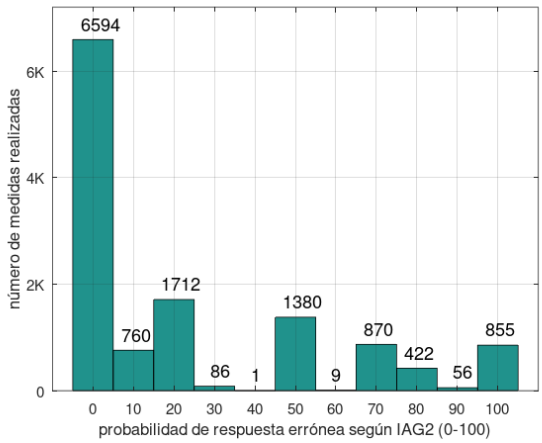


Fig. 4. Histograma de probabilidad de error en las respuestas.

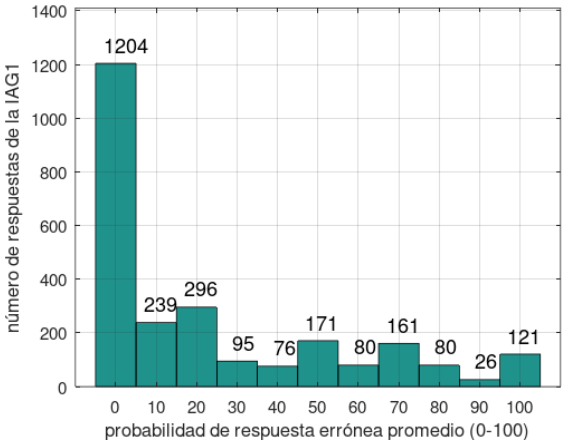


Fig. 5. Probabilidad de error promediada.

A la vista de los resultados comparados, queda patente que en el segundo paciente virtual se produce un mayor número de respuestas incorrectas. La única razón que puede justificar este

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

comportamiento es una definición poco precisa del caso clínico en este paciente por parte del profesor; que posibilita una mayor aleatoriedad en las respuestas de la IAG1 y también en las supervisiones de la IAG2. Indirectamente, el análisis de los resultados de las prácticas realizadas sirve como un mecanismo de control de calidad para los pacientes virtuales.

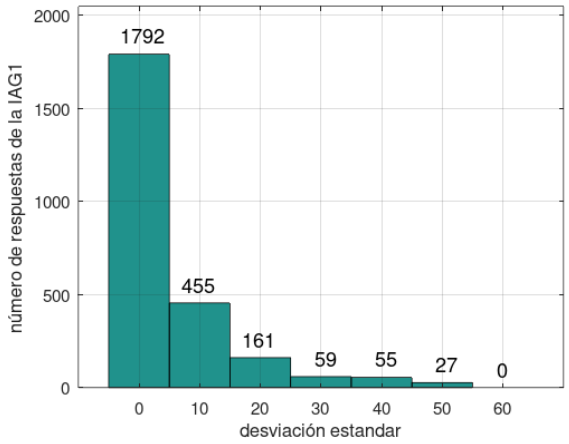


Fig. 6. Repetibilidad de las estimaciones de error.

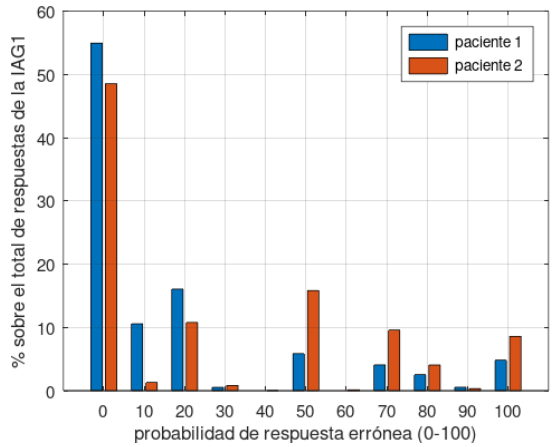


Fig. 7. Comparación de probabilidades de error entre dos pacientes virtuales.

VI. Trabajo Futuro

Las próximas etapas del proyecto contemplan, en primer lugar, analizar la influencia del parámetro temperatura de la IAG1 en la probabilidad de error de sus respuestas y; en segundo lugar, comprobar la efectividad del sistema de control propuesto. A continuación, detallamos ambos aspectos.

A. Influencia del parámetro temperatura

El esquema de control que proponemos se basa en reducir la temperatura en la IAG1 para regenerar sus respuestas cuando la IAG2 considera que la probabilidad de error supera un cierto umbral. Teóricamente, la reducción de la temperatura conllevará unas respuestas más predecibles y con menor probabilidad de alucinaciones, pero es necesaria una comprobación experimental. Para ello, partiremos de los resultados de este trabajo y regeneraremos todas las respuestas cuya probabilidad de error estimada ha sido igual o superior al 30%; utilizando para ello temperaturas de 0.5, 0.3 y 0.1. Los resultados nos permitirán cuantificar el efecto de la temperatura y con ello diseñar el sistema de control. También consideraremos el efecto del parámetro *top_p*, en conjunción o independientemente del parámetro temperatura.

B. Medida de efectividad del sistema de control

Con el sistema de control dimensionado correctamente, pondremos de nuevo los pacientes virtuales a disposición de los estudiantes, registraremos todos los resultados y analizaremos el funcionamiento, cuantificando el porcentaje de interacciones estudiante/IAG en las que el sistema de control actúa reduciendo la temperatura, los efectos sobre las probabilidades de error y los retrasos que supone la introducción del sistema de control. También se recogerán las valoraciones de la plataforma para evaluar si la introducción del sistema de control mejora la experiencia o si presenta inconvenientes, como un enlentecimiento de las interacciones.

VII. Conclusiones

Un modelo de IAG configurado para un funcionamiento natural y creativo (valores de temperatura o *top-p* elevados), está expuesto a respuestas poco fiables. Un control en bucle cerrado a través de una IAG secundaria podría garantizar la fiabilidad manteniendo la naturalidad de las conversaciones.

Nuestros resultados preliminares muestran la viabilidad de la supervisión mediante una IAG alternativa, aunque se requieren pruebas más exhaustivas.

Agradecimientos

Trabajo financiado por el Ministerio de Igualdad, proyecto "Investigación sobre sesgos de género en la generación de imágenes por Inteligencia Artificial" (LENA), exp. 19-3-ID24, subvención investigaciones feministas Instituto de las Mujeres 2024; y por el Ministerio de Ciencia, Innovación y Universidades, proyecto "Pacientes virtuales basados en Inteligencia Artificial y con características específicas de género para la formación en medicina y enfermería sobre entornos gamificados", exp. PID2024-157772OB-I00.

Referencias

[1] N. Posel, J.B. Mcgee, D.M. Fleiszer. "Twelve tips to support the development of clinical reasoning skills using virtual patient cases", *Medical Teacher*, 37, 813–818, 2014.

[2] M. Milne-Ives, C. De Cock, E. Lim, M.H. Shehadeh, N. De Pennington, G. Mole, E. Normando, E. Meinert. "The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review", *JMIR*, 22, e20346, 2020.

[3] D.A. Cook, M.M. Triola. "Virtual patients: a critical literature review and proposed next steps", *Medical Education*, 43, 303–311, 2009.

[4] R. Manesh, G. Dhaliwal. "Digital Tools to Enhance Clinical Reasoning", *Medical Clinics of North America*, 102, 559–565, 2018.

[5] A.A. Kononowicz, L.A. Woodham, S. Edelbring, N. Stathakarou, D. Davies, N. Saxena, L.T. Car, J. Carlstedt-Duke, J. Car, N. Zary. "Virtual Patient Simulations in Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration", *JMIR*, 21, e14676, 2019.

[6] F. Dolianiti, I. Tsoupouroglou, P. Antoniou, S. Konstantinidis, S. Anastasiades, P. Bamidis. "Chatbots in Healthcare Curricula: The Case of a Conversational Virtual Patient", *Lecture Notes in Computer Science*, 137–147, 2020.

[7] G. Eysenbach. "The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers", *JMIR Medical Education*, 9(1), e46885, 2023.

[8] C. Fernández, M.A. Vicente, I. Carrillo, M. Guilabert. "Análisis de especificidades de género en pacientes virtuales generados con las inteligencias artificiales ChatGPT y Mistral", *Los derechos de las mujeres en la sociedad digital*. Dykinson. ISBN: 9788410704480, 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Línea de investigación en "Análisis del sesgo de género en las herramientas de IA Generativa". Grupo Lovelace-Lena.

M. A. Vicente y C. Fernández
Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche, Universidad Miguel Hernández
suni@umh.es; c.fernandez@umh.es.

Presentamos la línea de investigación en el estudio del sesgo de género en las herramientas de inteligencia artificial compuesto por investigadores de diversas disciplinas y trayectorias, procedentes de tres destacadas instituciones académicas y de investigación: la Universidad Miguel Hernández (UMH), la Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO) y la Fundación San Juan de Dios (FSJD). Nuestro objetivo principal es abordar los sesgos de género que aparecen en los sistemas de inteligencia artificial generativa. Más allá de revelar la presencia de estos sesgos, queremos también crear conciencia en la comunidad tecnológica sobre la importancia de desarrollar tecnologías inclusivas y éticas.

***Palabras Clave*- inteligencia artificial generativa, sesgo de género, ética de la IA.**

I. Introducción

El grupo interdisciplinar Lovelace-Lena está formado por un grupo de investigadores que combinan perfiles de ingeniería, psicología, ciencias de la salud y estudios de género [1,2], procedentes de tres centros de investigación españoles (siendo la UMH el nodo coordinador), que colaboran en la detección y mitigación del sesgo de género en la Inteligencia Artificial Generativa (IAG).

II. Línea de Investigación y Proyectos

Nuestros objetivos son: (i) caracterizar los sesgos de género en modelos de texto e imagen de uso popular, (ii) diseñar protocolos objetivos y reproducibles para su evaluación, promoviendo un uso inclusivo de la IAG, y (iii) desarrollar tecnologías inclusivas y éticas.

En este contexto, el grupo ha trabajado en dos proyectos de investigación subvencionados por el Instituto de las Mujeres del Ministerio de Igualdad en sus convocatorias de 2023 y 2024. El primer proyecto titulado "*LOVELACE - Desarrollo de Métodos para Evaluar el Sesgo de Género en Chatbots de Inteligencia Artificial- Ref: 30-2-ID23."* se desarrolló entre octubre de 2023 y diciembre de 2024. Sus hallazgos clave son los siguientes: ChatGPT produce descripciones estereotipadas si no recibe instrucciones de lenguaje inclusivo; los *prompts* explícitos reducen el sesgo, aunque plataformas como Grok y Mistral muestran mayor flexibilidad que los filtros de ChatGPT; y los "pacientes virtuales" generados por IA omiten diferencias clínicas de género, lo que puede afectar a la formación sanitaria [3,4]. Además, se creó un producto: el *"Adaptador de Lenguaje Inclusivo LOVELACE*", basado en un modelo GPT, para facilitar a la comunidad universitaria la adaptación de textos siguiendo la Guía de Recomendaciones de la UMH para un Lenguaje No Sexista e Inclusivo [5].

Actualmente desarrollamos el proyecto *"LENA -Investigación sobre sesgos de género en la generación de imágenes por Inteligencia Artificial - Ref: 19-3-ID24"*, dedicado a evaluar el sesgo de género en plataformas de generación de imágenes. Examinamos tanto modelos de propósito general —DALL·E 3, Midjourney, Adobe Firefly— como generadores especializados —Comics Maker AI (cómic), Scenario (assets para videojuegos) y HeyGen (vídeo con avatares). Nuestra metodología combina baterías de prompts balanceadas con métricas automáticas de equidad y una revisión experta exhaustiva, lo que nos permite cuantificar y comparar la aparición de estereotipos visuales entre plataformas.

Referencias

[1] Disponible en la web: https://lovelace.umh.es/presentacion/ - Último acceso 27 mayo 2025.
[2] Disponible en la web: https://lena.umh.es/presentacion/ - Último acceso 27 mayo 2025.
[3] Carrillo, I., Fernández, C., Vicente, M. A., Guilabert, M., Sánchez, A., Gil, E., Arroyo, A., Calderón, M., Carratalá, M. C., López, A., Coves, A., Chilet, E., Valero, S., & Senabre, C. (2024). Detecting and Reducing Gender Bias in Spanish Texts Generated with ChatGPT and Mistral Chatbots: The Lovelace Project. Proceedings of The Global Conference on Women's Studies, 3(1), 29–42.
[4] Fernández Peris, C., Vicente Ripoll, M. A., Carrillo Murcia, I., & Guilabert Mora, M. (2024, julio). Análisis de especificidades de género en pacientes virtuales basados en Inteligencia Artificial. Ponencia presentada en el IV Congreso Internacional Feminismo Digital, Granada, España.
[5] Guía de Recomendaciones para un Lenguaje No Sexista e Inclusivo. Universidad Miguel Hernández. ISBN 978-84-09-50269-1 Versión abril 2023.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Orchestrated Service Provisioning in the GÉANT Global Platform for Labs

Jorge Sasiain
Department of Communications Engineering
University of the Basque Country (UPV/EHU)
48013 Bilbao, Spain
jorge.sasiain@ehu.eus

Frederic Loui
RARE/freeRtr Technical Leader
RENATER/GEANT
Paris, France
frederic.loui@renater.fr

In the context of the pan-European research and education network GÉANT, the Global Platform for Labs (GP4L) is a geographically distributed programmable network infrastructure dedicated to provide an ecosystem for experimentation in advanced network technologies. In this paper, we present ongoing work regarding the implementation of orchestration and automation for service provisioning in GP4L. Our work aims to improve network management operations in GP4L concerning the provision of the network configuration to support the lifecycle of experiments. In addition, we strive to create reusable code based on open source platforms that can be adopted by NRENs.

***Palabras Clave*—Network orchestration, Network automation, Network management, Service provisioning, NREN**

I. INTRODUCTION

The pan-Eruopean GÉANT network interconnects National Research and Education Networks (NRENs) across Europe, including a total of 38 NREN partners. Through a high performance backbone network and a portfolio of advanced services, GÉANT has been supporting the needs of Europe's Research & Education (R&E) community for nearly the last 30 years. In this context, the GÉANT Global Platform for Labs (GP4L)[1] is a large-scale experimental network infrastructure that has been continuously growing since the end of the last decade.

GP4L is dedicated to advancing network transformation, enabling experimentation, innovation, and collaboration in advanced network technologies. The GP4L experimental facility (Figure 1) is composed of over 40 networking and computing devices distributed primarily across Europe. This infrastructure includes data plane programming network equipment with ports of up to 400GE based on technologies like P4 and DPDK, and Data Transfer Nodes (DTNs). Thanks to this ecosystem, GP4L enables its users, ranging from individual researchers, to R&E organizations, to NRENs, and to network engineers, to run cutting-edge experiments in network programmability and new networking practices.

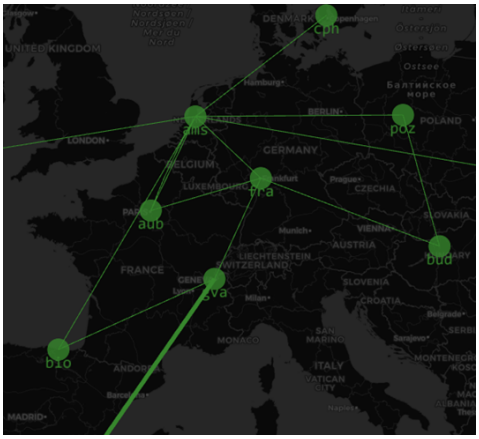


Fig. 1. GP4L Topology Map (Europe Only)

The GP4L community is actively working on the development of solutions and use cases aimed at improving the operational efficiency of network processes and operations through the implementation of automation and orchestration. These solutions are developed following the principles of flexibility, modularity and re-usability, in order to arrive at a blueprint ready to be adopted by NRENs and other organizations. In this paper, we present one of these in-progress developments, which targets achieving a generalized, reusable approach to orchestrating service provisioning, employing open-source automation components and TM Forum (TMF) standardized APIs.

II. BACKGROUND AND OBJECTIVES

Over the last years, GP4L has constituted a playground for validating innovative experiments in the fields of programmable networks (e.g. source routing [1], multicast forwarding [2], IPv6 packet processing [3], 5G User Plane Function [4]), leveraging its RARE (Router for Academia, Research & Education) data plane and freeRtr (freeRouter) control plane stack. Specifically, RARE/freeRtr[2]

[1]https://geant-netdev.gitlab-pages.pcss.pl/gp4ldocs/

[2]https://docs.rare.geant.org/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

is an open-source multifunctional router operating system, also developed under the framework of the GÉANT project, compatible with multiple dataplane targets such as hardware P4 switches and DPDK FPGAs. An efficient execution of on-demand experiments across this GP4L network substrate calls for the implementation of network orchestration, automation, and monitoring solutions that can configure the resources needed for the experiments, minimizing manual intervention.

Network automation encompasses methodologies and software tools aimed at automating the deployment, configuration, and monitoring of network devices and network services, replacing repetitive, cost-inefficient, and potentially error-prone manual operations with an Infrastructure as Code (IaC) approach [5]. Key components of the network automation stack include tools to inject configuration into network devices abstracted from hardware-specific details, tools for modeling the logic of business processes as workflows composed of tasks that represent network operations, and repositories to store up-to-date network information, known as Network Source of Truth (NSoT). With the consolidation of SDN/NFV frameworks, as well as the emergence of new paradigms such as Zero touch network & Service Management (ZSM), research efforts in network orchestration and automation have spiked in recent years in the context of next-generation networks such as 5G/6G [6] [7]. In our work, we target the application of these technologies to NRENs and R&E networks, where recent efforts such as [8] by the polish NREN PIONIER highlight how the implementation of flexible network automation and orchestration can significantly enhance the operation of large-scale research networks.

The orchestrated service provisioning work described in this paper complements recent successful initiatives carried out by GP4L in this area, such as the development of automated reservation management, and the validation of the Network Management as a Service (NMaaS)[3] platform. Like these initiatives, our ongoing work targets not only the benefit of GP4L, but also the fabrication of reusable solutions that can be adopted by NRENs and organizations with similar needs in their networks. We believe that our implementation in such a diverse and geographically dispersed environment as GP4L, jointly with the use of flexible data models based on open TFM standards, constitutes a unique contribution to advance the state of the art of network management and automation. The objectives pursued through the ongoing orchestrated service provisioning solution described in this paper are twofold:

- To design and develop workflows that automate the management of resources and the configuration of experiments over the infrastructure of the GP4L experimental testbed.
- To contribute towards providing modular and reusable solutions for network automation, exploring the use of different open source technologies and standardized APIs, and gathering lessons learned that can be documented and transferred to NRENs and organizations.

[3]https://nmaas.eu

III. TOWARDS ORCHESTRATED SERVICE PROVISIONING IN GP4L

In this section, the work carried out is described, first outlining the approach followed and the technologies selected to develop the service provisioning solution in GP4L. Later, a specific implementation consisting on the provisioning of layer-2 circuits between two GP4L devices is described and validated.

A. Orchestrated Service Provisioning Solution Design

The key components of the service provisioning solution proposed in this paper are as follows:

- A NSoT to act as the authoritative source of truth, representing the desired state of the network, including its topology, the configuration of the network devices, and the services, through a declarative approach. Maat[4] has been selected as the NSoT of choice for its flexibility and extensibility, featuring a JSON-based data model with cusutomizable schemas compliant with TMF APIs for resource and service management.
- Ansible[5] as the automation engine that interacts directly with the GP4L devices, retrieving and injecting configuration. We inject Ansible configuration via Jinja[6] templates for their flexibility in parametrizing the configuration. Furthermore, we employ Lightweight Service Orchestrator (LSO)[7], which acts as a proxy that can wait for a given Ansible playbook to complete, ensuring the smooth execution of service workflows.
- A workflow engine to allow the decomposition of the network services into several steps, automating their execution in an structured manner. For this purpose, we employ Apache Airflow[8], which enables the definition of processes in the form of Directed Acyclic Graphs (DAGs), collections of organized tasks each of which represents a step of the given process or workflow.
- A virtual network laboratory platform to deploy a digital twin of the GP4L network to test the workflows against before deploying them in the production network. This is carried out using Containerlab[9], which can deploy the core GP4L topology using the open source RARE/freeRtr image to emulate GP4L P4 devices.

Using these tools, a workflow that represents the provision of a service is depicted in Figure 2 and adheres to the following guidelines. The overarching entity is an Airflow DAG split into several tasks. When a user requests a new service, they first populate a form with

[4]https://geant-netdev.gitlab-pages.pcss.pl/MaatDocs/
[5]https://docs.ansible.com/
[6]https://jinja.palletsprojects.com/en/stable/
[7]https://github.com/workfloworchestrator/lso
[8]https://airflow.apache.org/
[9]https://containerlab.dev/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

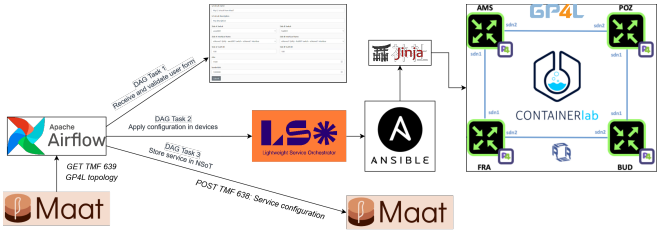


Fig. 2. Service Provisioning Workflow

the service parameters. This form is available using either the Airflow UI or, preferably, the Maat UI, which sends a POST request to the Airflow API with the parameters upon submission. The validation of these input parameters is the first DAG task, for which Airflow first retrieves resource data (devices and interfaces) from Maat.

After the input parameters are validated, the next DAG task is responsible for triggering the injection of the service configuration into the target devices. First, the DAG populates JSON objects for LSO with the configuration for each device. The main benefit of LSO is that it can wait until an Ansible playbook finalizes to return feedback to the DAG. LSO translates the JSON data into an Ansible inventory and variables, and launches the execution of an Ansible playbook, which in turn executes the tasks that belong to the role associated to the requested service against each device involved in the service. In these tasks, first Ansible renders the configuration from a Jinja template populated with the Ansible variables, and then applies the rendered configuration against an HTTP server listening on the device.

In the third and final DAG task, executed after LSO reports that the playbook completed successfully, Airflow updates the state in the NSoT, i.e. Maat, to reflect the newly configured service. Maat accepts schemas compliant with the TMF638 and TMF639 APIs for service and resource inventory, respectively. Hardware configuration of the devices and interfaces uses a TMF639 schema, while service-specific configuration uses a TMF638 schema. According to the data model of TMF638 and TMF639, resources and services are defined with a customizable set of *resourceCharacteristic* and *serviceCharactersitic*, and can be linked to each other using *resourceRelationship* and/or *serviceRelationship*, for example to stablish a relation between a device with one of its interfaces, or between an interface with a Service Attachment Point (SAP).

Each service provisioning workflow also contains its deprovisioning functionality counterpart. Provisioning and deprovisioning share the same Airflow DAG and the same Ansible role, but employ a different action in Ansible in order to trigger different tasks that make use of different Jinja templates and variables.

B. Implementation and Validation of Layer-2 Circuit Service

In this subsection we delve in more detail into the implementation of a specific service for the provision of layer 2 circuits, following the principles elaborated in the previous subsection.

Listing 1. LSO layer 2 circuit file

```
{
    "playbook_name": "configure_l2_circuit.yml",
    "callback": "http://localhost:8090/api/resume/",
    "inventory": {
        "rare": {
            "hosts": { /* Two hosts from Listing 2 */
            }
        }
    },
    "extra_vars": {
        "action": "",
        "l2circuit": {
            "name": "",
            "description": "",
            "virtual_circuit_id": "",
            "mtu": "",
            "bandwidth": ""
        }
    }
}
```

Listing 2. LSO layer 2 circuit host file

```
{
    "ansible_host": "",
    "sap": {
        "neighbor": "",
        "interface": "",
        "vrf": "",
        "template_name": "",
        "vlan_id": ""
    }
}
```

The user first submits the service request, denoting the device and interface endpoints for the circuit and additional parameters such as VLAN ID and bandwidth, for example using the Maat UI as shown in Figure 3. Airflow receives this input, validates it, and then builds the LSO files, populating the templates shown in Listing 1.

The Jinja template used by Ansible for the configuration of each of the two involved devices is shown in Listing 3. The Airflow handler listening in the LSO callback address notices when the playbook has completed and resumes the DAG, which finalizes by storing the service in Maat.

Listing 3. Jinja template for layer-2 circuit service

```
interface {{ SAP.INTERFACE }}.{{ SAP.VLAN_ID }}
 description {{ L2CIRCUIT.DESCRIPTION }}
 mtu {{ L2CIRCUIT.MTU }}
 bandwidth {{ L2CIRCUIT.BW }}
 vrf forwarding {{ SAP.VRF }}
 template {{ SAP.TEMPLATE_NAME }}
```

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

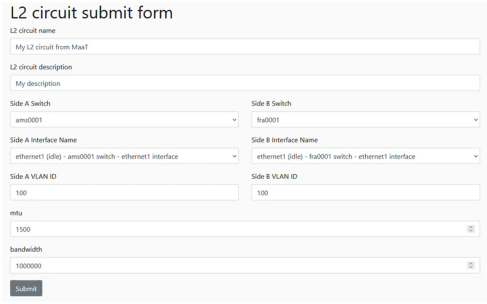


Fig. 3. Layer-2 Circuit Submission Form

```
 no shutdown
 exit
end
```

In Maat, services and resources are classified by their category field. This layer-2 circuit service involves a *l2.circuit* service and two *l2.circuit.sap* (SAP) services. Each SAP is linked to the top-level circuit service via a two-way *serviceRelationship*. It is also linked to the associated *interface* resource using a *serviceRelationship* in the interface and a mirror *resourceRelationship* in the SAP. An example of SAP service in the Maat UI is shown in Figure 4.

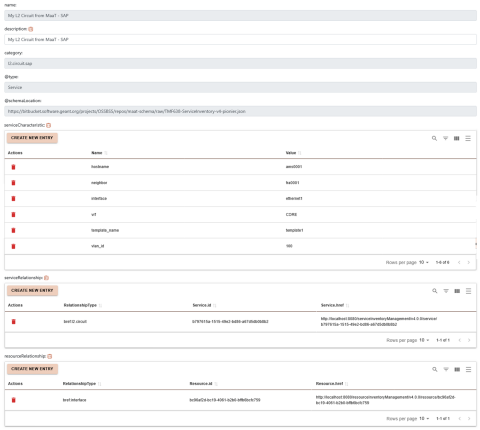


Fig. 4. SAP Service from a Layer-2 Circuit

Finally, the creation of the layer-2 circuit service is validated by ensuring that the configuration is in place in the GP4L devices and by executing a connectivity test between the two endpoints of the service.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented our work towards orchestrated service provisioning in GP4L. In the future, we plan to incorporate more services, such as Ethernet VPN and layer 3 services, as well as to support more types of devices beyond the RARE/freeRtr P4 devices used so far. To maximize re-usability, our ongoing work is aiming for generalized workflows that can be adapted to hardware from multiple vendors. The flexibility of the NSoT's data model makes it possible to abstract the internal device configuration into vendor-agnostic attributes towards the end user that submitted the service request.

Towards this goal, the approach we are following is to structure the logic of the automation solution in order to implement the functionality that is common to all DAGs (such as the APIs towards Maat and LSO) into a common library, separate from service-specific logic. Likewise, we are following a hierarchical approach with Ansible, so that, in addition to a shared playbook being responsible to include the corresponding role for the service, each role can execute a series of conditional subtasks according to the vendor of the devices and the service operation (deployment, deletion, modification, or verification). A clear separation between common logic, such as the rendering of Jinja templates, and vendor-specific configuration, contributes to facilitate the re-usability of the solution in different network environments.

In parallel, we also intend to evaluate other technologies emerging in the field of network automation, like the Kubernetes-based Kubenet and SDCIO. We are also beginning to evaluate the feasibility of using generative AI for the creation of workflows for new services.

ACKNOWLEDGEMENTS

This work was supported in part by the project GN5-2 HORIZON-INFRA-2024-GEANT-01-SGA in the Working Group 6 Task 2: Platform, and in part by the Spanish Ministry of Science and Innovation in the project EnablIng Native-AI Secure deterministic 6G networks for hyPer-connected envIRonmEnts (6G-INSPIRE) (PID2022-137329OB-C44). The authors also want to thank Marcin Adamski, Piotr Grabowski, and Tomasz Szewczyk from the Poznan Supercomputing and Networking Center (PSNC) for providing assistance and their reference implementation in the PIONIER network.

REFERENCES

[1] C. Dominicini, D. Mafioletti, A. C. Locateli, R. Villaca, M. Martinello, M. Ribeiro, and A. Gorodnik, "Polka: Polynomial key-based architecture for source routing in network fabrics," in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*. IEEE, 2020, pp. 326–334.

[2] S. Lindner, D. Merling, and M. Menth, "Learning multicast patterns for efficient bier forwarding with p4," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1238–1253, 2022.

[3] C. M. Moreira, E. Martelli, and T. Cass, "P4flow: A software-defined networking approach with programmable switches for accounting and forwarding ipv6 packets with user-defined flow label tags," in *EPJ Web of Conferences*, vol. 295. EDP Sciences, 2024, p. 07037.

[4] D. Franco, A. Atutxa, J. Astorga, E. Jacob, and F. Loui, "Towards an open-source and high-performance 5g upf on rare," in *JRES (Journées réseaux de l'enseignement et de la recherche) 2024*, 2024.

[5] M. Guerriero, M. Garriga, D. A. Tamburri, and F. Palomba, "Adoption, support, and challenges of infrastructure-as-code: Insights from industry," in *2019 IEEE International conference on software maintenance and evolution (ICSME)*. IEEE, 2019, pp. 580–589.

[6] E. Mulyana and G. Fakih, "Network automation with a single source of truth in a heterogeneous environment." *International Journal on Electrical Engineering & Informatics*, vol. 14, no. 1, 2022.

[7] M. El Rajab, L. Yang, and A. Shami, "Zero-touch networks: Towards next-generation network automation," *Computer Networks*, vol. 243, p. 110294, 2024.

[8] R. Lapacz, S. Filiposka, F. Loui, T. Szewczyk, M. Adamski, and K. Georgilakis, "A reusable approach to network service orchestration," in *JRES (Journées réseaux de l'enseignement et de la recherche) 2024*, 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Paving the Way for Digital Twins in Network Operations and Management

Daniel González-Sánchez*, Luis Bellido*, Ignacio D. Martinez-Casanueva*†, David Martínez-García*, David Fernández*, and Diego R. Lopez†
*Dpto. de Ingeniería de Sistemas Telemáticos, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, 28040, Spain
†Global CTIO, Telefónica Innovación Digital, Madrid, 28050, Spain
E-mails: daniel.gonzalez.sanchez@upm.es, luis.bellido@upm.es, ignacio.dominguezmartinez@telefonica.com, david.martinezgarci@upm.es, david.fernandez@upm.es, and diego.r.lopez@telefonica.com

***Abstract*—This work proposes a Network Digital Twin (NDT) solution based on a novel data integration approach to translate information from the network management domain into a formal, data-driven, and interoperable model, supporting improved control and management to facilitate the application of network automation systems.**

***Keywords*—Data integration, model-driven telemetry, network automation, Network Digital Twin (NDT), and network operations and management.**

SUMMARY

The Network Digital Twin (NDT) paradigm is a promising solution to efficiently analyze, diagnose, simulate, and control the behavior of real networks through data model-driven approaches within virtualized environments, enabling faster deployment of innovative applications without compromising the real-world network. However, challenges remain in mapping real network data to unified models and in ensuring efficient data collection, storage, and management processes.

To overcome these challenges, in [1] we presented a new methodology to model an NDT by translating data from model-driven telemetry sources into a knowledge-based data integration system that facilitates the management and operation of real networks. The proposed solution builds on the NGSI-LD standard specification defined within ETSI's Industry Specification Group on Context Information Management (ETSI ISG CIM), which enables the modeling of Digital Twins. The novelty of this work is threefold: i) an original NDT proposal focused on network operations and management purposes; ii) a novel data integration approach for mirroring the real network behavior; and iii) a new reference implementation for model-driven network management based on the NGSI-LD standard.

The work proposes a reference architecture and a prototype implementation of an NDT system for network operations and management. As shown in Figure 1, the architecture encompasses the four core elements considered to implement the logic of NDT solutions, providing: i) real-time and historical NDT data; ii) models for characterizing the NDT data; iii) interfaces to establish communication with the real network and with the NDT system; and iv) mapping mechanisms for the exchange of data between the real network and the NDT system. The prototype demonstrates how an NDT using the NGSI-LD standard can expose operational and configuration data from real networks in a formal and interoperable model, providing benefits to external network management systems (i.e., NMS) for network analytics and automation processes.

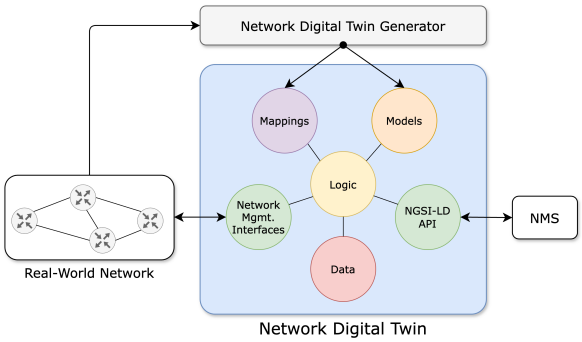


Fig. 1. NDT system for network operations and management.

This work was supported in part by the European Union's Horizon Europe Research and Innovation Program under Grant 101069732 (aerOS) and Grant 101097122 (ACROSS), and in part by the Ministry of Economic Affairs and Digital Transformation of the Spanish Government and the NextGenerationEU (Recovery, Transformation and Resilience Plan - PRTR) Projects of the UNICO-5G I+D Program under Grant TSI-063000-2021-79 (B5GEMINI-AIUC) and Grant TSI-063000-2021-81 (B5GEMINI-INFRA).

REFERENCES

[1] D. González-Sánchez, L. Bellido, I. D. Martinez-Casanueva, D. Martínez-García, D. Fernández and D. R. Lopez, "Toward Building a Digital Twin for Network Operations and Management," in IEEE Open Journal of the Communications Society, vol. 6, pp. 2583-2598, 2025, doi: 10.1109/OJCOMS.2025.3549873.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Análisis de Algoritmos Post-Cuánticos de Cifrado en Plataformas Heterogéneas

José Alberto Benito-Corral*, Jorge E. López de Vergara*, Luis de Pedro*, Iván González*, Florina Almenares†, Luis Cruz-Piris‡, Andrés Marín§

*Dept. Tecnología Electrónica y de las Comunicaciones, Universidad Autónoma de Madrid
†Dept. Ingeniería Telemática, Universidad Carlos III de Madrid
‡Dept. Automática, Universidad de Alcalá
§Dept. Ing. Sistemas Telemáticos, Universidad Politécnica de Madrid
josea.benito@estudiante.uam.es, {jorge.lopez_vergara, luis.depedro, ivan.gonzalez}@uam.es, florina@it.uc3m.es, luis.cruz@uah.es, andres.mlopez@upm.es

En este artículo se presenta un estudio exhaustivo del rendimiento de algoritmos criptográficos post-cuánticos estandarizados por el NIST (Kyber/ML-KEM, Dilithium/ML-DSA, FALCON/FN-DSA y SPHINCS+/SLH-DSA), comparándolos con RSA y ECDSA. Se evaluaron tiempos de generación de claves, encapsulación/desencapsulación, firma/verificación y consumo de CPU/memoria en las plataformas de: Ordenador PC, Raspberry Pi 5 y teléfono móvil Android. Se utilizaron distintos lenguajes de programación en las dos primeras plataformas (C, Java, C#, Python) y Kotlin en la tercera. Los resultados indican que las implementaciones en C optimizado con extensiones vectoriales ofrecen el mejor rendimiento, destacando a Kyber y Dilithium como opciones prácticas en sistemas reales, con resultados comparables a los que se obtienen en algoritmos de cifrado clásicos.

***Palabras Clave*—Criptografía post-cuántica, Kyber, Dilithium, FALCON, SPHINCS+, RSA, ECDSA, rendimiento, recursos, plataformas, lenguajes de programación.**

I. Introducción

La inminente aparición de ordenadores cuánticos, capaces de ejecutar el algoritmo de Shor en tiempo polinómico, es decir, con una eficiencia mucho mayor que los actuales algoritmos clásicos, pone en jaque la criptografía clásica basada en la factorización (RSA) y el logaritmo discreto (curvas elípticas, ECDSA) [1]. Para garantizar la confidencialidad y autenticidad de la información a largo plazo, la comunidad científica y los organismos de estandarización (NIST), están promoviendo esquemas de criptografía post–cuántica (PQC, *Post-Quantum Cryptography*) basados en problemas matemáticos intratables hoy en día tanto para procesadores clásicos como cuánticos.

Aunque en agosto de 2024 el NIST (*National Institute of Standards and Technology*) publicó varios estándares PQC [2], su integración en sistemas reales afronta desafíos prácticos: incompatibilidades con infraestructuras heredadas, limitaciones de hardware en dispositivos IoT y la necesidad de soporte multi-lenguaje en entornos empresariales heterogéneos. Esta brecha entre especificación y adopción motiva un análisis detallado de rendimiento y consumo de recursos, para que las organizaciones puedan planificar migraciones graduales sin comprometer su actividad durante el periodo de transición.

Por ello, este artículo presenta una evaluación práctica de los algoritmos estandarizados y a estandarizar por el NIST —ML-KEM (Kyber), ML-DSA (Dilithium), FN-DSA (FALCON) y SLH-DSA (SPHINCS+)— comparándolos con RSA-2048 y ECDSA-256 en tres plataformas heterogéneas: PC (Intel x86), Raspberry Pi 5 (ARM Cortex A76) y teléfono móvil Android. Se utiliza C, Java, C#, Python en las dos primeras plataformas y Kotlin en la tercera. Con el lenguaje C se ha probado a aplicar optimizaciones de compilación que aprovechen el hardware disponible. A partir de esta combinación de algoritmos, plataformas y lenguajes, se han medido los tiempos de generación de claves, encapsulación/decapsulación, firma/verificación y consumo de CPU y memoria.

Las contribuciones de este trabajo son:

- Comparar el rendimiento de esquemas post–cuánticos frente a RSA y ECDSA en operaciones criptográficas.
- Analizar el uso de recursos (CPU y memoria) de estos nuevos algoritmos criptográficos en diferentes arquitecturas y lenguajes de programación.
- Ofrecer recomendaciones prácticas para la adopción de criptografía resistente a ataques cuánticos en entornos reales.

El resto del artículo se estructura de la siguiente manera: en primer lugar, se revisa el estado del arte y los trabajos relacionados; después, se describe la metodología empleada y el entorno de pruebas; a continuación, se presentan

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

los resultados obtenidos; y, por último, se exponen las conclusiones junto con las principales líneas de trabajo futuro.

II. Estado del Arte

En esta sección, se muestran estudios y trabajos relacionados. La criptografía post-cuántica ha evolucionado rápidamente en la última década para responder a la amenaza que suponen los ordenadores cuánticos a los algoritmos clásicos. El NIST inició en 2016 un proceso de estandarización que, tras la selección de cuatro esquemas principales, tres basados en retículas (Kyber para intercambio de claves, Dilithium y FALCON para firmas), y otro basado en *hashes* sin estado (SPHINCS+), culminó en 2024 con la estandarización de ML-KEM (Kyber) [3], ML-DSA (Dilithium) [4], SLH-DSA (SPHINCS+) [5] y la preparación del borrador para FN-DSA (FALCON).

Varios estudios han evaluado el rendimiento de estos algoritmos en diferentes entornos. Benny *et al*. [6] realizan un meta-análisis sobre los riesgos de la computación cuántica y la llegada de los algoritmos post-cuánticos Kyber, Dilithium, FALCON y SPHINCS+, donde examinan la seguridad, características y rendimiento de estos algoritmos y las relaciones entre tamaño de clave pública y privada así como el tiempo de cifrado. De manera similar, Iglesias Hernández *et al*. [7] realizan un análisis comparativo entre Dilithium y FALCON, donde evalúan la seguridad, eficiencia y rendimiento, comparando el tamaño de las claves y midiendo el tiempo de ejecución en la generación de claves, firma y verificación para las distintas versiones de estos algoritmos.

En ámbitos más específicos, Vidaković y Miličević [8] analizan el desempeño y la aplicabilidad de Dilithium, FALCON y SPHINCS+ en entornos con recursos limitados, como IoT, tarjetas inteligentes, blockchain y comunicación entre vehículos (V2V). Se evalúa el equilibrio entre seguridad (tamaño de clave pública y firma), eficiencia computacional y consumo de recursos, destacando las ventajas y limitaciones de cada algoritmo según su utilización. Fitzgibbon y Octtaviani [9] presentan un estudio de rendimiento de algoritmos KEM: Kyber, HQC, McElliece y BIKE, y de firma digital: Dilithium, FALCON y SPHINCS+, para los niveles de seguridad 3 y 5, en distintos dispositivos: un PC, un portátil y una Raspberry Pi 4. Se usa la librería `liboqs` y demuestra de nuevo, que los algoritmos basados en retículas son los más eficientes. Abassi *et al*. [10] analizan Kyber, NTRU, BIKE, Dilithium y FALCON, en tres entornos distintos: servidor, portátil y un dispositivo equivalente a una Raspberry Pi 3B+, utilizando también `liboqs` e implementaciones originales de los autores. Los resultados confirman que los algoritmos muestran un rendimiento similar, siendo BIKE el de mayor consumo.

Este artículo presenta similitudes con los trabajos previos para analizar los algoritmos post-cuánticos, pero se distingue de estos por su análisis simultáneo a lo largo de tres ejes: algoritmos, arquitecturas y lenguajes. Ofrece así una visión más completa, evaluando PC, SBC (*Single Board Computer*) y teléfono móvil con C, Java, C#, Python y Kotlin, proporcionando tiempos de operación y uso de CPU y memoria de las distintas variantes de los algoritmos en todas estas plataformas y lenguajes.

III. Metodología

En esta sección se detalla el comportamiento de los esquemas post-cuánticos en entornos reales. Para ello se diseñó un banco de pruebas basado en la ejecución repetida de operaciones criptográficas y la monitorización de consumo de recursos. A continuación se describen los principales elementos de la metodología empleada.

Sistemas de prueba. Se emplearon tres plataformas:

- **PC:** Procesador AMD Ryzen 5 5500U (6 núcleos / 12 hilos, hasta 4.0 GHz), 16 GB RAM, Ubuntu 24.04.2 LTS, GCC 13.3.0, .NET SDK 8.0.115, Java 21.0.2 y Python 3.12.3.
- **SBC (Raspberry Pi 5):** Procesador ARM Cortex-A76 con 4 núcleos y 64 bits, 2.4 GHz, 8 GB RAM, Debian GNU/Linux 12, GCC 12.2.0, .NET SDK 8.0.405, Java 17.0.13 y Python 3.11.2.
- **Móvil Android:** Procesador MediaTek Dimensity 700 (ARM Cortex-A76 + A55) con 8 núcleos, 2.2 GHz, 8 GB RAM, Android 13, Kotlin 1.9.24.

Bibliotecas criptográficas. Se seleccionaron implementaciones oficiales o de referencia y se emplearon las siguientes bibliotecas:

- *C:* `OpenSSL` (RSA, ECDSA y Kyber (NEON en Raspberry PI 5) y código de referencia en C para Kyber, Dilithium, SPHINCS+ (pq-crystals) y FALCON (benchmarks NIST-PQC Round 3).
- *Java:* `bouncycastle.pqc.crypto` (SPHINCS+, Dilithium, FALCON), `bouncycastle.pqc.jcajce` (Kyber) y `bouncycastle.jce` (RSA, ECDSA).
- *Python:* `cryptography.hazmat` (RSA, ECDSA)
- *C#:* `bouncycastle.pqc.crypto` (PQC), `System.Security.Cryptography` (RSA, ECDSA).
- *Kotlin:* `bouncycastle.pqc.jcajce` para todos los esquemas post-cuánticos.

Adaptación del código. A partir de las versiones anteriores, se realizó una instrumentación en cada lenguaje para medir de forma precisa los tiempos que supone la generación de claves, la encapsulación/desencapsulación (intercambio de claves) o la firma/verificación (firmas digitales). El código fuente, disponible en GitHub[1], incluye las modificaciones necesarias para obtener estas métricas.

Obtención de uso de recursos. El consumo de CPU y memoria se registró con:

- `/usr/bin/time -v` en PC y SBC, obteniendo tiempo de ejecución, uso máximo de memoria residente (RSS) y porcentaje de CPU.
- Android Studio Profiler en el móvil, midiendo memoria y CPU tras activar depuración USB y perfilado.

Procedimiento. Para cada operación criptográfica:

[1] https://github.com/Alberto1-2/TFG_CODE

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

1. Se ejecutaron 1000 iteraciones (100 en móvil). Se calcularon valores promedio y desviación típica de tiempos (ms), CPU (%) y memoria (MB).

Este diseño asegura comparaciones homogéneas y minimiza ruido de medición, proporcionando datos fiables sobre el coste real para la integración algoritmos PQC.

IV. Resultados

En esta sección se presentan los resultados de las pruebas que se muestran en la Tabla I con los tiempos medios de operación (generación + encapsulación/desencapsulación para intercambio de claves, y generación + firma + verificación para firmas) para PC y Raspberry: con C (con optimizaciones AVX2 y NEON y FPU), C#, Java, Python; y en Android: Kotlin.

Para Kyber (512, 768 y 1024 bits), Python ofrece los peores tiempos: 7.9 ms a 17.7 ms en PC y de 11.1 ms a 25 ms en Raspberry Pi 5, siendo hasta 300 veces más lento respecto a C con AVX2. En cambio, Java y C# ofrecen resultados mucho más competitivos, con valores del submilisegundo en todas las variantes, alcanzando entre 0.18 y 0.4 ms en PC y 0.3 a 0.65 ms en Raspberry. La Fig. 1 muestra la comparativa de tiempos para Kyber 512.

En Dilithium (niveles 2, 3 y 5), de forma similar, Python presenta tiempos elevados, (≈42–73 ms) en PC y hasta 137 ms en Raspberry, mientras que Java y C# se sitúan entre 0.7 y 1.5 ms y entre 1.3 y 3.1 ms en ambos entornos, con degradaciones que oscilan entre 5 y hasta más de 300 veces (python) frente a C (PC). Kotlin ofrece cifras similares, con valores (≈6–14 ms). La Fig. 2 muestra la comparativa de tiempos para Dilithium 2.

Para FALCON 512 y 1024, Python resulta impráctico en PC (más de 10 s por iteración en FALCON 1024) y en Raspberry (1.7-20 s), mientras que Java y C# reducen esos tiempos a 11–25 ms (512 bits) y 28–68 ms (1024 bits). La Fig. 3 muestra la comparativa de tiempos para FALCON 512.

SPHINCS+ en su esquema de firma SHA2-128f (*fast robust*) alcanza en PC 7.9 ms (C), hasta 117 ms (C#), 107 ms (Java) y 67 ms (Python). En Raspberry Pi 5, llega a 40 ms (C), mientras que en C#, Java y Python alcanza 205 ms, 145 ms y 83 ms. En Android, Kotlin obtiene 391 ms, descartándolo para aplicaciones interactivas. La Fig. 4 muestra la comparativa de tiempos para SPHINCS+ SHA2-128f.

El uso de memoria en C es bastante bajo en relación con el resto de lenguajes, en un orden de magnitud similar entre PC y Raspberry, de unidades; en Python del orden de decenas, y en Java, C# y Kotlin del orden de centenas de MBytes. En uso CPU, Java y C# se superan el 100 % de uso, lo que indica que se aprovecha la existencia de varios cores en el procesador. C y Python muestran un uso cercano al 100 %. Mientras, en Kotlin se observan usos de CPU que alcanzan hasta el 200 %.

En algoritmos clásicos, RSA-2048 es mucho más lento, debido a la generación de clave: sobre 181 ms (C) en PC, y similares en C# y Java (183-143 ms), y 43 ms en Python. En Raspberry Pi 5, supera los 500 ms en todos los lenguajes. Kotlin en Android requiere 325 ms. ECDSA-256, en cambio, se mantiene por debajo de 1.2 ms en todas las plataformas y lenguajes, destacando por su eficiencia.

Estas cifras muestran que Kyber y Dilithium ofrecen un compromiso óptimo entre latencia y recursos, manteniéndose en rangos de ms incluso en hardware limitado como Raspberry Pi o Android, mientras que SPHINCS+ y RSA quedan fuera de escenarios de alto rendimiento.

V. Conclusiones

En este trabajo se ha presentado una evaluación exhaustiva de los esquemas de criptografía post-cuántica propuestos para estandarizar en el NIST (Kyber, Dilithium, FALCON y SPHINCS+) frente a los algoritmos clásicos RSA-2048 y ECDSA-256, en tres plataformas heterogéneas (PC, Raspberry Pi 5 y Android) con implementaciones con distintos lenguajes. Los resultados revelan que:

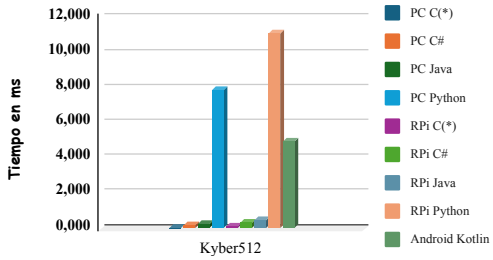


Fig. 1. Tiempos de ejecución para Kyber 512

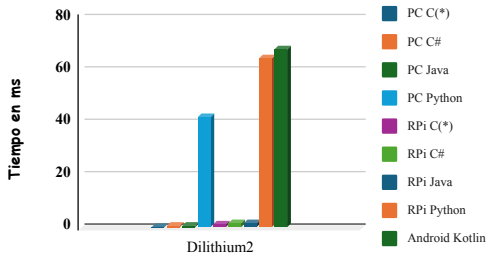


Fig. 2. Tiempos de ejecución para Dilithium 2

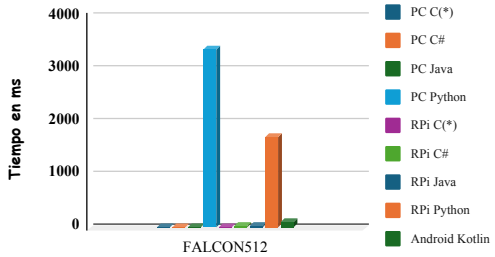


Fig. 3. Tiempos de ejecución para FALCON 512

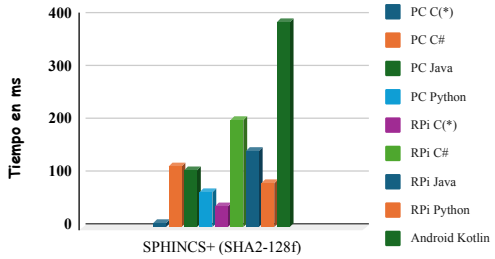


Fig. 4. Tiempos de ejecución para SPHINCS+ SHA2-128f

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
TIEMPO MEDIO TOTAL DE EJECUCIÓN Y DESVIACIÓN TÍPICA (MS) EN PC, RASPBERRY PI 5 Y ANDROID
C(*) INDICA OPTIMIZACIONES AVX2 EN C PARA PC, Y NEON Y FPU PARA RPI 5 EN C

Esquema	**PC**: C(*)	C#	Java	Python	**RPi 5**: C(*)	C#	Java	Python	**Android**: Kotlin
Intercambio de claves									
Kyber 512	0.027±0.003	0.185±0.028	0.265±0.128	7.900±0.071	0.073±0.002	0.290±0.179	0.447±0.332	11.111±0.180	4.909±7.204
Kyber 768	0.042±0.004	0.241±0.030	0.287±0.095	12.419±0.084	0.116±0.004	0.464±0.111	0.651±0.502	18.404±1.130	5.701±2.858
Kyber 1024	0.058±0.005	0.360±0.033	0.408±0.049	17.756±0.072	0.212±0.045	0.655±0.101	0.638±0.190	24.965±0.241	7.093±2.971
Firma digital									
Dilithium 2	0.124±0.039	0.678±0.312	0.776±0.268	42.254±20.532	0.810±0.359	1.329±0.631	1.345±0.542	64.722±29.253	6.810±5.265
Dilithium 3	0.203±0.062	1.186±0.559	1.244±0.481	68.721±35.142	1.316±0.631	2.336±1.135	2.160±0.913	112.230±59.474	10.624±8.095
Dilithium 5	0.283±0.066	1.543±0.554	1.555±0.475	88.771±37.510	1.649±0.555	3.128±1.185	2.806±0.919	136.899±54.810	14.086±9.285
SPHINCS+ (SHA2-128f)	7.914±0.149	116.600±0.555	106.922±8.019	66.621±5.194	40.182±0.057	204.739±10.919	145.073±9.440	83.469±0.095	391.044±25.067
FALCON 512	6.050±6.970	13.159±4.615	10.905±3.313	3.394±1.582 **seg**	10.953±2.918	25.790±9.378	22.767±6.263	1.709±1.113 **seg**	106.980±24.010
Algoritmos clásicos									
RSA-2048	181.449±96.890	183.340±100.645	146.266±93.880	43.265±25.652	648.012±329.217	646.377±271.578	538.567±347.396	241.216±111.766	325.385±143.890
ECDSA-256	0.105±0.012	0.338±0.036	0.675±0.261	0.145±0.005	0.342±0.072	1.107±0.344	1.188±0.577	0.373±0.025	0.642±0.049

- **Rendimiento.** Kyber 512 y Dilithium 2, optimizados con AVX2 en C, alcanzan latencias de intercambio de claves y firma/verificación inferiores a 0.3 ms en PC. En Raspberry Pi 5 con NEON, estos esquemas crecen entre 4 y 7 veces, y en Android se mantienen por debajo de 7 ms, superando claramente a RSA en velocidad de generación de claves y firmas. Además, con una desviación típica también mínima con valores por debajo de 5 milésimas y 7 décimas de milisegundo, respectivamente en C con AVX2, lo que proporciona una gran estabilidad de estos algoritmos en su ejecución en PC.
- **Consumo de recursos.** En Android, los esquemas post-cuánticos emplean un 140 %–200 % de CPU y 270–300 MB de memoria, exceso asumible en la mayoría de aplicaciones móviles. SPHINCS+ y FALCON presentan restricciones mayores (latencias de hasta cientos de ms) y, en el caso de SPHINCS+ es importante escoger un esquema ligero (SHA2-128f) para reducir los tiempos de proceso.
- **Portabilidad multi–lenguaje.** Las implementaciones en Java y C# ofrecen latencias submilisegundo para Kyber e inferior a los 2 ms para Dilithium, favoreciendo su integración en aplicaciones empresariales. Python resulta adecuado solo para prototipos por sus penalizaciones de hasta dos órdenes de magnitud.

De este modo, Kyber y Dilithium se perfilan como las opciones más equilibradas para una transición segura y eficiente hacia la era post-cuántica en entornos reales. A partir de los resultados obtenidos, como trabajo futuro se plantean las siguientes ideas:

- **Optimización de implementaciones:** mejorar el rendimiento de FALCON y SPHINCS+ mediante técnicas de bajo nivel y optimización de código para ampliar su accesibilidad en entornos de baja capacidad computacional.
- **Estudio en entornos reales a gran escala:** analizar el comportamiento de los algoritmos en sistemas con distinto número de iteraciones, alta concurrencia de usuarios, múltiples dispositivos y diversidad de arquitecturas, para evaluar su escalabilidad y robustez.
- **Evaluación de nuevos candidatos PQC:** aplicar la misma metodología a los esquemas de *NIST Round 4* (BIKE, Classic McEliece, HQC, o NTRU), con el fin de comparar sus características de rendimiento y eficiencia con los estándares actuales.
- **Esquemas híbridos clásico–post-cuántico:** diseñar y validar protocolos que combinen algoritmos tradicionales y post-cuánticos en un mismo flujo criptográfico, facilitando la compatibilidad y reforzando la seguridad durante la transición.

AGRADECIMIENTOS

Esta actuación ha sido financiada mediante el programa de actividades de I+D con referencia TEC-2024/COM-504 y acrónimo RAMONES-CM concedido por la Comunidad de Madrid a través de la Dirección General de Investigación e Innovación Tecnológica a través de la Orden 5696/2024.

REFERENCIAS

[1] D. J. Bernstein and T. Lange, "Post-quantum cryptography," *Nature*, vol. 549, pp. 188–194, 2017.
[2] National Institute of Standards and Technology, "Post-quantum cryptography standardization fips," https://csrc.nist.gov/pqc-standardization, NIST, Tech. Rep., 2024.
[3] ——, "FIPS 203 module-lattice-based key-encapsulation mechanism standard," https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.203.pdf, NIST, Federal Information Processing Standards Publication FIPS 203, 2024.
[4] ——, "FIPS 204 module-lattice-based digital signature standard," https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.204.pdf, NIST, Federal Information Processing Standards Publication FIPS 204, 2024.
[5] ——, "FIPS 205 stateless hash-based digital signature standard," https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.205.pdf, NIST, Federal Information Processing Standards Publication FIPS 205, 2024.
[6] S. Benny, I. Desai, L. Uriarte, I. Tsai, and L. McMahan, "A meta-analysis on NIST post-quantum cryptographic primitive finalists," *Emerging Investigators*, 2024.
[7] E. Iglesias Hernández, L. Hernández Álvarez, L. Hernández Encinas, and J. Sánchez García, "Análisis comparativo de las firmas digitales postcuánticas basadas en retículos," in *IX Jornadas Nacionales de Investigación en Ciberseguridad (JNIC)*, 2024, pp. 404–411.
[8] M. Vidaković and K. Miličević, "Performance and applicability of post-quantum digital signature algorithms in resource-constrained environments," *Algorithms*, vol. 16, no. 11, p. 518, 2023.
[9] G. Fitzgibbon and C. Ottaviani, "Constrained device performance benchmarking with the implementation of post-quantum cryptography," *Cryptography*, vol. 8, no. 2, 2024.
[10] M. Abbasi, F. Cardoso, P. Váz, J. Silva, and P. Martin, "A practical performance benchmark of post-quantum cryptography across heterogeneous computing environments," *Cryptography*, vol. 9, no. 2, 2025.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Diseño e implementación de una arquitectura abierta basada en la nube para planificar rutas personalizadas en función de la calidad del aire

E. Meneses-Albala*, S. Felici-Castell*, JJ. Perez-Solano*, P. Salva-Garcia†
* ETSE, Universitat de València, † University Western of Scotland (Reino Unido),
emeal2@uv.es, felici@uv.es, jjperezs@uv.es, pablo.salva-garcia@uws.ac.uk

La contaminación urbana representa un riesgo significativo para la salud respiratoria, lo que hace necesario el uso de sistemas inteligentes para la planificación de rutas. Este estudio presenta el diseño y la evaluación de una Arquitectura de Microservicios (MSA) escalable, abierta y nativa en la nube, orientada al desarrollo de un planificador de rutas saludables. Aprovechando tecnologías de código abierto, principalmente Kubernetes para la orquestación y Docker para la contenedorización, la arquitectura permite procesar grandes volúmenes de datos de estaciones de monitoreo de calidad del aire. Además, se emplean técnicas de interpolación espacial para generar mapas de contaminación de alta resolución, y se incorporan perfiles de salud personalizados mediante una Métrica Compleja Personalizada, con el fin de ofrecer rutas optimizadas de forma ad-hoc. Esta MSA y su planificador han sido validados en Valencia, España, logrando una reducción promedio del 18% en la exposición a la contaminación durante las pruebas realizadas, con un incremento del 10% en la distancia recorrida y con un tiempo medio de respuesta inferior a un segundo (997,96 ms).

***Palabras Clave*—calidad del aire, rutas, cloud, microservicios, open**

I. INTRODUCCIÓN

El diseño y uso de servicios urbanos inteligentes dependen de la gestión eficiente de grandes volúmenes de datos en tiempo real, provenientes de fuentes distribuidas y heterogéneas como sensores ambientales, sistemas de transporte e infraestructuras urbanas. Procesar estos datos de forma eficiente y adaptable a tareas complejas y personalizadas para múltiples usuarios plantea importantes desafíos arquitectónicos.

Tradicionalmente, aplicaciones complejas como estas podrían haberse abordado mediante arquitecturas monolíticas. Sin embargo, a medida que los requisitos de escalabilidad aumentan, los flujos de datos se vuelven más dinámicos y la necesidad de actualización y evolución rápida se intensifica, los monolitos revelan limitaciones significativas: se vuelven rígidos, difíciles de escalar selectivamente (obligando a escalar toda la aplicación incluso si solo un microservicio es el cuello de botella), poco resilientes (un fallo en una parte puede colapsar el sistema entero) y complejos de mantener y desplegar con agilidad. Además de que se ralentiza el proceso de incorporación de nuevas tecnologías en partes específicas del sistema.

En respuesta a estas limitaciones y buscando una mayor agilidad, resiliencia y, sobre todo, escalabilidad granular, la Arquitectura de Microservicios (*Micro Services Architecture*, MSA) ha ganado un auge sin precedentes en la última década como el paradigma preferido para construir determinados sistemas distribuidos modernos y preparados para la nube [1]. La MSA propone dividir una aplicación en un conjunto de servicios pequeños, autónomos, desplegables de forma independiente y diseñados para cumplir con una funcionalidad específica del sistema. Este enfoque, facilitado enormemente por la contenedorización (que empaqueta cada servicio con sus dependencias) con Docker y sistemas de orquestación como Kubernetes (que gestiona el despliegue, escalado y auto-reparación de los contenedores), confiere ventajas cruciales: flexibilidad, modularidad, mayor resiliencia (los fallos se aíslan) y la capacidad fundamental de escalar cada servicio horizontalmente e independientemente de los demás, optimizando el uso de recursos y garantizando el rendimiento bajo picos de carga en microservicios específicos.

Aprovechando estas capacidades ofrecidas por las MSAs y combinándolas con la infraestructura que proporcionan las tecnologías en la nube, este artículo presenta el diseño, implementación y evaluación de un Planificador de Rutas Saludables (o *Healthy Route Planner*, HRP) personalizado, como caso de uso frente a los desafíos previamente identificados. A diferencia de soluciones comerciales de enrutamiento ampliamente utilizadas, como Google Maps, que son sistemas cerrados (*closed-source*) y limitan la personalización profunda o la transparen-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

cia de su funcionamiento interno, nuestra propuesta se basa íntegramente en una MSA realizada con tecnologías abiertas (*open-source*). Este trabajo se centra en el diseño de dicha arquitectura, cuya implementación no solo permite una operación eficiente frente a datos dinámicos y múltiples usuarios, sino que, al estar basada en una arquitectura abierta, ofrece un alto grado de adaptabilidad, personalización y replicabilidad. Esta propuesta representa un avance sustancial respecto a desarrollos previos considerados como pruebas de concepto [2], [3], los cuales no contemplaban una preparación integral para entornos en la nube ni abordaban plenamente los requerimientos asociados al despliegue escalable y multiusuario.

El resto del trabajo se organiza de la siguiente manera. En la Sección II, se presenta una revisión del estado del arte en planificadores de rutas saludables y arquitecturas de microservicios para aplicaciones geoespaciales. La Sección III detalla el análisis de los requisitos del sistema, los desafíos tecnológicos abordados y las decisiones de diseño fundamentales, incluyendo el flujo de trabajo y la selección tecnológica. Seguidamente, la Sección IV describe en profundidad la MSA propuesta, especificando sus componentes de bases de datos, lógica de negocio y mecanismos de acceso externo. La Sección V presenta los resultados de la evaluación de la aplicación, incluyendo la calidad de las rutas generadas bajo diferentes perfiles y condiciones, así como el rendimiento computacional del sistema. Finalmente, la Sección VI resume las conclusiones principales de este estudio y las líneas de trabajo futuro.

II. ESTADO DEL ARTE

A. Calidad del aire

La calidad del aire (Air Quality, AQ) en nuestro HRP se basa en datos de estaciones de monitoreo (*Air Quality Monitoring Stations*, AQMS). La reciente Directiva Europea 2024/2881 [4] define contaminantes clave como NO_2, SO_2, CO, O_3, $PM2.5$ y $PM10$. Los datos para nuestra validación en Valencia provienen de redes reguladas [5], [6], con actualizaciones cada 10 minutos. La Tabla I resume las Guías de Calidad del Aire (*Air Quality Guidelines*, AQG) de la Organización Mundial de la Salud (OMS) [7] usadas para la normalización de los contaminantes.

Contaminante	Valor recomendado	Promedio
$PM2.5$	15 $\mu g/m^3$	24-horas
$PM10$	45 $\mu g/m^3$	24-horas
O_3	100 $\mu g/m^3$	8-horas
NO_2	25 $\mu g/m^3$	24-horas
SO_2	40 $\mu g/m^3$	24-horas
CO	4 mg/m^3	24-horas

Tabla I: Guías de Calidad del Aire (AQG) de la OMS con los valores máximos recomendados de concentración [7] y su respectivo tiempo de promediado.

B. Planificadores de rutas orientados a la AQ

Últimamente, la planificación de rutas urbanas que buscan reducir la exposición de las personas a contaminantes atmosféricos ha generado un interés considerable. Incluso aplicaciones tan extendidas y de uso diario como Google Maps han comenzado a mostrar datos de calidad del aire para ciertas ubicaciones [8]. Sin embargo, la información sobre la contaminación suele tener retrasos y no se integra directamente en el cálculo de rutas. Presentándose simplemente como un dato complementario, sin guiar activamente al usuario hacia trayectos menos contaminados. Paralelamente, la investigación académica ha explorado este campo. Enfoques iniciales [9], [10] se basaron en datos estáticos o históricos, limitando su aplicabilidad a las condiciones ambientales dinámicas y cambiantes del mundo real. Trabajos más recientes han incorporado datos en tiempo real, usando mediciones móviles [11], o explorando aplicaciones específicas como rutas para ejercicio [12], o considerando otros factores como el ruido [13], [14]; no obstante, la mayoría de las aplicaciones no incorporan personalización basada en el estado de salud del usuario al calcular las rutas. A diferencia de estas aproximaciones (incluyendo la funcionalidad actual de Google Maps), nuestra propuesta se distingue por ofrecer rutas personalizadas calculadas en tiempo real, integrando un set completo de contaminantes regulados por la OMS y adaptándose dinámicamente tanto a las condiciones ambientales, como al estado de salud específico de cada usuario.

C. Arquitecturas de microservicios para aplicaciones geoespaciales

Las MSAs han demostrado ser especialmente útiles en aplicaciones geoespaciales. Un ejemplo destacado es el modelo de 'servicios atómicos' propuesto en [15] para entornos de ciudades inteligentes, que promueve el uso de módulos especializados en funciones únicas, como análisis de datos, integración y visualización. Esta estructura modular no solo mejora el aislamiento de fallos, sino que también permite escalar los servicios de forma independiente [1]. Además, la contenedorización ayuda a garantizar entornos de despliegue consistentes y portables en infraestructuras distribuidas [16].

El procesamiento de información ambiental a gran escala, como la interpolación espacial o la actualización de modelos en tiempo real, implica una alta carga computacional. La computación distribuida permite dividir estas tareas entre múltiples nodos, lo que mejora el rendimiento general del sistema. Herramientas de orquestación como Kubernetes facilitan una distribución eficiente de la carga entre instancias de microservicios. Iniciativas como SmartAirQ [17] demuestran cómo Kubernetes puede respaldar la ejecución escalable de procesos geoespaciales complejos.

Con todo ello, nuestra contribución se distingue por integrar en un planificador de rutas tres aspectos clave: a) una arquitectura abierta, nativa en la nube y basada en este paradigma, b) el uso en tiempo real de datos provenientes

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de múltiples contaminantes regulados por la OMS y c) la personalización de rutas urbanas basada en el estado de salud del usuario y su sensibilidad a contaminantes. Esta combinación permite calcular rutas que minimizan la exposición individual a la contaminación atmosférica de forma dinámica, eficiente y adaptable a contextos urbanos complejos, superando las limitaciones de soluciones actuales tanto comerciales como académicas.

III. ANÁLISIS

Esta sección presenta un análisis conciso de las principales características de la arquitectura propuesta y del caso de uso asociado. Se analizan los requisitos no funcionales, los posibles cuellos de botella, el flujo de trabajo de la aplicación y la justificación de las tecnologías seleccionadas. Asimismo, se incluye una descripción resumida del uso de la Métrica Compleja Personalizada (*Personalized Complex Metric*, PCM) para el enrutamiento personalizado y del método para aumentar la resolución espacial de la distribución de contaminantes, para el planificador, utilizando técnicas de Kriging.

A. Flujo de trabajo, requisitos y cuellos de botella

El flujo principal de trabajo se detalla en la Figura 1. En ella se muestra cómo cada 10 minutos se inicia un ciclo que recolecta y procesa los datos de las AQMS en tiempo real. Si se detectan valores erróneos o faltantes (por ejemplo, debido a mantenimiento de sensores), se realiza una estimación basada en medidas históricas. Una vez validados, estos datos se guardan en la base de datos, alimentando el cómputo de Kriging para la interpolación espacial de la calidad del aire y la creación de un grafo del entorno espacial con los niveles de contaminantes. Este grafo se almacena en un volumen persistente, listo para ser utilizado por el servicio de rutas personalizadas que calcula y entrega las rutas solicitadas por el usuario.

Los requisitos no funcionales asociados al caso de uso se resumen en la Tabla II. Para poder cumplir con estos requisitos, es fundamental identificar y mitigar los posibles cuellos de botella del sistema. Entre los principales se encuentran la dependencia de fuentes de datos externas, la carga computacional asociada a la planificación de rutas, así como la necesidad de mecanismos eficientes de precomputación y de capacidades de autoescalado para garantizar el rendimiento bajo condiciones variables de demanda.

B. Análisis tecnológico

1) Algoritmo de enrutamiento y PCM: Para calcular rutas que minimicen la exposición a la contaminación adaptadas a la salud del usuario, se utiliza el algoritmo de Dijkstra sobre una representación gráfica de la ciudad. El coste $c(u, v)$ de recorrer cada arista (segmento de calle) entre los nodos u y v se define mediante la PCM. Esta métrica se construye combinando dos factores:

a) Un valor de contaminación ponderado $w_p(u, v)$ para la arista, calculado como la suma ponderada de las concentraciones normalizadas $p_i(u, v)$ de cada contaminante i (por ejemplo, NO_2, O_3, PM2.5) con los valores de la Tabla I [7]. Además, para acentuar las situaciones de riesgo, cada concentración normalizada se ajusta mediante una función exponencial tratando de penalizar concentraciones elevadas, de forma que:

$$p_i^{\text{ajustado}}(u, v) = p_i^k; \quad k \geq 1; \tag{1}$$

y el valor ponderado para la arista se obtiene como:

$$w_p(u, v) = \sum_{i \in P} w_i \cdot p_i^{\text{ajustado}}(u, v) \tag{2}$$

donde P es el conjunto de contaminantes considerados y $\sum w_i = 1$. Los pesos w_i se ajustan según el perfil de salud del usuario (personas asmáticas, embarazadas o mayores), lo que permite priorizar los contaminantes más perjudiciales para grupos particularmente vulnerables.
En la práctica, el valor utilizado de k en la Ecuación (1) es de 6, penalizando los valores superiores a 1 tras la normalización, que son aquellos que exceden los máximos de la OMS [7]. k se ajusta en función de los niveles medios de contaminación.

b) La distancia física $d(u, v)$ de la arista.

Esto se combina linealmente para obtener el coste final de la arista:

$$c(u, v) = (\alpha \cdot w_p(u, v) + (1 - \alpha)) \cdot d(u, v) \tag{3}$$

donde α es un parámetro configurable ($0 \leq \alpha \leq 1$) que permite equilibrar la prioridad entre minimizar la exposición a la contaminación (α alto) y minimizar la distancia del recorrido (α bajo).

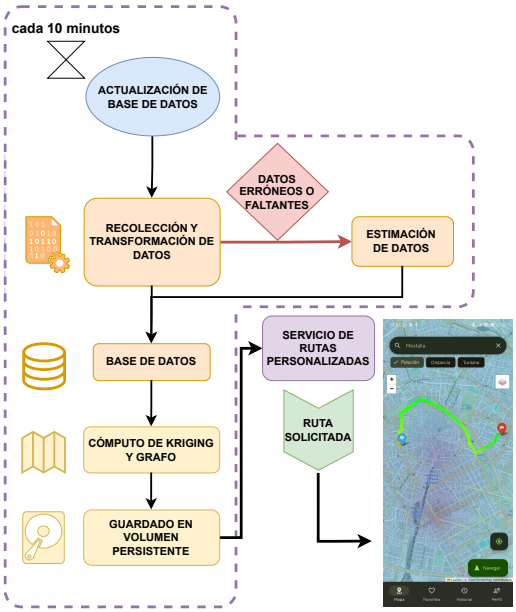


Fig. 1: Flujo principal de trabajo de la aplicación

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Requisito	**Objetivo**
Tiempo de respuesta (Response Time, RT) del cálculo de rutas de extremo a extremo vía API	$\leq 2\,$s con 100 usuarios concurrentes
Frecuencia de actualización de la interpolación	$\leq 10\,$min después de la llegada de datos del AQMS
Solicitudes concurrentes	≥ 50 solicitudes/segundo con $< 1\%$ de errores
Disponibilidad del sistema	$\geq 99.9\%$ por mes
Almacenamiento de datos históricos de calidad del aire (AQ)	$\geq 1\,$año con intervalos de 10 minutos

Tabla II: Requisitos no funcionales para HRP

2) Interpolación espacial de los contaminantes: Estimar los niveles de contaminación a través del mapa, especialmente entre las ubicaciones de los sensores, es necesario para nuestro caso de uso. El método principal utilizado para esta estimación espacial es Kriging, una técnica geoestadística que permite estimar valores en ubicaciones no muestreadas a partir de datos obtenidos en puntos cercanos, aprovechando la autocorrelación espacial inherente a los datos. Esto permite generar de manera estadística un mapa continuo de la distribución de contaminantes.

3) Backend, Nube y Frontend: Antes de detallar la arquitectura técnica, es importante contextualizar brevemente los conceptos de *frontend* y *backend*. En el desarrollo de aplicaciones, el *frontend* hace referencia a la interfaz que interactúa directamente con el usuario, mientras que el *backend* comprende los sistemas y servicios que procesan la lógica interna, almacenan datos y responden a las solicitudes del *frontend*. Ambos componentes se comunican mediante Interfaces de Programación de Aplicación (*Application Programming Interface*, APIs) y constituyen juntos la estructura completa de una aplicación moderna.

Principalmente se hará uso de un diseño políglota, empleando distintas tecnologías según la naturaleza y requisitos de cada componente del sistema. El *backend* utiliza Java con Spring Boot para determinados servicios y gestión de bases de datos. Además hace uso de Python con FastAPI para tareas computacionalmente intensivas como el enrutamiento y la interpolación, beneficiándose de librerías como networkx y pykrige.

Para la infraestructura, cada unidad funcional se empaqueta en un entorno aislado (contenedores Docker) por su consistencia y aislamiento. Kubernetes se utiliza para la orquestación de dichos entornos, proporcionando autoescalado automatizado (*Horizontal Pod Autoscaler*), balanceo de carga, auto-recuperación y políticas de red para poder cumplir con los requisitos no funcionales de concurrencia y tiempo de actividad.

El *frontend*, se desarrolla usando React Native, permitiendo una única base de código para aplicaciones móviles multiplataforma (iOS y Android). Esta interfaz interactúa con el *backend* a través de APIs, lo que ofrece flexibilidad en la elección de tecnologías.

IV. DISEÑO DE LA ARQUITECTURA

La MSA propuesta se orquesta mediante Kubernetes, utilizando contenedores para gestionar eficientemente sus distintos componentes o microservicios. Cada contenedor está diseñado para cumplir una función específica dentro del sistema, como la gestión de bases de datos, la recolección de datos en tiempo real o el cálculo de rutas. Este diseño se organiza en tres secciones: bases de datos, lógica de negocio y acceso externo.

A. Bases de datos

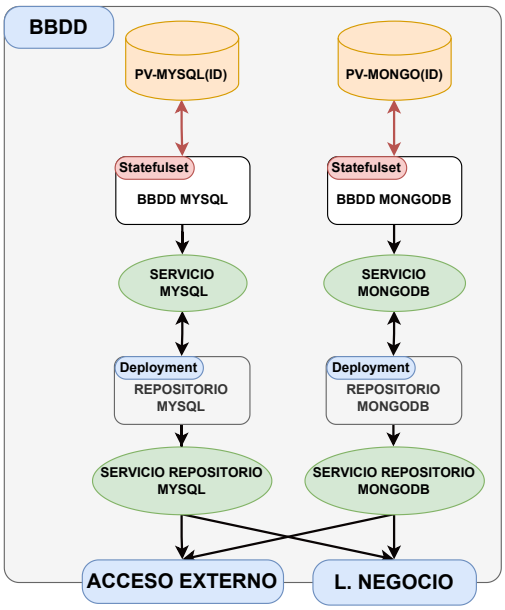


Fig. 2: Diagrama de la arquitectura de las bases de datos.

Adoptando un enfoque de persistencia políglota, se hace uso de dos sistemas de gestión de bases de datos (*Database Management System*, DBMS) principales: MySQL, para almacenar información estructurada de las AQMS y sensores; y MongoDB, dedicado a la gestión de datos en tiempo real provenientes de las AQMS.

Como se puede observar en la Figura 2, la persistencia de los datos en ambos casos se garantiza mediante

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

volúmenes persistentes de Kubernetes (*Persistent Volumes*, PVs). Las instancias de cada DBMS son gestionadas por *StatefulSets*, que aseguran una identidad de red estable y se exponen a través de Servicios de Kubernetes para su acceso. Además, para que las bases de datos sean fácilmente accesibles para los demás microservicios, se agregan unos componentes adicionales de repositorios de datos. Estos actúan como intermediarios, proporcionando una interfaz unificada (basada en llamadas API) para la gestión y operación de los datos, desplegándose en el clúster como Deployments de Kubernetes y expuestos a través de Servicios de Kubernetes. Estos repositorios son utilizados tanto por los módulos de la capa de Acceso Externo como por los de Lógica de Negocio.

B. Lógica de negocio

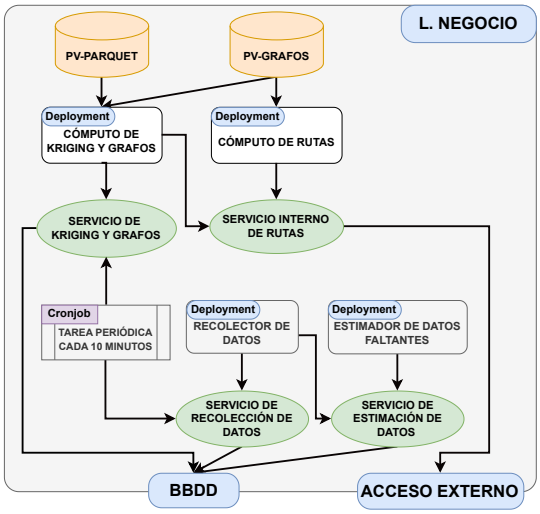


Fig. 3: Diagrama de la arquitectura de lógica de negocio.

Esta sección cubre el conjunto de reglas y procesos que determinan cómo se crean, almacenan y manipulan los datos dentro de la arquitectura, En la Figura 3 se muestra la capa de lógica de negocio, donde cada componente se implementa como un Deployment. Estos utilizan PVs cuando requieren almacenar datos que deben persistir, y se exponen mediante un Servicio de Kubernetes para facilitar la comunicación entre ellos y con otros componentes. Las funcionalidades de esta sección son las siguientes:

- **Obtención de datos en tiempo real** Esta funcionalidad está compuesta por dos componentes distintos. Cada 10 minutos, mediante una llamada periódica programada con un Kubernetes CronJob, se obtienen nuevos datos de las estaciones AQMS. Si se detectan datos faltantes o erróneos (debido a mantenimiento o fallos en los sensores), se infieren los valores en tiempo real basándose en los promedios históricos registrados a la misma hora y en el mismo día. Finalmente, los datos se almacenan en la base de datos MongoDB.
- **Generador de mapa de contaminación**. El microservicio de cómputo de Kriging y grafos, realiza la interpolación espacial Kriging con los datos de las AQMS previamente almacenados. Después, integra estos valores de contaminación en el grafo del área de interés (Valencia), calcula los valores ponderados para cada perfil predeterminado de la PCM y genera mapas de calor para cada contaminante y perfil.
  El grafo y mapas de calor resultantes, se almacenan en un PV, como un archivo.
- **Cálculo de ruta saludable**. El componente de cómputo de rutas carga en memoria el grafo previamente generado. Al recibir una solicitud del usuario con los parámetros correspondientes, se calcula la ruta óptima. Estos parámetros incluyen las coordenadas de origen y destino, el perfil de salud del usuario y sus preferencias de ruta, como por ejemplo optar por la ruta menos contaminada o la más corta en distancia.
  Después de calcular la ruta óptima, esta es devuelta al usuario y se representa en el mapa de la aplicación, junto con los mapas de calor correspondientes a cada contaminante.

C. Acceso externo

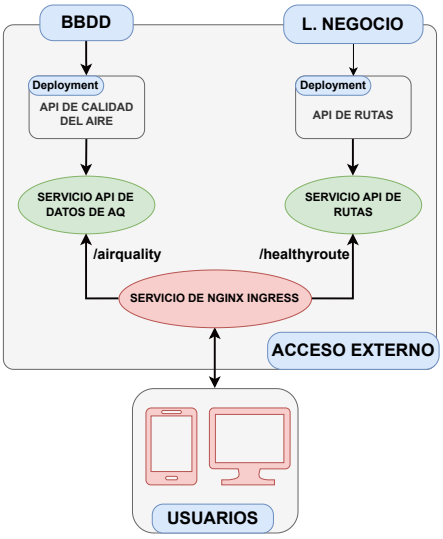


Fig. 4: Diagrama de la arquitectura de acceso externo.

En la Figura 4 se muestra la parte de acceso externo, compuesta por dos puntos de acceso principales que exponen APIs. Estos se despliegan en dos Deployments independientes: uno dedicado a la gestión y consulta de bases de datos, y otro responsable del cálculo de rutas en función de parámetros definidos por el usuario. Además, se hace uso de un servicio Ingress de Nginx que redirige solicitudes a los componentes internos según el sufijo de la URL.

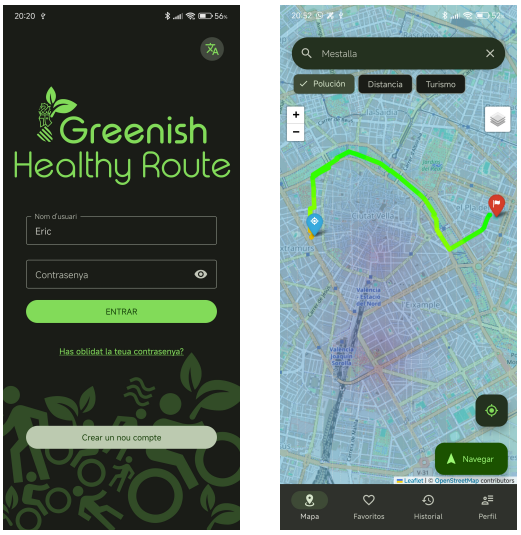
This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

V. RESULTADOS

Esta sección presenta los resultados de la evaluación de la aplicación, cubriendo la interfaz de usuario, la calidad de las rutas generadas según los diferentes perfiles de salud, y el rendimiento computacional. El código del proyecto y un vídeo de demostración están disponibles en [18].

A. *Interfaz de usuario*

La aplicación multiplataforma fue desarrollada utilizando *React Native*, adoptando la estética moderna de Material Design 3. La Figura 5 muestra varias capturas de pantalla de la interfaz gráfica de usuario de la aplicación móvil.



(a) Inicio de sesión (b) Ruta saludable

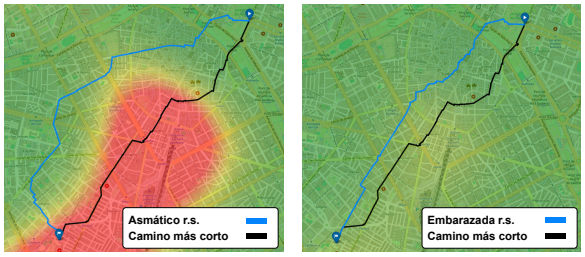
Fig. 5: Capturas de pantalla de la interfaz de usuario

B. *Evaluación de la calidad de las rutas*

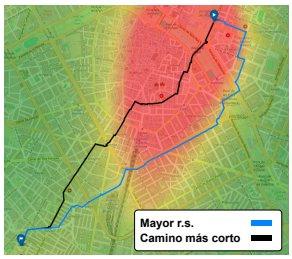
Evaluamos la calidad de las rutas generadas por el HRP comparando el Incremento en la Distancia recorrida (*Increased Distance*, ID) y la Reducción de la Contaminación (*Pollution Reduction*, PR) con respecto al camino más corto. Para ello, se consideraron los perfiles de salud de personas asmáticas, embarazadas y mayores, además de ajustar el parámetro α en la Ecuación (3) y establecer $k = 6$ en la Ecuación (1).

En una primera fase, se utilizaron datos sintéticos de calidad del aire, simulando niveles de contaminación extremadamente altos en las 12 AQMS de Valencia, con el objetivo de demostrar la capacidad del HRP para identificar rutas más saludables bajo condiciones adversas. La Figura 6 ilustra ejemplos correspondientes a distintos perfiles cuando se establece un valor $\alpha = 0{,}5$. En estos escenarios, la priorización de la salud mostró una reducción sustancial de la exposición a la contaminación (aproximadamente un 18% de PR), aunque también puede conllevar incrementos en la distancia recorrida (hasta casi un 10% de ID) en comparación con la ruta más corta.

En la Figura 7 se muestran los resultados para las rutas de un perfil asmático utilizando los mismos datos sintéticos y distintos valores de α. Las rutas se pueden clasificar en: a) rutas más cortas pero con mayor exposición a contaminación y b) rutas con la menor exposición posible.



(a) Perfil asmático (b) Perfil de embarazada



(c) Perfil de persona mayor

Fig. 6: Ejemplos de rutas con el camino más corto ($\alpha = 0$) y rutas saludables según los distintos perfiles (r.s.) ($\alpha = 0.5$) empleando datos sintéticos.

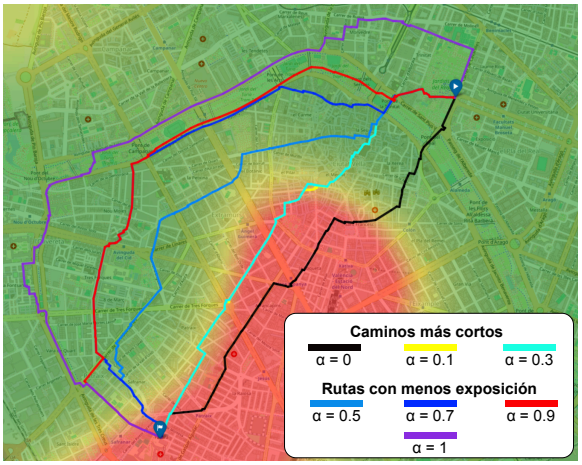


Fig. 7: Ejemplos de rutas para el perfil de asmático variando el parámetro α empleando datos sintéticos.

Para la evaluación bajo condiciones reales de calidad del aire, se emplearon datos en tiempo real de las 12 AQMS correspondientes a dos días representativos (un lunes y un sábado), para tres rutas distintas (A, B y C) mostradas en la Tabla III. Estas pruebas se llevaron a cabo utilizando un valor de $\alpha = 0{,}5$ en la Ecuación (3), lo que implica una ponderación práctica del 50 % al componente

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Ruta	Origen (lat, lon)	Destino (lat, lon)
A	39.469148, -0.405475	39.471525, -0.354835
B	39.479564, -0.366457	39.454503, -0.394550
C	39.475181, -0.387617	39.45787, -0.33689

Tabla III: Configuraciones de rutas evaluadas bajo condiciones reales de calidad del aire, representadas mediante coordenadas geográficas (latitud, longitud).

de salud (PCM) y del 50 % a la distancia, con el objetivo de alcanzar un equilibrio funcional entre exposición a la contaminación y longitud del trayecto.

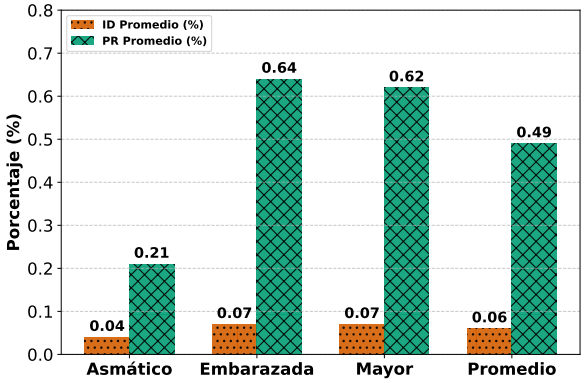


Fig. 8: Promedio del Incremento de Distancia (ID) y Reducción de Contaminación (PR) para los perfiles asmático, embarazada, persona mayor y el promedio general considerando $\alpha = 0{,}5$ en escenarios reales

La Figura 8 presenta los valores promedio del ID y PR obtenidos en todas las pruebas realizadas con el valor de $\alpha = 0{,}5$, diferenciando entre los perfiles de salud y el promedio general. Como se observa, el HRP logró una ligera reducción de la exposición, a costa de un incremento mínimo en la distancia recorrida en los distintos perfiles.

En promedio, considerando todos los escenarios reales y perfiles, se obtuvo una PR media del 0,49%, con un coste medio de solo un 0,06% de aumento en la distancia. Esta reducción tan pequeña se debe a que, durante los dos días de evaluación, los niveles de contaminación no fueron elevados; en consecuencia, muchas de las rutas con menor exposición coincidían con la ruta más corta, lo que limitó la posibilidad de lograr una reducción media apreciable.

C. Evaluación del tiempo de respuesta

El rendimiento del componente de cómputo de rutas fue evaluado utilizando un único nodo de Kubernetes con las siguientes especificaciones: CPU Intel Core i7-12700 y 16 GB de RAM DDR5. Se simularon solicitudes concurrentes mediante Apache JMeter, variando de 5 a 100 usuarios simultáneos, con un total de 5000 solicitudes por prueba.

La Tabla IV presenta el tiempo de respuesta promedio, el *throughput* y la tasa de errores debidos a *timeouts*. Los resultados muestran que tanto los tiempos de respuesta como las tasas de error aumentan con el incremento en la cantidad de solicitudes concurrentes. No obstante, el microservicio mostró un buen rendimiento hasta alcanzar las 80 solicitudes concurrentes, logrando un *throughput* de 57,76 solicitudes por segundo, con un tiempo de respuesta promedio inferior a 1000 ms (997,96ms) y una tasa de error menor al 1%. Este desempeño cumple con los requisitos no funcionales de tiempo de respuesta y escalabilidad, según lo establecido en la Tabla II.

Cabe destacar que, dado que las pruebas se realizaron sobre un único nodo de Kubernetes, los resultados representan un escenario base. Una de las principales ventajas de utilizar Kubernetes es que, al disponer de múltiples nodos en un clúster, la aplicación puede escalar dinámicamente para atender a un mayor número de usuarios. Esto se logra sin necesidad de realizar cambios en la arquitectura, simplemente aumentando el número de réplicas del microservicio dentro del clúster, lo que permite mejorar los tiempos de respuesta y la capacidad de procesamiento.

Solicitudes	RT prom. (desv.)[ms]	Err. [%]	Throughput
5	105.77 (67.20)	0.00	46.85
10	161.91 (109.90)	0.00	52.36
20	277.12 (186.90)	0.00	49.22
40	545.18 (318.60)	0.00	53.21
60	768.07 (716.85)	0.00	56.40
80	997.96 (870.98)	0.00	57.76
100	1318.21 (971.78)	1.68	54.57

Tabla IV: Resultados de rendimiento para un solo nodo sin réplicas con diferentes solicitudes concurrentes, con detalle del Tiempo de Respuesta (RT) (en promedio (prom.) y desviación estándar (desv.)) en ms, porcentaje de solicitudes con errores (Err.) debido a *timeouts* y *throughput* (solicitudes por segundo).

VI. CONCLUSIONES

Este trabajo presenta el diseño y la evaluación de una aplicación para la planificación de rutas saludables personalizadas, basada en el estado de salud del usuario y en datos de calidad del aire en tiempo real, orientada a su despliegue a gran escala y en la nube.

Entre las principales contribuciones se encuentra el desarrollo de una arquitectura escalable basada en microservicios, construida sobre tecnologías abiertas y nativas de la nube. El sistema integra datos en tiempo real de contaminantes regulados por la OMS, obtenidos a partir de estaciones de monitoreo de calidad del aire, y emplea técnicas de interpolación espacial basadas en Kriging para generar mapas dinámicos de concentración. Esta información permite calcular rutas que minimizan la exposición a contaminantes específicos en base al perfil de cada usuario (como personas asmáticas, embarazadas o mayores), proponiendo un enfoque que difiere de las soluciones basadas únicamente en niveles bajos de contaminación.

La aplicación fue validada en diferentes escenarios de contaminación (sintéticos y reales), mostrando una reducción media durante las pruebas realizadas de 18% en la exposición a la contaminación, con un incremento del 10% en la distancia recorrida y con un tiempo medio

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de respuesta inferior a un segundo (997,96 ms) para 80 usuarios concurrentes en un entorno de despliegue de nodo único. Si utilizamos escenarios reales, estos porcentajes son inferiores, ya que los niveles de contaminación son mucho más bajos.

En comparación con soluciones actuales, tanto comerciales como académicas, esta propuesta destaca por su capacidad de adaptación dinámica, personalización y por su arquitectura flexible y abierta. Como trabajo futuro, se priorizará la evaluación de la misma en entornos distribuidos con múltiples nodos, así como mejoras continuas en eficiencia, seguridad y usabilidad.

AGRADECIMIENTOS

El artículo ha sido parcialmente financiado por el proyecto PID2021-126823OB- I00 por MCIN/AEI/ 10.13039/ 501100011033 y por European Union NextGeneration EU/ PRTR; por Generalitat Valenciana con ayudas CIAICO/2022/ 179, CIACIF/2023/ 416 y CIAEST/2024/ 71 así como el Ministerio de Educación en las becas de movilidad PRX23/00589.

REFERENCIAS

[1] S. Li, H. Zhang, Z. Jia, C. Zhong, C. Zhang, Z. Shan, J. Shen, and M. A. Babar, "Understanding and addressing quality attributes of microservices architecture: A systematic literature review," *Information and Software Technology*, vol. 131, p. 106449, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950584920301993

[2] R. Fayos-Jordan, R. Araiz-Chapa, S. Felici-Castell, J. Segura-Garcia, J. J. Perez-Solano, and J. M. Alcaraz-Calero, "Eco4rupa: 5g-iot inclusive and intelligent routing ecosystem with low-cost air quality monitoring," *Information*, vol. 14, no. 8, 2023. [Online]. Available: https://www.mdpi.com/2078-2489/14/8/445

[3] S. Felici Castell, J. J. Pérez Solano, J. Segura Garcia, A. Soriano Asensi, R. Fayos Jordán, J. López Ballester, and G. Mas Requena, "Eco4rupa: Buscando tu ruta," in *Libro de Actas. XVI Jornadas de Ingeniería Telemática (JITEL 2023)*, 2023. [Online]. Available: https://web.salleurl.edu/docsmkt/JITEL/libro-actas-jitel.pdf

[4] The European Parliament and the Council of the European Union, "Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe (recast)," Official Journal of the European Union, L Series, Nov. 2024, oJ L, 20.11.2024, pp. 1–70. ELI: http://data.europa.eu/eli/dir/2024/2881/oj. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024L2881

[5] Conselleria d'Agricultura, Desenvolupament Rural, Emergència Climàtica i Transició Ecològica, "RED VALENCIANA DE VIGILANCIA Y CONTROL DE LA CONTAMINACIÓN ATMOSFÉRICA." https://agroambient.gva.es/va/web/calidad-ambiental/datos-on-line, 2025, accessed: 27/01/2025.

[6] Ajuntament de Valencia, minut a minut, "Estaciones contaminación atmosféricas." https://valencia.opendatasoft.com/, 2025, accessed: 21/01/2025.

[7] WHO agency, "WHO global air quality guidelines," https://apps.who.int/iris/bitstream/handle/10665/345329/9789240034228-eng.pdf, 2025, accessed: 15/01/2025.

[8] Google, "Air quality on google maps," https://support.google.com/maps/answer/11270845?hl=en&sjid=1260932663551655139-EU, 2025, [Accessed 23-05-2025].

[9] J. G. Su, M. Winters, M. Nunes, and M. Brauer, "Designing a route planner to facilitate and promote cycling in metro vancouver, canada," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 7, pp. 495–505, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965856410000546

[10] J. Y. Wang, K. N. Dirks, M. Ehrgott, J. Pearce, and A. K. Cheung, "Supporting healthy route choice for commuter cyclists: The trade-off between travel time and pollutant dose," *Operations Research for Health Care*, vol. 19, pp. 156–164, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211692317301157

[11] Y. Wang, Y. Wu, Z. Li, K. Liao, C. Li, and G. Song, "Route planning for active travel considering air pollution exposure," *Transportation Research Part D: Transport and Environment*, vol. 103, p. 103176, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361920922000086

[12] F. Casino, C. Patsakis, E. Batista, F. Borràs, and A. Martínez-Balleste, "Healthy routes in the smart city: A context-aware mobile recommender," *IEEE Software*, vol. 34, no. 6, pp. 42–47, 2017.

[13] L.-N. Gao, F. Tao, P.-L. Ma, C.-Y. Wang, W. Kong, W.-K. Chen, and T. Zhou, "A short-distance healthy route planning approach," *Journal of Transport & Health*, vol. 24, p. 101314, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214140521003443

[14] J. Helle, A. Poom, E. Willberg, and T. Toivonen, "The green paths route planning software for exposure-optimised active travel," *Journal of Open Research Software*, Oct 2023.

[15] F. Cirillo, D. Gómez, L. Diez, I. Elicegui Maestro, T. B. J. Gilbert, and R. Akhavan, "Smart city iot services creation through large-scale collaboration," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5267–5275, 2020.

[16] F. J. López-Andreu, J. A. López-Morales, J. F. Atenza Juárez, R. Alcaraz, M. D. Hernández, M. Erena, J. A. Domínguez-Gómez, and S. García Galiano, "Monitoring system of the mar menor coastal lagoon (spain) and its watershed basin using the integration of massive heterogeneous data," *Sensors*, vol. 22, no. 17, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/17/6507

[17] A. Kaginalkar, S. Kumar, P. Gargava, N. Kharkar, and D. Niyogi, "Smartairq: A big data governance framework for urban air quality management in smart cities," *Frontiers in Environmental Science*, vol. Volume 10 - 2022, 2022. [Online]. Available: https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2022.785129

[18] E. M. Albalá, "Greenish routes vlc," https://github.com/greenish-lab-vlc/greenish-routes-access, 2024, access to this repository requires permission, contact the main author for access.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Gestión Segura de Datos Médicos en Entornos IoMT mediante Blockchain y Smart Contracts

Miguel Vinaixa Fernández, Jaume Ramis Bibiloni, M. Magdalena Payeras Capellà, Macià Mut Puigserver
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears, Carretera de Valldemossa, Km. 7,5, 07122, Palma
miguel.vinaixa@uib.cat, jaume.ramis@uib.cat, mpayeras@uib.cat, macia.mut@uib.cat

***Palabras Clave*—Blockchain, IoMT (Internet of Medical Things), Smart Contracts, Protección de datos sanitarios**

Presentación de líneas de investigación

El desarrollo del *Internet de las Cosas Médicas* (IoMT) ha propiciado el uso de sensores biomédicos integrados en redes corporales inalámbricas (WBANs) para la monitorización remota y continua del estado de salud de los pacientes, en tiempo real. Esta tecnología genera una cantidad masiva de datos sensibles que requieren ser gestionados de forma segura, privada y automatizada. En este contexto, se propone una arquitectura basada en *blockchain* y *smart contracts* para la recolección, almacenamiento y procesamiento de los datos médicos, preservando al mismo tiempo la soberanía de los datos del paciente y su derecho a decidir con quién desea compartir dicha información.

El sistema considera un conjunto de actores con roles diferenciados [1]: autoridades reguladoras, centros médicos, laboratorios clínicos, médicos, pacientes, y dispositivos IoMT. En una primera fase, se establece un mecanismo de registro y autenticación de estos actores. Posteriormente, los dispositivos IoMT, a través de la WBAN, recolectan datos clínicos que se almacenan de forma distribuida utilizando sistemas como IPFS. En la *blockchain* sólo se registran los *hashes* correspondientes para asegurar la integridad de los historiales clínicos electrónicos (EHRs).

Para la gestión de credenciales de acceso, se plantea el uso de **RejSBTs** (Rejectable Soulbound Tokens), una evolución de los Soulbound Tokens que garantiza la intransferibilidad de las credenciales, la aceptación explícita por parte del usuario y la prueba de recepción. Esta tecnología responde a limitaciones observadas en el uso de NFTs para representar credenciales digitales, tal como se discute en trabajos previos [2]. La implementación de RejSBTs permite asignar términos y condiciones durante la emisión de la credencial, facilitando una gestión robusta del acceso a los EHRs.

Los *smart contracts* se encargan de procesar automáticamente los datos recolectados, evaluando condiciones clínicas predeterminadas, como los niveles de batería de un marcapasos o signos vitales críticos según el sistema EWS (*Emergency Warning Score*). Estos contratos pueden disparar eventos que activen respuestas médicas automatizadas, notificaciones o autorizaciones de acceso, permitiendo así una actuación rápida y segura ante situaciones de emergencia. Además, mediante la notificación certificada de dichas situaciones al paciente, se busca fomentar su participación activa en el cuidado de su salud, motivándolo a adoptar las medidas necesarias para mejorar su estado clínico.

El sistema propuesto se estructura en tres fases principales:

1. Registro e identificación de los actores participantes.
2. Recolección y almacenamiento seguro de datos clínicos mediante WBANs.
3. Compartición selectiva de datos conforme a las políticas de acceso establecidas por el paciente.

Este enfoque proporciona una solución innovadora que combina privacidad, integridad, interoperabilidad y automatización en el tratamiento de datos médicos, alineada con los principios de protección de datos como el RGPD y con gran potencial en aplicaciones de eSalud.

Agradecimientos

This work was supported by Grant PID2021-122394OB-I00 (BLOBSEC) funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU.

Referencias

[1] Pericàs-Gornals, R., Mut-Puigserver, M., Payeras-Capellà, M.M. "Highly private blockchain-based management system for digital COVID-19 certificatesÍnt. J. Inf. Secur. 21, 1069–1090 (2022). https://doi.org/10.1007/s10207-022-00598-3

[2] Pericàs-Gornals, R., Mut-Puigserver, M., Payeras-Capellá, M.M., Cabot-Nadal, M.À., Ramis-Bibiloni, J. "Digital credentials management system using rejectable soulbound tokens.Ann. Telecommun. 79, 843–855 (2024). https://doi.org/10.1007/s12243-024-01032-6

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Integración de Herramientas de IA Generativa en Tareas Académicas: un Proyecto de Innovación Docente en Ingeniería Telemática

Jaume Ramis Bibiloni, M. Magdalena Payeras Capellà, Rosa Pericàs Gornals,
Loren Carrasco Martorell, Macià Mut Puigserver, Miquel Àngel Cabot Nadal
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
Carretera de Valldemossa, Km. 7,5, 07122, Palma
jaume.ramis@uib.cat, mpayeras@uib.cat, rosa.pericas@uib.cat,
loren.carrasco@uib.cat, macia.mut@uib.cat, miquel.cabot@uib.cat

Resumen

En el contexto de la educación superior, la Inteligencia Artificial (IA) se consolida como una herramienta clave en la transformación del aprendizaje. Este proyecto de innovación docente tiene como objetivo fomentar el uso responsable, ético y creativo de herramientas de IA generativa por parte del alumnado del Grado de Ingeniería Telemática, integrándolas en las tareas académicas para mejorar la calidad, la eficiencia y el desarrollo de competencias técnicas y profesionales. La iniciativa propone una metodología participativa centrada en el estudiante, que incluye la selección de herramientas adecuadas, el rediseño de actividades académicas, la elaboración de materiales de apoyo y el análisis de resultados. Además de contribuir a la mejora del proceso de enseñanza-aprendizaje, el proyecto aborda desafíos como el riesgo de fraude académico o la pérdida de habilidades críticas, y sus resultados podrían ser extrapolables a otros ámbitos universitarios más allá del Grado en Ingeniería Telemática.

***Palabras Clave*—Inteligencia Artificial generativa, innovación docente, ingeniería telemática, competencias digitales, aprendizaje centrado en el estudiante**

I. Introducción

En el contexto actual de la educación superior, la adopción de tecnologías basadas en Inteligencia Artificial (IA) se ha convertido en una realidad en constante expansión. La abundancia de datos y el creciente potencial computacional han facilitado la aparición y consolidación de un amplio abanico de aplicaciones de IA en el ámbito educativo, lo que está transformando las dinámicas tradicionales de enseñanza y aprendizaje [1]. Este fenómeno resulta especialmente relevante en los estudios de Ingeniería Telemática, que, por su naturaleza tecnológica, se encuentran en una posición privilegiada para integrar estas herramientas de forma significativa y estratégica.

Este proyecto de innovación docente surge ante la necesidad de preparar al estudiantado del Grado en Ingeniería Telemática para afrontar los desafíos de un entorno cada vez más digitalizado y automatizado. Aunque la IA ofrece un considerable potencial para enriquecer los procesos de aprendizaje y fortalecer las competencias profesionales, también plantea riesgos y retos importantes que deben abordarse desde la formación universitaria. Entre ellos destacan el uso inadecuado o acrítico de estas herramientas, la generación de desigualdades entre el alumnado en función de su grado de familiarización con la IA, y la posible aparición de conductas asociadas al fraude académico.

Uno de los principales riesgos asociados a la incorporación de herramientas de IA generativa en el ámbito educativo es la dependencia excesiva de estas tecnologías sin una comprensión adecuada de su funcionamiento ni una reflexión crítica sobre sus implicaciones éticas, legales y sociales. Esta situación puede limitar el desarrollo de competencias fundamentales como el pensamiento crítico, la creatividad o la capacidad analítica. Asimismo, la accesibilidad desigual a estas herramientas y la carencia de formación específica pueden generar nuevas desigualdades dentro del aula, reproduciendo brechas digitales entre el estudiantado. También preocupa el uso indebido de sistemas de generación automática de contenidos como medio para eludir el trabajo académico individual, lo cual comprometería la integridad del proceso formativo. En este contexto, resulta de especial interés el informe de la Universidad de Cornell [2], que proporciona orientaciones y recomendaciones para un uso responsable de la IA generativa en la educación, abordando tanto sus oportunidades y beneficios como sus limitaciones, con especial

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

atención a aspectos clave como la integridad académica, la accesibilidad y las implicaciones éticas.

En este sentido, la proliferación de herramientas de IA generativa —capaces de producir texto, código, imágenes o incluso soluciones a problemas técnicos— obliga a los docentes a repensar el diseño de las tareas académicas tradicionales. Actividades como la redacción de informes, la resolución de problemas o la elaboración de proyectos pueden ser fácilmente automatizadas, lo que exige una redefinición de los objetivos de aprendizaje y de las estrategias de evaluación. En lugar de centrarse exclusivamente en la producción de resultados, es necesario fomentar enfoques que prioricen la reflexión crítica, la resolución colaborativa de problemas y el aprendizaje significativo.

Este proyecto tiene como finalidad principal la incorporación de herramientas de IA generativa en el ámbito universitario, promoviendo su uso tanto por parte del estudiantado como del personal docente en contextos académicos reales. La iniciativa apuesta por el diseño, aplicación y evaluación de metodologías educativas innovadoras que favorezcan la mejora del aprendizaje, el desarrollo de habilidades técnicas y profesionales, así como la calidad y eficiencia en la realización de tareas académicas. Además, se fomenta la creatividad y la innovación mediante la exploración de nuevas formas de interacción con la tecnología en entornos de enseñanza y aprendizaje.

Si bien la experiencia se desarrolla en el marco específico del Grado en Ingeniería Telemática, se considera que los resultados y aprendizajes derivados del proyecto pueden ser extrapolables a otras titulaciones universitarias, especialmente aquellas del ámbito tecnológico. En este sentido, se pretende no solo mejorar el proceso de enseñanza-aprendizaje dentro del grado, sino también ofrecer un modelo replicable que pueda inspirar futuras iniciativas en contextos educativos similares. Esta visión se alinea con propuestas recientes que promueven la integración crítica y pedagógica de herramientas de IA en la educación, como el informe de la Universidad de Cornell [2], las recomendaciones planteadas por la UNESCO [3], o la guía elaborada por el Ministerio de Educación de Chile [4], que destacan el potencial de estas tecnologías para transformar positivamente las prácticas docentes y fomentar el aprendizaje activo.

El presente artículo se organiza de la siguiente manera. En la Sección II se presentan los objetivos principales del proyecto de innovación docente. La Sección III describe la metodología adoptada para la integración de herramientas de IA generativa en el proceso de enseñanza-aprendizaje y detalla el plan de trabajo desarrollado para su implementación. En la Sección IV se exponen los resultados preliminares obtenidos durante las primeras fases del proyecto. Finalmente, en la Sección V se presentan las conclusiones y se esbozan algunas líneas de trabajo futuro orientadas a la consolidación y ampliación de esta experiencia innovadora.

II. Objetivos

El proyecto tiene como finalidad principal promover el uso responsable, eficiente y creativo de herramientas de IA generativa en el ámbito académico del Grado en Ingeniería Telemática. En este marco, se definen los siguientes objetivos específicos:

- **Fomentar la familiarización del estudiantado con herramientas de IA generativa**, acercando estas tecnologías al entorno universitario de manera práctica y contextualizada.
- **Integrar el uso de herramientas de IA en actividades académicas** pertenecientes a diferentes asignaturas y cursos del grado, con el propósito de promover su aplicación activa en la realización de tareas tanto académicas como con proyección profesional.
- **Mejorar la calidad y la eficiencia en la realización de tareas académicas y profesionales**, mediante el aprovechamiento del potencial de la IA para asistir en procesos de redacción, análisis, resolución de problemas o desarrollo técnico.
- **Fomentar la creatividad y la innovación en el alumnado**, estimulando el uso de herramientas de IA como medio para abordar problemas de manera original y generar soluciones creativas en contextos académicos diversos.
- **Evaluar el impacto del uso de herramientas de IA en el aprendizaje**, considerando tanto el rendimiento académico como el desarrollo de competencias técnicas y profesionales, con el fin de obtener evidencias sobre su eficacia educativa.

Estos objetivos están orientados a proporcionar una formación integral y adaptada a las demandas de un entorno académico y profesional cada vez más influenciado por la IA. A través de su implementación, se busca no solo mejorar las habilidades técnicas de los estudiantes, sino también fomentar una actitud ética y reflexiva hacia el uso de tecnologías emergentes, contribuyendo así a su desarrollo como profesionales competentes y responsables en el ámbito de la ingeniería telemática.

III. Metodología y Plan de Trabajo para la Implementación del Proyecto

La metodología propuesta para la implementación de este proyecto de innovación docente se basa en un enfoque participativo y centrado en el estudiante, que fomenta la colaboración activa entre profesores y alumnos a lo largo del proceso de aprendizaje. Esta metodología se implementa mediante un plan de trabajo que se estructura en una serie de tareas y fases que guiarán el proceso desde la selección de herramientas de IA hasta la difusión de los resultados obtenidos. Este enfoque metodológico se fundamenta en un proceso iterativo de diseño, implementación, evaluación y mejora continua. A continuación, se describen las tareas y fases que componen el plan de trabajo.

A. Tareas iniciales y de preparación

- **Selección de herramientas de IA.** Se ha llevado a cabo una investigación exhaustiva para identificar y seleccionar las herramientas de IA más adecuadas para el contexto de los estudios de Ingeniería Telemática. Se han priorizado aquellas que son accesibles,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

fáciles de usar y que tienen una alta aplicabilidad en las actividades académicas y profesionales previstas.

- **Descripción detallada de las herramientas seleccionadas.** Para cada herramienta de IA seleccionada, se ha elaborado una descripción detallada especificando las tareas académicas y de evaluación en las que se puede aplicar, como la generación de documentos, traducción, corrección de textos, entre otras.
- **Materiales de soporte.** Se han buscado, revisado y seleccionado guías y recursos ya existentes sobre el uso de herramientas de IA generativa en contextos educativos. A partir de estos materiales, se han elaborado recomendaciones específicas de uso adaptadas a las actividades y objetivos de aprendizaje de las asignaturas participantes en el proyecto. Estas orientaciones han servido como base para integrar de forma coherente y pedagógica las herramientas de IA en el desarrollo de las tareas académicas.
- **Formación de los docentes.** Los profesores implicados en el proyecto han llevado a cabo un proceso de formación para familiarizarse con las herramientas de IA seleccionadas. A través de la exploración autónoma y el aprendizaje entre iguales, han adquirido los conocimientos necesarios para valorar la aplicabilidad de estas herramientas en sus respectivas asignaturas y comenzar a integrarlas en su práctica docente.
- **Análisis de las actividades académicas y de evaluación.** Una vez seleccionadas las herramientas de IA, se ha procedido a analizar de forma detallada todos los elementos docentes y actividades de evaluación en los que se pueda integrar su uso. Se han establecido las relaciones entre las actividades académicas de las distintas asignaturas participantes en el proyecto (informes, proyectos, presentaciones, etc.) y las herramientas de IA susceptibles de ser utilizadas en cada una de ellas, tales como herramientas para la generación de documentos, traducciones, búsquedas bibliográficas, esquemas, generación y revisión de código, entre otras.

La Tabla I muestra la relación entre distintas tipologías de actividades académicas y de evaluación, y los tipos de herramientas de IA generativa que pueden utilizarse para apoyarlas. Las actividades se han agrupado en los siguientes bloques temáticos: actividades escritas y análisis, aprendizaje basado en problemas y casos, comunicación y presentación, trabajo colaborativo y reflexivo, y actividades experimentales o prácticas. Cada celda marcada indica que, para esa actividad concreta, el uso de la herramienta de IA correspondiente puede aportar valor pedagógico, ya sea facilitando la tarea, mejorando su calidad o promoviendo nuevas formas de aprendizaje. Esta matriz permite visualizar de forma clara y sistemática las posibilidades de integración de la IA generativa en el contexto educativo universitario, destacando su versatilidad y aplicabilidad transversal en múltiples formatos de enseñanza y evaluación.

Es importante destacar que, dado el ritmo acelerado de evolución de las herramientas de IA generativa, este trabajo no se centra en describir plataformas o aplicaciones concretas. En lugar de ello, se ha optado por identificar y analizar las tipologías de actividades académicas que pueden beneficiarse del uso de estas tecnologías. Esta aproximación permite ofrecer una propuesta metodológica más duradera y adaptable, centrada en el desarrollo de competencias y estrategias de integración pedagógica, independientemente de las herramientas específicas disponibles en cada momento. Lo relevante no es tanto qué herramienta se utiliza, sino cómo y para qué se emplea dentro del proceso de enseñanza-aprendizaje.

B. Diseño e implementación en el aula

- **Diseño de actividades académicas y de evaluación.** Se ha llevado a cabo la definición de nuevas actividades de aprendizaje y se han redefinido algunas de las ya existentes, con el fin de incorporar el uso de las herramientas de IA seleccionadas. Estas actividades permiten que los estudiantes se familiaricen con el funcionamiento y la aplicación práctica de las herramientas de IA. Las actividades han sido adaptadas a las necesidades e intereses específicos de los estudiantes, y se han integrado en las asignaturas seleccionadas.
- **Implementación del plan metodológico en el aula.** La implementación de esta metodología en el aula se lleva a cabo de manera iterativa, lo que permite realizar ajustes a medida que se obtengan resultados y feedback de los estudiantes. La implementación está acompañada de un seguimiento y evaluación continua a lo largo de todo el proceso.

C. Iteraciones del proyecto

El proceso se lleva a cabo en cuatro iteraciones, cada una centrada en la mejora y consolidación de la metodología propuesta. A continuación se indica, de forma resumida, en qué consisten dichas iteraciones.

- **1ª iteración: Primera fase del Plan Piloto (curso 2024-2025).** La primera fase se ha centrado en la incorporación de la metodología en las siguientes asignaturas del primer semestre:
  - 3er curso, 1er semestre, Modelización de Redes – depuración de código.
  - 4to curso, 1er semestre, Laboratorio de Redes y Proyectos de Telecomunicación – presentaciones y proyecto técnico, ambos en inglés.

  En esta fase, se ha llevado a cabo un seguimiento y evaluación tanto cualitativa como cuantitativa, comparando los resultados con los obtenidos en cursos anteriores. Se han utilizado encuestas y entrevistas con los estudiantes para conocer sus opiniones y experiencias. Esta información ha permitido realizar ajustes en la metodología con el fin de mejorar su efectividad.
- **2ª iteración: Segunda fase del Plan Piloto (curso 2024-2025).** En esta fase se ha integrado la me-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Cuadro I
RELACIÓN ENTRE TIPOS DE ACTIVIDADES ACADÉMICAS Y HERRAMIENTAS DE IA UTILIZADAS.

Actividad académica y de evaluación	Gen. documentos	Mejora textos	Traducciones	Resúmenes	Búsqueda biblio.	Esquemas	Código	Soporte LaTeX	Gráficos	Imágenes
Actividades escritas y análisis										
Redacción de documentos técnicos	X	X	X	X	X			X		
Elaboración de resúmenes	X	X		X		X				
Análisis crítico de artículos científicos	X	X		X	X					
Revisión de literatura	X	X		X	X	X				
Creación de glosarios temáticos	X	X	X		X	X		X		
Aprendizaje basado en problemas y casos										
Resolución de problemas	X	X			X	X	X		X	
Simulación de casos prácticos	X	X	X		X	X	X		X	X
Propuestas de solución a proyectos técnicos	X	X	X		X	X	X	X	X	X
Análisis de impacto o viabilidad de soluciones	X	X		X	X		X		X	
Comunicación y presentación										
Preparación de presentaciones (slides + guion)	X	X	X	X	X	X		X	X	X
Simulación de exposiciones orales	X	X	X	X					X	X
Traducción y adaptación de contenidos	X	X	X		X					
Trabajo colaborativo y reflexivo										
Autoevaluaciones y coevaluaciones	X	X		X						
Diseño colaborativo de proyectos	X	X	X	X	X	X	X	X	X	X
Tareas con roles (líder, revisor, ...)	X	X								
Actividades experimentales o prácticas										
Informes de prácticas de laboratorio	X	X	X	X	X	X		X	X	X
Desarrollo de código (scripts)							X		X	
Interpretación de resultados de simulaciones				X			X		X	X
Generación de tablas y análisis estadístico						X	X	X	X	
Comparación entre herramientas tecnológicas	X	X		X	X	X			X	
Evaluación de seguridad en sistemas					X	X	X		X	

todología en las siguientes asignaturas del segundo semestre:

- 2do curso, 2do semestre, Redes de Operadora – resolución de problemas.
- 3er curso, 2do semestre, Desarrollo de Aplicaciones Telemáticas – presentaciones y proyecto de programación web.

Al igual que en la primera iteración, se ha realizado un seguimiento, evaluación y reajuste de la metodología basado en el feedback de los estudiantes y los resultados obtenidos.

- **3ª iteración: Primera fase del Plan de Consolidación (curso 2025-2026).** En esta fase se consolidará la metodología en las asignaturas de la primera iteración y además se incorporará en la asignatura del 1r semestre de 4to curso *Privacidad e Identidad Digital* (depuración de código y presentaciones). En todas ellas está previsto integrar las modificaciones y mejoras derivadas de los resultados del seguimiento y evaluación previos.
- **4ª iteración: Segunda fase del Plan de Consolidación (curso 2025-2026).** Se implementará la metodología en las asignaturas de la segunda iteración y además se incorporará en la asignatura del 2do semestre de 4to curso *Seguridad del Software* (depuración de código, predicción de riesgos), integrando las lecciones aprendidas de las fases anteriores.

D. Evaluación del efecto de la innovación

La evaluación del impacto de la metodología innovadora propuesta se lleva a cabo de manera continua a lo largo del proyecto, con el objetivo de medir su eficacia y la percepción de los estudiantes sobre su uso. A continuación, se detallan los procesos para evaluar el efecto de la innovación implementada:

1. **Cuestionarios de percepción y aceptación.** Al final de cada una de las cuatro iteraciones contempladas en el proceso de implementación de la propuesta en el aula, se realiza un cuestionario anónimo a los estudiantes para conocer su percepción y grado de aceptación de la metodología utilizada. El objetivo es llevar a cabo un seguimiento y evaluación continuos durante todo el proceso.
2. **Análisis de los resultados de las encuestas habituales.** Se analizan los resultados y comentarios obtenidos en las encuestas de evaluación docente que se realizan habitualmente en la UIB. Se anima a los estudiantes a incluir sus comentarios sobre la metodología utilizada para fomentar las buenas prácticas en el uso de las herramientas de IA en las tareas académicas encargadas.
3. **Análisis de la calidad de las tareas académicas.** Los profesores responsables realizan un análisis de la calidad de las tareas académicas realizadas por los estudiantes en el año académico en curso, comparándolas con la calidad observada en los cursos anterio-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

res. Además, se revisan las calificaciones numéricas obtenidas y se comparan con las correspondientes a los cursos pasados.

4. **Análisis de la carga de trabajo.** Es fundamental analizar la carga de trabajo que esta propuesta implica tanto para los profesores como para los estudiantes, y valorar los beneficios que se derivan de la misma. Este análisis permite identificar los desafíos y esfuerzos adicionales requeridos, asegurando que los beneficios potenciales, como la mejora de las competencias académicas y profesionales, justifican los recursos y el tiempo invertidos. Evaluar cuidadosamente estos aspectos garantiza una implementación equilibrada y efectiva, maximizando los beneficios educativos y minimizando el impacto negativo en las cargas de trabajo.

E. Difusión de resultados y conclusiones

Finalmente, entre los objetivos de este proyecto se encuentra la difusión de los resultados y las buenas prácticas obtenidas, tanto dentro como fuera de la institución educativa. El propósito es compartir la experiencia sobre el uso de herramientas de IA en la docencia como metodología innovadora para mejorar la calidad y la eficiencia del trabajo de los estudiantes, así como promover la adopción de herramientas de IA en otros programas de estudio y en otros contextos educativos.

Este plan de trabajo asegura un enfoque riguroso y flexible para la implementación de las herramientas de IA en el aula, con una evaluación continua y una mejora iterativa que garantizará su efectividad en el proceso de enseñanza-aprendizaje.

IV. Resultados Preliminares

En esta sección se presentan los resultados obtenidos de la implementación del proyecto en sus primeras fases, concretamente en la primera iteración y los resultados preliminares de la segunda iteración. Estos resultados proporcionan una visión inicial sobre el impacto de la incorporación de herramientas de IA en las actividades académicas y el aprendizaje de los estudiantes.

A. Primera Iteración: Resultados Iniciales

Durante la primera iteración del proyecto, que se desarrolló en el primer semestre del curso 2024-2025, se implementaron las herramientas de IA en las siguientes asignaturas:

- **3r curso, 1r semestre, Modelización de Redes:** Se utilizó la depuración de código como actividad principal. Los estudiantes mostraron una rápida adopción de las herramientas de IA, con un incremento en la eficiencia en la depuración del código en comparación con años anteriores.
- **4t curso, 1r semestre, Laboratorio de Redes y Proyectos de Telecomunicación:** Se implementaron herramientas de IA para la creación de presentaciones y la elaboración de proyectos técnicos. Los estudiantes demostraron una mayor creatividad en el diseño de sus presentaciones, así como una mejora en la calidad de los proyectos técnicos entregados.

En general, los primeros resultados reflejan un uso positivo de las herramientas de IA, con mejoras en la eficiencia de las tareas y en la calidad de los productos entregados por los estudiantes. Sin embargo, también se identificaron algunos desafíos, como la necesidad de mayor familiarización con las herramientas por parte de los estudiantes y algunos aspectos relacionados con la integración fluida de la IA en las actividades académicas.

B. Segunda Iteración: Resultados Preliminares

En la segunda iteración, correspondiente al segundo semestre del curso 2024-2025, se ha incorporado el uso de herramientas de IA en las siguientes asignaturas:

- **2n curso, 2n semestre, Redes de Operadora:** Se está utilizando IA para la resolución de una colección de problemas. Los resultados preliminares indican que los estudiantes están utilizando eficazmente la IA para resolver problemas complejos, mejorando tanto su velocidad de resolución como la precisión de sus respuestas.
- **3r curso, 2n semestre, Desarrollo de Aplicaciones Telemáticas:** Se ha implementado IA como herramienta de apoyo para la elaboración de presentaciones y en el desarrollo de proyectos web. Los estudiantes destacan su utilidad para organizar y estructurar ideas, así como los beneficios que ofrece en el ámbito de la programación. No obstante, reconocen la importancia de ser conscientes de los riesgos asociados con su uso.

Aunque los resultados preliminares son prometedores, es necesario continuar monitorizando la integración de la IA en estas asignaturas para ajustar las estrategias pedagógicas y asegurar que se maximicen los beneficios para los estudiantes.

C. Resultados de las encuestas

Se ha llevado a cabo un seguimiento continuo de los resultados obtenidos en las dos primeras iteraciones. Las encuestas realizadas a los estudiantes ayudan a identificar áreas de mejora, tanto en el uso de las herramientas como en la implementación de las actividades. Estas encuestas incluyen preguntas relacionadas con los usos que han dado a la IA en diferentes tipos de pruebas evaluables, como sesiones de prácticas, presentaciones orales, trabajos de investigación, proyectos, redacción en inglés, generación de gráficos..., la valoración de los resultados obtenidos, su percepción sobre el impacto de la ayuda recibida, su opinión general sobre el uso de la IA, su opinión sobre si deberían ser utilizadas en otras asignaturas, su opinión sobre el aprendizaje del uso crítico y ético de la IA en las asignaturas de grado y su opinión general sobre el proyecto. A continuación se muestran y analizan algunos de los resultados obtenidos en las encuestas para las asignaturas que forman parte de estas dos primeras iteraciones del proceso.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

TIPOLOGIA: ¿Para qué tareas has utilizado las herramientas de IA generativa en los diferentes tipos de pruebas evaluables? (Selecciona todas las que correspondan):

1. Traducción de enunciados o terminología.
2. Corrección o mejora del formato, estructura o diseño.
3. Generación de frases o textos.
4. Creación del índice u organización de los contenidos.
5. Mejora de la redacción del texto técnico.
6. Generación de imágenes o gráficos de apoyo.
7. Detección de errores y corrección de código.
8. Ayuda para la programación o elaboración de fragmentos de código.
9. Realización de diagramas, organigramas o elementos dinámicos.
10. Otros.

La figura 1 recoge los resultados agregados de esta pregunta en las diferentes asignaturas. Cada columna expresa el porcentaje de alumnos que han marcado esta opción en la encuesta. Estos resultados muestran el uso combinado en diferentes tipologías de pruebas. Cabe decir que el resultado no es homogéneo por asignaturas. El resultado combinado permite valorar el uso que los alumnos realizan en sus estudios a través de diferentes asignaturas.

IMPACTO: ¿Cómo valorarías el uso de herramientas de IA generativa para las tareas evaluables?

1. Muy útiles.
2. Moderadamente útiles.
3. Poco útiles.
4. Nada útiles.

La figura 2 recoge los resultados agregados de esta pregunta en las diferentes asignaturas. Cada columna expresa el porcentaje de alumnos que han marcado esta opción en la encuesta.

FACILIDAD DE USO: ¿Cómo valoras la facilidad de uso de herramientas de IA generativa para resolver tareas académicas?

1. Muy fácil.
2. Fácil.
3. Moderadamente difícil.
4. Muy difícil.

Los resultados agregados de esta pregunta en las diferentes asignaturas se representan en la figura 3. Cada columna expresa el porcentaje de alumnos que han marcado esta opción en la encuesta. Seguidamente, en la Sección D, se describirán las principales dificultades encontradas por los alumnos.

HERRAMIENTAS UTILIZADAS: En las encuestas se ha preguntado sobre la herramienta de IA utilizada. En el caso de pruebas de evaluación individuales los alumnos podían escoger libremente la herramienta a utilizar mientras que en aquellas realizadas por parejas, las herramientas utilizadas por los dos miembros de la pareja debían ser diferentes. En los resultados, recogidos en la figura 4, se observa como este requisito ha llevado a los alumnos a explorar nuevas herramientas aunque en los trabajos individuales el uso se ha concentrado en aquellas herramientas más populares. La figura 4 muestra los resultados agregados de esta pregunta en las diferentes asignaturas, donde cada columna expresa el porcentaje de alumnos que han marcado esta opción en la encuesta.

Para finalizar el análisis de los resultados de las encuestas, se detallan a continuación las respuestas a diversas cuestiones planteadas a los alumnos con el fin de evaluar su experiencia con la utilización sistemática de herramientas de IA generativa en sus tareas académicas. En lo que respecta al grado de satisfacción general con la metodología implementada en este proyecto, el 43 % se muestra muy satisfecho, el 51 % satisfecho y el restante 6 % considera neutra o negativa la experiencia. Cuando los alumnos son preguntados acerca de si les gustaría incrementar el uso de este tipo de herramientas en las actividades académicas del grado, el 40 % está completamente a favor, el 53 % consideran que sí pero de forma moderada y tan solo el 7 % se manifiesta en contra de su uso. Estos resultados reflejan una percepción general mayoritariamente positiva de los alumnos de la metodología propuesta.

D. Resultados del seguimiento y valoración del alumnado

Los primeros resultados sugieren que la integración de herramientas de IA generativa está teniendo un impacto positivo en el proceso de aprendizaje, si bien los estudiantes han señalado también una serie de retos y consideraciones importantes que deben ser tenidas en cuenta. En primer lugar, destacan que estas herramientas no son infalibles: las respuestas generadas pueden contener errores, no ajustarse a la pregunta planteada o incluso incluir información inventada, lo que puede confundir en lugar de ayudar. También se ha señalado que, en ocasiones, las respuestas son demasiado generales, carecen de concreción o presentan elementos superfluos, lo que obliga a realizar un análisis crítico y una revisión cuidadosa por parte del usuario. Igualmente, algunos estudiantes han identificado que la generación de imágenes, en particular, todavía presenta limitaciones importantes en cuanto a calidad y precisión.

Varios estudiantes advierten del riesgo de una excesiva automatización del trabajo académico, lo que puede limitar el estudio efectivo de los contenidos y fomentar una dependencia poco crítica de estas tecnologías. Esta tendencia puede tener consecuencias en el desarrollo del pensamiento crítico y la originalidad, generando textos previsibles o mecánicamente construidos. Asimismo, se han señalado dificultades a la hora de integrar de forma adecuada la producción generada por la IA con las aportaciones personales del alumnado, lo que, en algunos casos, puede derivar en un sentimiento de fraude o la percepción de estar "haciendo trampas".

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

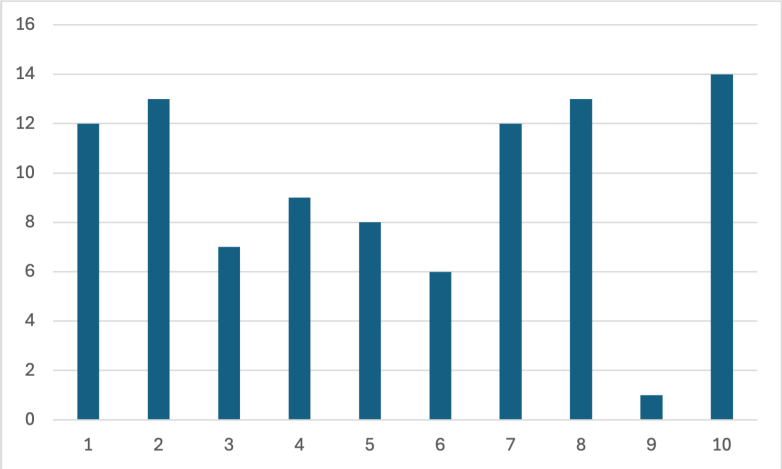


Figura 1. Porcentajes de las tipologías de usos.

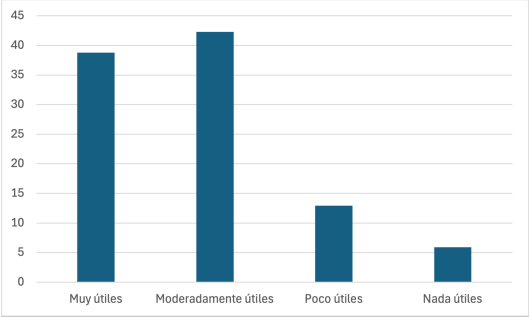


Figura 2. Porcentajes de las valoraciones del impacto del uso de la IA.

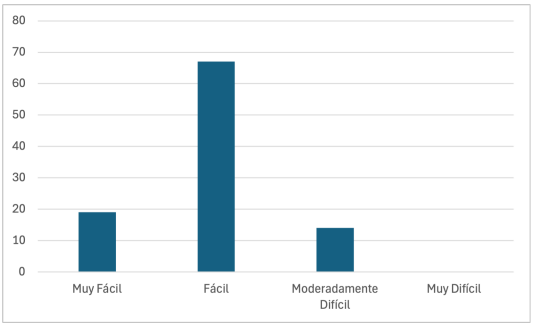


Figura 3. Porcentajes de las valoraciones de la facilidad de uso de la IA.

Por otro lado, se valora positivamente el uso de la IA generativa como herramienta para obtener información, redactar y resumir textos, mejorar la productividad y organizar contenidos tanto en documentos como en presentaciones. Muchos estudiantes indican que estas herramientas pueden ayudar a desbloquear momentos de duda, ahorrar tiempo en tareas mecánicas y aportar una primera base sobre la que trabajar. No obstante, subrayan la importancia de contextualizar adecuadamente las preguntas y formular el *prompt* correcto para obtener resultados relevantes, lo cual requiere claridad en los objetivos y en el tipo de respuesta esperada.

Finalmente, el alumnado insiste en la necesidad de referenciar siempre el uso de herramientas de IA generativa en los trabajos académicos, y de contrastar cuidadosamente la información obtenida, dado que estas tecnologías pueden reproducir sesgos o proporcionar datos incorrectos. En conjunto, los comentarios reflejan tanto el potencial como las limitaciones de la IA generativa, así como la necesidad de un uso consciente, moderado y ético por parte del estudiantado.

Los resultados completos se evaluarán de forma más exhaustiva al final de las cuatro iteraciones y se compararán con los resultados de cursos anteriores para analizar el impacto en el rendimiento académico y la percepción de los estudiantes sobre el uso de la IA en sus estudios.

V. Conclusiones y trabajo futuro

El auge de las herramientas de IA generativa representa un punto de inflexión en los procesos de aprendizaje y producción académica. Tomar conciencia de su potencial transformador resulta esencial, no solo por su impacto inmediato en el rendimiento académico, sino también por su influencia en la futura inserción laboral y profesional de los estudiantes. En este nuevo contexto, es previsible la aparición de una brecha significativa entre quienes adquieren competencia en el uso crítico y eficaz de estas tecnologías y quienes carecen de dicha preparación. Esta brecha, que puede agravar desigualdades existentes, constituye a la vez una oportunidad de mejora y una amenaza si no se gestiona adecuadamente desde el ámbito educativo. Por ello, se hace imprescindible promover un uso ético, consciente y formativo de la IA en entornos universitarios, asegurando que los estudiantes no solo conozcan las herramientas, sino que comprendan sus limitaciones, implicaciones y el papel activo que deben asumir como usuarios responsables.

El aprovechamiento efectivo de las herramientas de IA generativa no depende únicamente de su disponibilidad, sino de la capacidad del usuario para interactuar con ellas de forma adecuada. Resulta fundamental saber formular

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

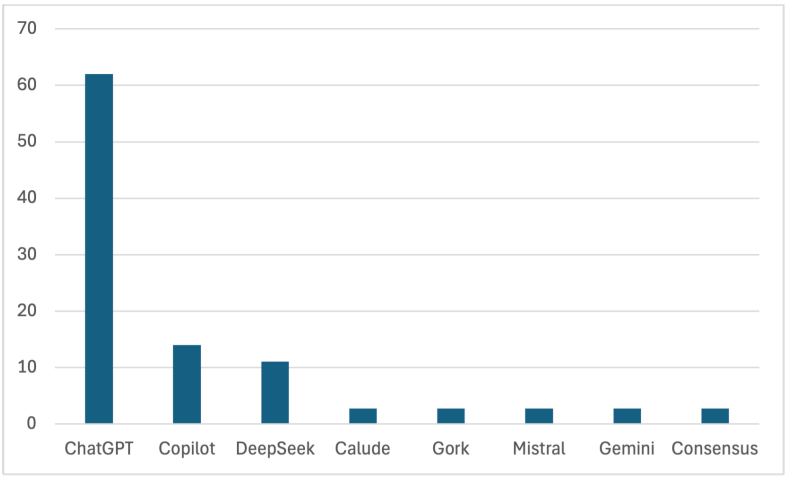


Figura 4. Porcentajes de las herramientas de IA utilizadas.

correctamente las preguntas o indicaciones para obtener respuestas relevantes, útiles y coherentes. Esta habilidad se convierte en una competencia clave dentro del entorno educativo actual. No obstante, el uso de estas herramientas debe ir acompañado de una actitud crítica por parte del estudiante, quien debe evaluar cuidadosamente la calidad, la veracidad y la pertinencia de la información generada. La verificación de los datos y la capacidad de discernimiento son esenciales para evitar la propagación de errores o sesgos, y para asegurar un aprendizaje significativo basado en el uso responsable de la IA.

En el contexto académico y científico actual, es esencial establecer prácticas claras y transparentes en relación con el uso de herramientas de IA generativa. Una de las recomendaciones emergentes más relevantes es la de citar explícitamente el uso de dichas herramientas cuando se han empleado en la elaboración de trabajos académicos o publicaciones científicas. Diversas revistas y congresos ya exigen a los autores la inclusión de un apartado específico (*AI disclosure statement*) donde se detalle qué herramientas de IA se han utilizado, con qué propósito y en qué medida han contribuido al resultado final. Asimismo, se exige que los autores asuman la responsabilidad completa sobre el contenido presentado, dejando constancia de que han revisado, interpretado y validado críticamente cualquier aportación generada por estas herramientas. Esta práctica no solo garantiza la integridad académica, sino que también promueve un uso ético, consciente y formativo de la IA en entornos educativos.

Este proyecto se encuentra actualmente en el ecuador de su desarrollo, habiéndose implementado parcialmente durante el curso académico 2024-2025. Si bien ya contamos con algunos resultados preliminares que nos permiten realizar una primera evaluación de las actividades y herramientas empleadas, aún disponemos de un curso académico más para continuar con su implementación y maximizar los beneficios de esta experiencia de innovación docente. Durante este próximo curso, se prevé realizar ajustes basados en el análisis de los resultados obtenidos hasta el momento, con el fin de optimizar las metodologías empleadas y asegurar una integración más efectiva de las herramientas de IA generativa en el proceso de enseñanza-aprendizaje. Así, se espera continuar con el seguimiento y la evaluación de su impacto, con el objetivo de consolidar los aprendizajes obtenidos y contribuir a la evolución de las prácticas educativas en el ámbito tecnológico.

AGRADECIMIENTOS

Actividad realizada en el marco del proyecto de innovación concedido per el IRIE (UIB) nº PID242654, Título "Fomento de las buenas prácticas en el uso de herramientas de inteligencia artificial generativa en tareas académicas del grado de ingeniería Telemática".

Declaración sobre el uso de herramientas de IA generativa y asistida en el proceso de redacción: Durante la preparación de este trabajo, los autores utilizaron herramientas de IA generativa, específicamente OpenAI ChatGPT, para asistir en la mejora de la claridad, legibilidad y fluidez del texto. Después, los autores revisaron y editaron el contenido de manera exhaustiva, asegurándose de que toda la información presentada fuera precisa y relevante. Los autores asumen plena responsabilidad sobre el contenido y los resultados presentados en este artículo.

REFERENCIAS

[1] Rose Luckin, Wayne Holmes, Mark Griffiths, Laurie B. Forcier, Intelligence Unleashed: An Argument for AI in Education, Pearson Education, url=https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence-Unleashed-Publication.pdf, 2016.

[2] Generative Artificial Intelligence for Education and Pedagogy, 2023, url=https://teaching.cornell.edu/generative-artificial-intelligence

[3] Emma Sabzalieva, Arianna Valentini, ChatGPT e Inteligencia Artificial en la educacion superior, UNESCO, 2023, URL=https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa

[4] Guía para Docentes: Cómo usar ChatGPT para potenciar el aprendizaje activo, Centro de Innovación, Ministerio de Educación, Gobierno de Chile, 2023, url=https://ciudadaniadigital.mineduc.cl/wp-content/uploads/2023/05/Guia-para-Docentes-Como-usar-ChatGPT-Mineduc.pdf

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XV Jornadas
Actas de las XVII Jornadas
de Ingeniería Telemática
(JITEL 2025),
Cáceres (España),
12-14 de noviembre de 2025.

ISBN

Estafa a través de ataques BEC: Un análisis de casos reales

David Melendi[1], Xabiel G. Pañeda[1], Víctor Corcoba[1], Alejandro G. Pañeda[2], Roberto García[1]

(1) Departamento Informática
Universidad Oviedo
Campus de Gijón, Gijón/Xixón, Asturies, España

(2) Bosch Service Solutions SAU
Hermanos Garcia Noblejas, 19, Madrid, España

melendi@uniovi.es, xabiel@uniovi.es, corcobavictor@uniovi.es, alejandro.garciapaneda@es.bosch.com, garciaroberto@uniovi.es

Las situaciones de compromiso de correo electrónico empresarial (BEC) se han convertido en los últimos años en una gran preocupación para todo tipo de organizaciones. Existen múltiples variantes de estos ataques que provocan incidentes de seguridad con un gran coste económico y reputacional. En este trabajo se analiza un tipo de situación concreta, relacionado con la interceptación de comunicaciones entre empresas y la manipulación de información de facturación, con la finalidad de cometer una estafa. El estudio incluye una descripción detallada de la forma de actuar de los estafadores, así como de las diferentes variantes de este ataque que hemos podido analizar en los últimos años, debido a nuestra actividad en el campo pericial. El trabajo permite a expertos en ciberseguridad y forensia digital comprender este tipo de situaciones. Por ello, puede ser de gran ayuda tanto en el campo de la prevención como en el desarrollo de una investigación sobre uno de estos incidentes. Adicionalmente, el trabajo también incluye algunas propuestas desde el punto de vista de la regulación en el despliegue de servicios de Internet.

***Palabras Clave*- BEC, correo electrónico, forensia, evidencia digital, informe pericial, estafa**

I. Introducción

Las situaciones de compromiso de correo electrónico empresarial (BEC) se dan cuando unos atacantes manipulan unas comunicaciones de correo electrónico o comprometen unas cuentas de correo electrónico con la finalidad de engañar a una compañía o a sus empleados. Este tipo de incidentes de seguridad se ha convertido en uno de los riesgos más significativos para todo tipo de organizaciones en los últimos años, con un gran coste económico y reputacional. Solo en Estados Unidos, en 2023 se denunciaron al Internet Crime Complaint Center del FBI 21.489 casos de incidentes BEC [1]. Se estima que esos casos supusieron unas pérdidas de 2.900 millones de dólares, siendo el segundo tipo de delito más costoso del año en ese país. En el caso de España no hay cifras específicas de incidentes BEC, pero la cifra más elevada de delitos informáticos en el año 2023 corresponde a los procedimientos judiciales por estafas y defraudaciones que ascienden a 19.464 [2]. Esto supone un 82,87% del total de los procedimientos incoados en el año por delitos informáticos, con un considerable crecimiento en los últimos años (72,43% en 2020). Estos datos se reflejan igualmente en el Informe sobre Cibercriminalidad del Ministerio del Interior correspondiente al año 2023, con un incremento del 26% en las investigaciones policiales por estafa (el 90,5% de las iniciadas por cualquier clase de ciberdelito) [3].

Un incidente BEC también puede terminar en un acceso no autorizado a información confidencial. Puede ser que la información protegida forme parte de unas comunicaciones intervenidas, o que los atacantes se hagan con esta información utilizando técnicas de ingeniería social, o porque a partir del incidente se obtienen los mecanismos de acceso a datos protegidos. Sea como sea, un tercero puede tener acceso a información confidencial a través de una situación BEC. Los daños de una exfiltración pueden ser económicos, estimándose en una media global de 4,88 millones de dólares por incidente [4]. Pero los daños también son reputacionales para la organización afectada, además de poder ser la fuente de nuevos crímenes que puedan realizar con la información obtenida (por ejemplo, una estafa [5]).

Existen distintas tipologías de ciberataques basados en situaciones BEC, siendo las clases más relevantes las siguientes [6]:

- Cuentas comprometidas, que consiste en hacerse con el control de una cuenta de un empleado de una empresa para utilizarla como vehículo para realizar operaciones fraudulentas. Por ejemplo, para solicitar pagos a los clientes de la organización.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- El fraude del CEO, que consiste principalmente en que un atacante suplanta la identidad de uno de los responsables de la compañía para ordenar a los empleados de la misma realizar transferencias o pagos fraudulentos.
- Suplantación de servicios jurídicos, que es similar al fraude del CEO. Los atacantes suplantan la identidad de los servicios jurídicos de la compañía para solicitar a un empleado de la misma que realice determinadas operaciones, típicamente transferencias bancarias o envíos de datos.
- Estafa de las tarjetas regalo, que es otra variante del fraude del CEO en la que los atacantes suplantan la identidad de un responsable de la empresa para ordenar a algún empleado adquirir distintas tarjetas regalo (códigos) que deben enviarles. Supuestamente, estas tarjetas forman parte de un programa de recompensas en la compañía.
- Envío de facturas fraudulentas, que consiste en que los estafadores envían facturas falsas o solicitudes de pago fraudulentas haciéndose pasar por un proveedor de la organización estafada.

Muchas de estas situaciones se producen porque se utiliza un servicio inseguro por naturaleza, para soportar la operativa de la empresa. Así, los ataques BEC suelen explotar factores humanos aprovechándose de ciertas vulnerabilidades. De ahí que muchos de ellos explotan aspectos psicológicos como la autoridad, la urgencia o la confianza [7]. Aunque es complicado erradicar totalmente este tipo de ataques por la ausencia de cargas maliciosas en los mensajes y su alto grado de personalización [8], existen mecanismos de protección que pueden, al menos, mitigar su impacto. No obstante, dotar al correo electrónico de seguridad depende de muchos agentes: usuarios finales, personal técnico de las organizaciones (cuando existe), proveedores, proveedores de proveedores, etc. La responsabilidad de implantar estas medidas se difumina y cualquier punto débil en la cadena de agentes involucrados propicia la aparición de incidentes. Algunas de estas situaciones se podrían evitar mediante la utilización de mecanismos de firma electrónica. No obstante, muchos usuarios desconocen su funcionamiento y su utilidad.

En este trabajo se analiza una variante de esta última tipología de situaciones BEC. Esta variante concreta consiste en unos atacantes que intervienen las comunicaciones entre un proveedor y un cliente para alterar la información de pago de un servicio o suministro. La finalidad última es la de engañar al cliente para que realice el pago en una cuenta controlada por el estafador. Hemos podido analizar múltiples casos de incidentes de este tipo, debido a su experiencia como analistas forenses y peritos en más de 150 casos reales de distinta naturaleza. Por ello, el trabajo incluye una descripción detallada de la forma de actuar de los estafadores, así como de las diferentes variantes del ataque a los que hemos tenido acceso en los últimos años. El trabajo permite a expertos en ciberseguridad y forensia digital comprender este tipo de situaciones. Por ello, puede ser de gran ayuda tanto en el campo de la prevención como en el desarrollo de una investigación sobre uno de estos incidentes. Adicionalmente, el trabajo también incluye algunas propuestas desde el punto de vista de la regulación en el despliegue de servicios de Internet.

El resto del artículo está organizado de la siguiente forma. La Sección II describe el tipo de situación BEC concreto objeto de este estudio, consistente en la interceptación y manipulación de datos de facturación. En la Sección III se analizan distintos casos reales en los hemos trabajado. En la Sección IV se realiza una discusión sobre los motivos detrás de este tipo de situaciones. Por último, la Sección V presenta las principales conclusiones del estudio y algunas líneas futuras de trabajo.

II. Incidentes BEC de Interceptación y Manipulación de Facturas

Este trabajo estudia una variante de situación BEC de envío de facturas fraudulentas. En concreto, la variante está relacionada con la interceptación de comunicaciones y la manipulación de información de facturación con la finalidad de cometer una estafa. La secuencia de un incidente de este tipo se detalla a continuación:

1. Existe una relación mercantil entre dos organizaciones, en la que una actúa de proveedor de la otra, que actúa de cliente.
2. Tras la prestación de un servicio o la realización de un suministro, el proveedor envía una o varias facturas por correo electrónico al cliente. El texto del correo, o las facturas que se adjuntan al mismo, contienen información sobre la cuenta bancaria en la que es necesario realizar el pago.
3. El cliente recibe las facturas por correo electrónico, y procede con el pago siguiendo las instrucciones que se han recibido.
4. El proveedor, días más tarde, se percata de que no ha recibido el pago de las facturas. Por ello, contacta con el cliente para interesarse por el pago de las mismas.
5. El cliente responde al proveedor que ya ha pagado las facturas. Frecuentemente, envía uno o varios justificantes bancarios de los pagos realizados.
6. A partir de ese punto, proveedor y cliente intercambian una serie de comunicaciones hasta que se percatan de que las instrucciones de pago recibidas por el cliente no son las mismas que había enviado el proveedor. Es habitual que las facturas se envíen en formato PDF y contengan la cuenta bancaria en la que se debe realizar el pago. Tal y como se observa en la Figura 1, las facturas enviadas y las recibidas son idénticas, salvo por el número de cuenta bancaria.

Lo que ha sucedido es que un tercero ha tenido acceso al correo electrónico (o correos) en el que se enviaban las facturas y ha modificado la información de pago. En ocasiones manipulando el texto del mensaje y en otras ocasiones modificando los adjuntos que se envían. Es decir, este incidente consiste en:

1. Una interceptación de un mensaje legítimo. El atacante intercepta el mensaje de correo electrónico original para poder manipularlo con posterioridad.

Actas de las XV Jornadas
Actas de las XVII Jornadas
de Ingeniería Telemática
(JITEL 2025),
Cáceres (España),
12-14 de noviembre de 2025.

ISBN

2. Una suplantación de identidad del remitente. Frecuentemente, el atacante suplanta al remitente para hacer llegar al destinatario un mensaje idéntico al original, pero con los datos de pago modificados.

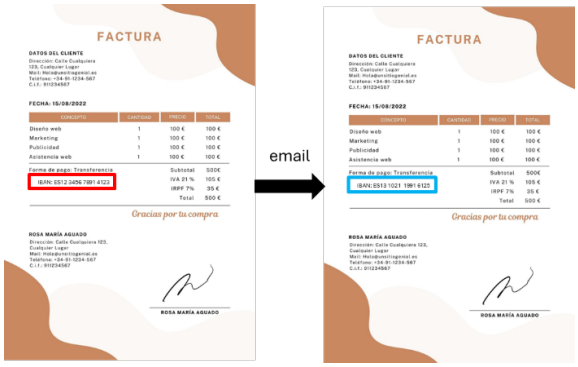


Fig. 1. Comparación de facturas legítima y fraudulenta

Lo habitual es que las empresas se echen la culpa la una a la otra. El proveedor indica que envió las facturas con las instrucciones de pago correctas y el cliente indica que se realizó el pago de esas facturas siguiendo las instrucciones enviadas por el proveedor. Las dos partes suelen conservar los mensajes de correo electrónico que han enviado o recibido para mantener su postura. Como quiera que el servicio se ha prestado (o el bien se ha suministrado), el incidente termina en los tribunales a través de una demanda ordinaria por el impago de las facturas realizada por el proveedor. Es decir, se inicia un procedimiento civil entre proveedor y cliente, quedando fuera del mismo los proveedores de los servicios de correo electrónico y los autores de la manipulación de la factura. En este proceso, las partes aportan las comunicaciones realizadas y, frecuentemente, informes periciales de parte que analizan lo sucedido. El proveedor solicita que se dicte sentencia por la que se condene al cliente al pago de la cantidad adeudada con los intereses correspondientes e imposición de costas. Por otro lado, el cliente demandado solicita la desestimación de la demanda alegando que el pago tiene efectos liberatorios al haberse efectuado de buena fe. Ambas partes coinciden en que se ha producido una estafa informática, pero no se ponen de acuerdo en cuál de ellas ha sido la estafada. Es decir, surge un conflicto entre las partes afectadas sin investigar el origen real del fraude. Es decir, este proceso no busca condenar al autor de los hechos, si no ver quién asume las pérdidas.

III. Análisis de Casos Reales

En nuestra experiencia como peritos y analistas forenses, na investigación de este tipo se realiza principalmente en base a las siguientes evidencias:

- Cabeceras de correo electrónico, tanto de los mensajes fraudulentos como de mensajes legítimos que se han intercambiado entre las partes implicadas.
- Metadatos en archivos adjuntos.
- Registros DNS con información de acceso público sobre las políticas de seguridad implantadas por las organizaciones involucradas.
- Registros de auditoría de los sistemas desplegados, cuando esto es posible.
- Registros de autenticación y acceso al servicio, cuando estos registros están disponibles.
- Parametrizaciones del servicio, cuando es posible el acceso a las mismas [9].

A partir de las evidencias recogidas en numerosos casos de esta naturaleza que han podido ser investigados, podemos deducir que la intervención de los mensajes interceptados se produce habitualmente por alguna de las siguientes situaciones:

- Robo de credenciales del remitente del mensaje, tal y como se muestra en la Figura 2.
- Interceptación del mensaje en un punto intermedio de la infraestructura, como se observa en la Figura 3.
- Manipulación de un mensaje una vez que se ha depositado en el buzón de destinatario, tal y como se muestra en la Figura 4.

La parte clave está en la recolección e interceptación de los mensajes originales. En Mitre se dispone de una categoría específica de potenciales técnicas que han sido detectadas en casos reales con esta finalidad [10].

A. Robo de credenciales del remitente

En este caso, los estafadores han robado las credenciales del remitente y han interceptado el mensaje original. La interceptación se ha realizado probablemente mediante una regla configurada en la cuenta de correo afectada, que reenvía los mensajes de interés a otra cuenta controlada por los estafadores. Recibido el mensaje original que tenía la información de facturación, los estafadores han modificado su contenido y lo han reenviado utilizando la cuenta legítima del remitente. Es decir, el mensaje fraudulento es indistinguible del mensaje original, tanto a simple vista, como por la información de control que incluye, dado que parte del mismo punto exacto y sigue el mismo camino.

B. Interceptación en un punto intermedio

En este caso, el mensaje se ha interceptado mediante una intervención en algún punto intermedio de la infraestructura de mensajería. Es decir, los estafadores han conseguido un acceso de un determinado nivel de privilegios a alguno de los servidores que intervienen en la fase de transferencia del mensaje. Este acceso se pudo realizar de distintas formas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

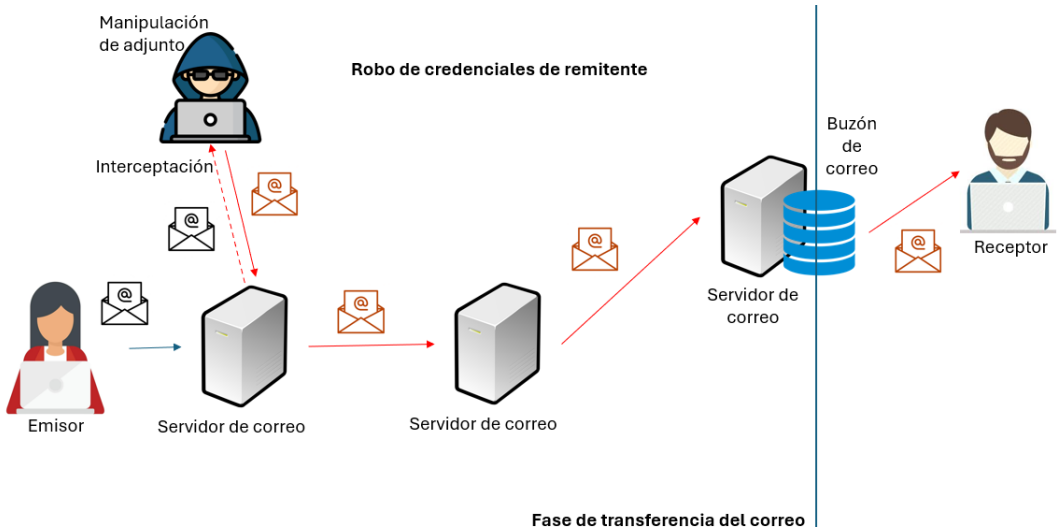


Fig. 2. Robo de credenciales en el remitente

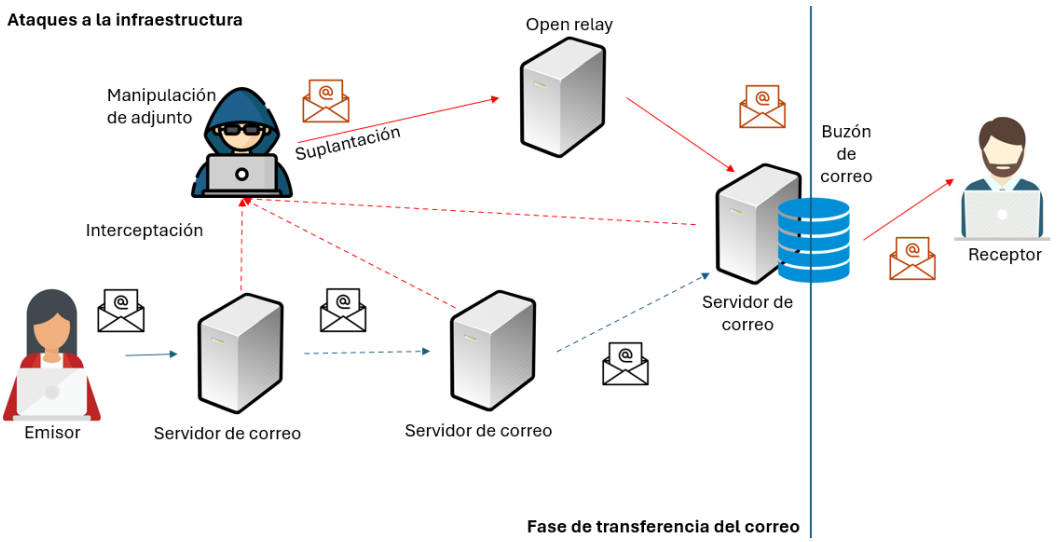


Fig. 3. Interceptación mediante un ataque a la infraestructura

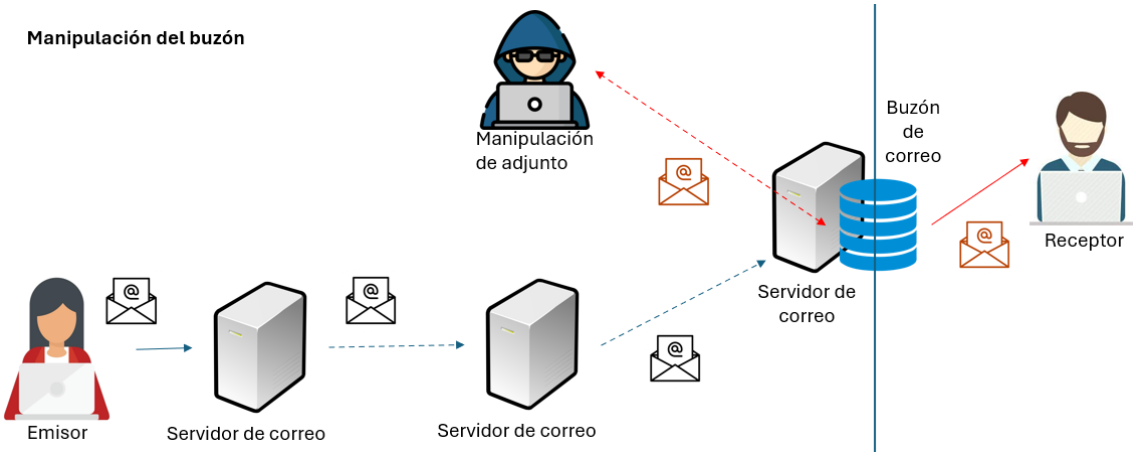


Fig. 4. Manipulación de un mensaje en el buzón del destinatario

Actas de las XV Jornadas
Actas de las XVII Jornadas
de Ingeniería Telemática
(JITEL 2025),
Cáceres (España),
12-14 de noviembre de 2025.

ISBN

Puede haber sido una intrusión en esos servidores, puede haberse aprovechado una vulnerabilidad del software o puede haberse producido un robo de credenciales del destinatario. Según cómo se haya producido, la intrusión se realizó en la infraestructura del remitente de la factura o en la infraestructura del destinatario.

En estos casos, el modus operandi suele ser el de interceptar los mensajes a través de esos accesos a la infraestructura, mediante la configuración de reglas temporales de reenvío. Estas reglas reenvían los mensajes de interés a una cuenta de correo intermedia en poder de los estafadores. Una vez que los mensajes se han recibido en esa cuenta intermedia, se manipula su contenido para modificar las instrucciones de pago y los mensajes alterados se reenvían al destinatario legítimo.

Frecuentemente, los estafadores reenvían el mensaje original a la cuenta intermedia impidiendo que este sea remitido por la infraestructura al destinatario legítimo. No obstante, en uno de los casos estudiados, el destinatario de la información recibió tanto el mensaje original como el mensaje modificado por los estafadores, con una diferencia de una hora y media aproximadamente.

Nos hemos encontrado con distintos métodos utilizados por los estafadores para enviar el mensaje alterado al destinatario original:

- A través de una cuenta con una apariencia similar a la original (por ejemplo, mediante homoglifos).
- Empleando un *open relay* y utilizando la cuenta de correo original tanto en la cabecera de los mensajes como en el sobre SMTP.
- Utilizando la infraestructura y una cuenta de un servicio legítimo, de la que se han apoderado los estafadores. Adicionalmente, reflejando la cuenta de correo original en la cabecera de los mensajes.

En el primer caso, cuando el mensaje se envía desde una cuenta de apariencia similar, el destinatario puede percibir algún problema según la similitud entre las direcciones de correo electrónico. No obstante, dependiendo del servicio utilizado, es posible que no haya indicios técnicos que puedan revelar la ilegitimidad del mensaje. Esta situación, es la menos frecuente de las que nos hemos encontrado.

En los dos últimos casos, el mensaje puede ser indistinguible del original a simple vista. Es decir, el destinatario del mensaje no puede diferenciarlo del mensaje legítimo, de haberlo recibido. Solamente es posible detectar las diferencias a través de un análisis de cabeceras de los mensajes, dado que las *Received* muestran un camino distinto del habitual. No obstante, este tipo de análisis está fuera del alcance de un usuario sin conocimientos técnicos. Hemos podido determinar la forma en la que se realizan este tipo de estafas debido al rastro que los sistemas utilizados por los estafadores suelen dejar en cabeceras como *Tthread-Index*, *In-Reply-To* o *References* en los mensajes.

Cuando se utiliza una cuenta de un servicio legítimo para enviar el mensaje fraudulento, también existe un reflejo de las medidas de seguridad implantadas por la organización responsable de dicho servicio en cabeceras como *Received-SPF*, *DKIM-signature* o *Authentication-Results*. Un ejemplo de cabeceras se muestra en la Figura 5. En este caso, se utiliza una cuenta real cuyas credenciales han sido robadas con anterioridad, en un servicio legítimo con autenticación de remitentes, SPF y DKIM configurados. Gracias al análisis de cabeceras, se observa que se ha llevado a cabo la autenticación del usuario remitente (*Authenticated sender*) y que la dirección de remitente utilizada en el sobre SMTP (*smtp.mailfrom*) es la del servicio legítimo. También se observa que el mensaje fraudulento se envía desde una infraestructura con SPF y DKIM, pero la ausencia de DMARC impide verificar el alineamiento entre la dirección de remitente utilizada en el sobre SMTP (*smtp.mailfrom*) y la dirección utilizada en la cabecera *From* del mensaje (*header.from*). En cuanto al uso de un *open relay*, los síntomas son similares.

La efectividad de estos ataques depende totalmente de que la organización remitente del mensaje original tenga activados los mecanismos de validación adecuados y de que la organización destinataria realice las comprobaciones adecuadas de dichos mecanismos. En ambos casos, la cabecera *From* del mensaje (*header.from*) contiene la dirección del remitente original del correo. La ausencia de validaciones es la que permite engañar al usuario al que va destinada la información de facturación. De hecho, las organizaciones implicadas en los incidentes que se han podido analizar tenían implantadas medidas de seguridad de correo muy básicas:

- Algunas de las organizaciones remitentes de los mensajes disponían de registros SPF publicados, pero muchas no lo hacían. Por otro lado, muchos de los destinatarios de los mensajes no realizaban validaciones de registros SPF, lo que permite la suplantación de remitentes de correo con independencia de lo diligentes que éstos hubieran sido. En cuanto a la configuración de estas políticas, solo encontramos una organización que tenía un *hardfail* por defecto en sus registros SPF. En este caso particular, la estafa pudo ser perpetrada porque los destinatarios no realizaban validaciones SPF. No obstante, esta política *hardfail* por defecto sí generaba errores cuando los estafadores enviaban los mensajes fraudulentos con copia a una cuenta en Gmail. Es decir, los estafadores enviaban copias de sus mensajes a una cuenta de Gmail. Era la infraestructura de Gmail la que, al recibir esas copias, detectaba el origen fraudulento de los mensajes y generaba mensajes de error automatizados. Estos mensajes de error eran recibidos por los remitentes originales de los mensajes, pero estos usuarios no sabían cómo interpretar los avisos, procediendo a su descarte.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

```
...
Received: from wmx-2.organizacion_puente.eu (AAA.BBB.CCC.119) by
 DM6NAM11FT027.mail.protection.outlook.com (10.13.172.205) with Microsoft SMTP
 Server (version=TLS1_2, cipher=TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384) id
 XX.XX.XXXX.17 via Frontend Transport; Thu, 17 Aug 2023 12:39:39 +0000
Received: from localhost (localhost [127.0.0.1])
    by mail2.organizacion_puente.fr (Postfix) with ESMTP id CXXXXXXXXX
    for <destinatarioOriginal@hotmail.com>; Thu, 17 Aug 2023 14:39:37 +0200 (CEST)
Received: from mail2.organizacion_puente.fr ([127.0.0.1])
    by localhost (organizacion_puente.fr [127.0.0.1]) (amavisd-new, port 10024)
    with ESMTP id YYYY_Y-YYYYY for <destinatarioOriginal@hotmail.com>;
    Thu, 17 Aug 2023 14:39:37 +0200 (CEST)
Received: from mail.organizacion_puente_dominio_2.fr (localhost [127.0.0.1])
    (using TLSv1.2 with cipher ECDHE-RSA-AES256-GCM-SHA384 (256/256 bits))
    (No client certificate requested)
    (Authenticated sender: contact@organizacion_puente_dominio_2.fr)
    by mail2.organizacion_puente.fr (Postfix) with ESMTPSA id ZZZZZZZZZZ
    for <destinatarioOriginal@hotmail.com>; Thu, 17 Aug 2023 14:39:37 +0200 (CEST)
From: =?utf-82B2...?= <usuario@organizacionRemitente.com>
To: "destinatarioOriginal@hotmail.com" <destinatarioOriginal@hotmail.com>
...
Thread-Index: AAAAAAAAAAAAAAA/eeeeeeeeeee/NN    ==
...
Message-ID: <22222222222222222222222222222222@organizacionRemitente.com>
Reply-To: "usuario@organizacionRemitente.com" <usuario@organizacionRemitente.com>
...
received-spf: Pass (protection.outlook.com: domain of organizacion_puente.fr designates
 AAA.BBB.CCC.119 as permitted sender) receiver=protection.outlook.com;
 client-ip=AAA.BBB.CCC.119; helo=wmx-2.organizacion_puente.eu; pr=C
dkim-signature: v=1; a=rsa-sha256; c=simple/simple; d=organizacion_puente.fr; s=default;
    t=1692275977; bh=YbAW...yma5c=;
    h=Date:From:To:Subject:Reply-To:From;
    b=...
...
authentication-results: spf=pass (sender IP is AAA.BBB.CCC.119)
 smtp.mailfrom=organizacion_puente.fr; dkim=pass (signature was verified)
 header.d=organizacion_puente.fr;dmarc=none action=none header.from=organizacionRemitente.com;
...
```

Fig. 5. Cabeceras en un mensaje fraudulento

- Respecto al uso de DKIM, solo hemos encontrado una organización que lo utilizase en sus mensajes, pero sin una política DMARC asociada. Pese a que muchas de las cabeceras de los mensajes originales se conservaban en los mensajes fraudulentos, las cabeceras *DKIM-signature*, en este caso particular, se eliminaban durante la manipulación de los mensajes para evitar validaciones en el destinatario de los mensajes.
- Tampoco se observó la utilización de firmas electrónicas para garantizar la integridad de los mensajes ni la de los documentos adjuntos.

C. *Manipulación en el buzón del destinatario*

En este caso, la manipulación del mensaje se realizaba una vez que ya había concluido la fase de transferencia de correo, cuando éste ya se había almacenado en el buzón del destinatario. Esto es posible tras un robo de credenciales del destinatario o bien mediante una intrusión en el servidor en el que está almacenado el mensaje. A través de un acceso al buzón, se modifica el contenido del mensaje para cambiar la información de pago. Los cambios pueden llegar a ser prácticamente indetectables, de forma que el mensaje fraudulento sea indistinguible del mensaje original salvo por algunas incoherencias que se pueden detectar tras un análisis de cabeceras. Por ejemplo, hemos encontrado un caso en el que la cabecera *Authentication-results* reflejaba validaciones satisfactorias de DKIM, sin la presencia de un *DKIM-signature*. Es decir, durante la manipulación del mensaje se eliminó esa cabecera para evitar la comprobación de la integridad del mensaje.

IV. Discusión

Situaciones como la descrita se producen por varios motivos sobre los que es necesario reflexionar.

A. *Errores conceptuales sobre los servicios de correo-e*

Existen errores conceptuales acerca de estos servicios, lo que provoca que los usuarios, inconscientemente, usen métodos de comunicación inseguros para cosas que requieren de ciertas garantías de seguridad. Así, pese a que los servicios de correo electrónico se diseñaron hace décadas, algunos de los principios de su diseño se mantienen vigentes a día de hoy. En su concepción original se tomó como referencia el servicio postal y, como en el servicio postal, el envío de los mensajes ordinarios (o las cartas) no puede considerarse fiable ni seguro. Tampoco se diseñaron estos servicios para verificar el origen de los mensajes siendo, la suplantación de identidad de un remitente, un problema ampliamente estudiado desde un plano técnico. Es cierto que se han realizado numerosos avances en la seguridad en los servicios de correo electrónico en los últimos años. Se han diseñado soluciones que van desde el cifrado total de los contenidos de remitente a destinatario, hasta la verificación del origen de los mensajes o la firma de los mismos. Pero el problema está en que estas medidas de seguridad no se aplican de forma total ni de manera

Actas de las XV Jornadas
Actas de las XVII Jornadas
de Ingeniería Telemática
(JITEL 2025),
Cáceres (España),
12-14 de noviembre de 2025.

ISBN

generalizada. Como ya se ha comentado, en los casos analizados hemos encontrado organizaciones con muy distintos grados de implantación de medidas de protección de correo electrónico. Adicionalmente, de poco sirve que una organización sea muy estricta con sus medidas de seguridad si se comunica con otras organizaciones que no lo son.

B. Implantación heterogénea de medidas de seguridad

El grado de implantación de las medidas de protección es diverso por distintos motivos. En primer lugar, no hay un actor global que imponga su implantación a todos los usuarios de este servicio. Las soluciones existen, pero no es obligatorio utilizarlas para que el correo electrónico funcione. En segundo lugar, las empresas destinan a la ciberseguridad recursos de forma muy diversa. Hay empresas que tienen una dimensión que les permite contar con un equipo de técnicos especializados y otras que operan con dificultades económicas constantemente. Las primeras se pueden permitir destinar recursos a la protección de las comunicaciones de correo electrónico, pero las segundas difícilmente. ¿Es exigible a todas las empresas cumplir unos mínimos de seguridad de forma análoga a lo que se hace con las administraciones públicas? Esta pregunta quedaría abierta, aunque podría ser de difícil contestación real. Dado que muchas empresas, sobre todo las más pequeñas, contratan este tipo de servicios a terceros y dado que existe una legislación específica para los proveedores de servicios de internet, ¿es exigible la aplicación de estas medidas de seguridad a los proveedores? ¿Qué alcance tiene lo previsto en el artículo 8 de la Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico? ¿No supone la situación actual una amenaza para la seguridad pública? ¿Por qué se recoge en dicha Ley la obligación de información sobre seguridad y no sobre su implantación? ¿Qué sentido tiene informar sobre las medidas de seguridad cuando el destinatario de la información no entiende lo que se le transmite? ¿Es la responsabilidad civil achacable al proveedor del servicio cuando éste no ha aplicado medidas de protección?

Incluso cuando el servicio de correo se proporciona dentro de unos parámetros de seguridad aceptables, siempre existe el riesgo de que se produzca algún incidente. Un prestador de servicios no puede controlar totalmente las amenazas existentes. Esto está ligado al hecho indiscutible de que los sistemas siempre pueden tener vulnerabilidades no detectadas o corregidas. Es decir, los incidentes de seguridad se van a producir y en algunos casos los atacantes tendrán éxito. Por ello, en la ciberseguridad moderna se trabaja en tres aspectos fundamentales: la prevención, la detección y la recuperación. No obstante, el hecho de que se puede producir un incidente BEC pese a que el proveedor de servicios emplee buenas prácticas en el ámbito de la ciberseguridad, podría limitar el alcance de la responsabilidad civil exigible. Pero si hay una negligencia del proveedor de servicios, éste debe ser penalizado, para lo que son necesarios algunos ajustes legislativos.

En una de las posibilidades materiales que se describían con anterioridad, se comentaba que la suplantación de identidad puede realizarse mediante una dirección de correo similar a la de la dirección original. Esto se puede hacer de diferentes maneras, incluyendo la utilización de nombres de dominio similares o el uso de homoglifos. No es una suplantación de identidad real, sino una aproximación basada en la similitud de las direcciones de correo. Esta situación es difícil de controlar debido a que no se puede limitar la reserva de nombres de dominio en Internet bajo el pretexto de un supuesto parecido con un dominio ya existente. Es cierto que existen precedentes en el ámbito de la protección intelectual, pero siempre resueltos con posterioridad y no de forma preventiva. Por otro lado, no se puede obviar la existencia de servidores *open relay*, tanto por militancia como por vulnerabilidades en los sistemas subyacentes. Pese a que las vulnerabilidades son difíciles de evitar totalmente, en un caso de negligencia grave en el mantenimiento de sistemas, podríamos hablar de que la organización propietaria ha colaborado en el incidente. Sin la utilización de esta infraestructura, el incidente no hubiese sido posible. Esto vendría a reforzar lo ya comentado respecto a la legislación en materia regulatoria y de responsabilidad civil. En el caso de la militancia, cuando se despliegan elementos de infraestructura que carecen de cualquier tipo de control de manera intencionada, se puede estar colaborando en el desarrollo de un delito. Por ello, estos casos quizás deberían contemplarse en el Código Penal, aun reconociendo que su aplicación a nivel global es problemática.

C. Elementos relacionados con el proceso legal

Un último aspecto está relacionado con el proceso legal propiamente dicho. En general, cuando una de las partes afectadas decide interponer una denuncia por los hechos ocurridos es porque se siente estafada, pudiendo esto tener consecuencias en el proceso civil (al asumir que ha sido ésta la parte que ha sufrido la intrusión). En cualquier caso, se presenta una denuncia en la policía y, al no tratarse de algo masivo, la investigación se archiva por falta de autor conocido. Se puede perseguir la transferencia bancaria, pero esto suele terminar en la localización de uno o varios muleros que participaron en el proceso, sin llegar al autor real del ataque. En cuanto al proceso civil, la investigación de los hechos debe enfocarse a determinar lo ya comentado: si la parte demandada está liberada del pago o no. Más allá de lo que ya se ha mencionado, esta es la vía que se sigue normalmente. Los autores de la estafa salen indemnes porque este proceso se olvida de ellos y, en ocasiones, dificulta la obtención de evidencias que podrían esclarecer lo sucedido.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

V. Conclusiones y Trabajos Futuros

Como se puede observar, los incidentes BEC plantean numerosos retos para la seguridad en la actualidad. Se producen frecuentes incidentes de este tipo, con un elevado coste económico y reputacional para todo tipo de organizaciones. Como efecto colateral está el hecho de que la frecuente exposición a ciberamenazas genera una desconfianza total en la población y en las organizaciones en la utilización de medios tecnológicos. Esta desconfianza puede convertir en ineficaces las iniciativas tecnológicas que empresas y administraciones públicas llevan desarrollando durante décadas. Si un ciudadano no se fía de los servicios digitales, de nada sirve la inversión realizada en implementar servicios de administración electrónica. Recuperar la confianza de la ciudadanía es otro de los retos para la seguridad, como ya ha planteado recientemente el Foro Económico Mundial [11].

Otro de los retos es el incremento en la complejidad en la prevención y detección de incidentes. Los criminales evolucionan sus tácticas para superar las barreas tecnológicas desplegadas y pasar desapercibidos hasta que la estafa se materializa finalmente. Esto implica que las técnicas de protección y recuperación deben ser mejoradas. Esta evolución es necesaria tanto en el plano tecnológico, mediante el diseño de nuevas soluciones o la evolución de las existentes, como en el plano organizativo mediante la actualización de los protocolos de actuación y las estrategias corporativas.

Como ya se apuntaba en la sección de discusión, es necesario realizar cambios legales y regulatorios. No es posible que los operadores de servicios de Internet y las empresas que tienen infraestructuras informáticas descuidadas no asuman las consecuencias de una actuación negligente. No se puede pretender que esto se solucione proporcionando información a los usuarios si éstos no son capaces de entender en qué consisten las medidas de seguridad implementadas. Es exigible que los servicios de Internet se presten con las debidas medidas de seguridad. También es necesario realizar algunos cambios legislativos de cara a prevenir este tipo de situaciones, como se ha indicado y a facilitar su resolución con la localización de los estafadores y la recuperación de los fondos estafados. La reciente normativa en materia de facturación electrónica para el sector privado recogida en la Ley "Crea y Crece" [12], puede mejorar esta situación. El problema es que la información de pago no necesariamente se tiene que incorporar en la factura, que esta Ley exime del uso de la factura electrónica a algunos tipos de actividad económica y que no es de aplicación general en transacciones internacionales.

Más allá del impacto global que tienen los incidentes BEC, la minimización de estos incidentes y de sus consecuencias, así como la investigación de las estafas cometidas gracias a estos incidentes requiere de un refuerzo en la cooperación internacional. Sobre todo, en el ámbito de la investigación de ciberdelitos puesto que los estafadores actúan sin fronteras. También es necesaria la colaboración en la elaboración de ese marco legislativo actualizado que permita sancionar algunas conductas específicas y mejorar los procedimientos actuales.

A nivel interno, una organización debería intentar mejorar los siguientes aspectos:

- Trabajar en la concienciación y en la formación de sus empleados en materia de ciberseguridad.
- Implantar o exigir a sus proveedores de servicios TIC la implantación de los mecanismos de seguridad debidos: SPF, DKIM, DMARC y ARC, cuando sea de aplicación, así como sistemas avanzados de detección y prevención. Asimismo, exigir, dentro de lo posible, la implantación de estas medidas a sus clientes y proveedores.
- Trabajar en la adopción de estándares de seguridad, al amparo de la normativa nacional y europea.

Además de seguir estudiando estos casos, así como la evolución en las técnicas empleadas por los estafadores, se plantea el análisis futuro de otros tipos de situaciones BEC, gracias al acceso a casos reales en el ámbito de nuestra actividad forense y pericial.

Agradecimientos

Este trabajo ha sido financiado por la Cátedra Telecable de Redes de Nueva Generación de la Universidad de Oviedo y el Ministerio de Ciencia, Innovación y Universidades, la Agencia y del Fondo Europeo de Desarrollo Regional (Proyecto PID2024-155581OB-C22 financiado por MICIU /AEI /10.13039/501100011033 / FEDER, UE).

Referencias

[1] Internet Crime Complaint Center, FBI (2024) "2023 Internet Crime Report", [Online] Disponible en https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

[2] Fiscalía General de Estado (2024) "Memoria de la Fiscalía General de Estado (Ejercicio 2023)", [Online] Disponible en https://www.fiscal.es/memorias/memoria2024/FISCALIA_SITE

[3] Ministerio del Interior (2024) "Informe sobre la cibercriminalidad en España, 2023", [Online] Disponible en https://www.interior.gob.es/opencms/export/sites/default/.galleries/galeria-de-prensa/documentos-y-multimedia/balances-e-informes/2023/Informe-Cibercriminalidad_2023.pdf

[4] IBM (2024) "Cost of a Data Breach Report 2024", [Online] Disponible en https://www.ibm.com/reports/data-breach

[5] INCIBE (2024) "La DGT no está mandando SMS para notificar recordatorios de multas pendientes de pago", [Online] Disponible https://www.incibe.es/ciudadania/avisos/la-dgt-no-esta-mandando-sms-para-notificar-recordatorios-de-multas-pendientes-de

[6] B. L.-Bergmans (2024) "What Is Business Email Compromise (BEC)?", Cybersecurity 101, Crowdstrike, [Online] Disponible en https://www.crowdstrike.com/cybersecurity-101/business-email-compromise-bec/

[7] C. Cross y R. Gillett. (2020) "Exploiting trust for financial gain: An overview of business email compromise (BEC) fraud", en Journal of Financial Crime, 27(3), pp. 871-884.

[8] A. Cidon, L. Gavish, et al (2019) "High Precision Detection of Business Email Compromise", en Actas del 28th USENIX Security Symposium, Santa Clara, CA, USA.

[9] Microsoft (2025) "Alert classification for suspicious email forwarding activity", [Online] Disponible en https://learn.microsoft.com/en-us/defender-xdr/alert-grading-playbook-email-forwarding

[10] Mitre, "Email Collection", ID T1114, 2025 [Online] Disponible en https://attack.mitre.org/versions/v17/techniques/T1114/

[11] Foro Económico Mundial – Foro de Davos (2023), "Trust in tech has eroded: Here are 3 ways to rebuild it", [Online] Disponible en: https://www.weforum.org/agenda/2023/01/davos23-trust-in-tech-3-ways-rebuild-it/

[12] (2022) "Ley 18/2022, de 28 de septiembre, de creación y crecimiento de empresas.", BOE, nº 234, pp. 133603-133675.

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Hacia una Arquitectura para Apoyar el Aprendizaje Ubicuo en Múltiples Dominios Utilizando Datos Abiertos Enlazados

Pablo García-Zarza, Miguel L. Bote-Lorenzo, Juan I. Asensio-Pérez,
Eduardo Gómez-Sánchez, Guillermo Vega-Gorgojo
Dept. Teoría de la Señal, Comunicaciones e Ingeniería Telemática, E.T.S.I. de Telecomunicación, Univ. de Valladolid
Campus Miguel Delibes, Paseo de Belén 15, 47011 Valladolid
{pablogz@gsic, [migbot, juaase, edugom, guiveg]@tel}.uva.es

El aprendizaje ubicuo (*ubiquitous learning*, UL) puede tener lugar en cualquier lugar y momento. Ofrece ventajas como el acceso remoto a los materiales educativos y su adaptación a los estudiantes. A menudo los profesores quieren diseñar situaciones de UL que implican la creación y gestión de cosas espaciales (lugares de interés desde una perspectiva educativa) y tareas de aprendizaje. Este trabajo no es trivial ya que requiere recopilar datos como las coordenadas de las cosas espaciales y enlazar las tareas de aprendizaje a esas ubicaciones. Afortunadamente, los repositorios de datos abiertos, en particular los Datos Abiertos Enlazados (LOD), pueden utilizarse como fuentes de información para situaciones de UL. Esto facilita el trabajo de autoría de los profesores. Por ello, existen aplicaciones que utilizan LOD para apoyar el UL en dominios específicos. Sin embargo, el diseño de este tipo de aplicaciones basadas en LOD no es sencillo. Además, a los profesores les podría resultar útil permitir utilizarlas en diferentes contextos de aprendizaje, incluyendo otras áreas geográficas o nuevos temas educativos. Con este objetivo, en este póster se propone una arquitectura software que simplifica la puesta en marcha de situaciones de UL mediante el uso de LOD. Esta arquitectura proviene del análisis de cuatro aplicaciones de UL. La arquitectura consta de componentes que controlan la gestión de los recursos educativos, las acciones de los usuarios y el análisis de estas acciones. Además, se ha llevado a cabo una implementación de la arquitectura a través de una aplicación utilizada por profesores y estudiantes en una situación de aprendizaje real.

***Palabras Clave*—servicios de apoyo al aprendizaje, arquitectura orientada a servicios, aprendizaje ubicuo, web semántica, LOD**

I. INTRODUCCIÓN

Las telecomunicaciones y la educación se han vuelto cada vez más interdependientes, especialmente con el auge del aprendizaje digital. Esta tendencia fue la que llevó a la Unión Internacional de Telecomunicaciones (*International Telecommunication Union*, ITU) a elaborar un informe en 2012 [1] sobre el estado de la estandarización de las tecnologías de apoyo al aprendizaje, con la participación de diversas organizaciones como el Sector de Normalización de las Telecomunicaciones de la ITU (ITU-T) y el Comité de Estándares de Tecnología Educativa del IEEE. El informe destacó la necesidad de nuevas normas para abordar la «desconexión entre el desarrollo y la implementación de tecnologías educativas» que era evidente en aquel momento. Desde entonces, la ITU-T ha publicado numerosos estándares oficiales relacionados con la tecnología educativa. Estos estándares abordan cuestiones como la especificación de requisitos para los servicios de apoyo al aprendizaje y la definición de arquitecturas de referencia para servicios de aprendizaje que puedan dar soporte a diversos paradigmas educativos, incluido el aprendizaje ubicuo (por ejemplo, [2], [3]).

El aprendizaje ubicuo (*ubiquitous learning*, UL) se define como aquel que, apoyado por la tecnología, puede tener lugar en cualquier lugar y momento a través de múltiples espacios [4], [5]. Este enfoque de aprendizaje difumina las distinciones entre las situaciones de aprendizaje que se producen dentro y fuera del aula gracias a la popularización y el uso de las tecnologías móviles. Bomsdorf [6] destaca los principales beneficios del UL para los estudiantes. Los profesores que deseen que sus estudiantes se aprovechen del UL tendrán que crear los recursos educativos necesarios. Estos recursos consistirán en las cosas espaciales (*spatial things*, ST) en las que quieran desarrollar la situación educativa (lugares de interés para el dominio como catedrales en el dominio del patrimonio cultural, árboles en el dominio forestal, etc.), las tareas de aprendizaje asociadas a estos lugares y el orden en el que los estudiantes deberán realizar las tareas. La autoría de estos recursos no es trivial [7] y puede tener un impacto significativo en la carga de trabajo de los profesores.

Existen muchas aplicaciones de UL que tienen como

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

objetivo apoyar a los profesores. Para reducir la carga de los profesores algunas aplicaciones utilizan datos abiertos (*Open Data*, OD) para facilitar la puesta en marcha de situaciones de UL. Los Datos Abiertos Enlazados [8] (*Linked Open Data*, LOD) destacan para ser utilizados en UL. Esto es debido al gran volumen de datos de distintos dominios que se almacenan como LOD. La creciente relevancia de LOD puede observarse en Wikidata –el repositorio LOD más importante en la actualidad– que ha pasado de almacenar unos 15 millones de entidades en agosto de 2014 a más de 117 en junio de 2025[1]. Además, otra de las características de LOD son las relaciones que se pueden establecer entre repositorios de distintas organizaciones ya que permiten la navegabilidad entre recursos. Esto resulta útil para la creación de situaciones de UL y facilita la reutilización de datos. Por ejemplo, un repositorio LOD que almacena información sobre árboles singulares de un país puede vincular sus entidades con la especie arbórea correspondiente de Wikidata. Esto permite que una aplicación que utilice dicho repositorio de árboles singulares obtenga datos adicionales que ofrecer a los profesores cuando diseñen la situación de aprendizaje.

Las aplicaciones que ofrecen un mayor nivel de apoyo a los docentes pueden ser complicadas de diseñar y desarrollar. En este artículo proponemos una arquitectura software para facilitar el diseño de aplicaciones de apoyo al aprendizaje ubicuo multidominio que utilicen LOD. Se han analizado las funcionalidades ofrecidas por cuatro aplicaciones de UL que utilizan distintos tipos de datos para encontrar elementos comunes que se reflejen en la arquitectura. Un elemento común a todas estas aplicaciones es el uso de un mapa para visualizar ST y la creación de recursos de aprendizaje. Además, en algunos casos es posible realizar un seguimiento de los estudiantes mientras completan las tareas. Para comprobar la validez de la propuesta, se ha implementado un prototipo basado en la arquitectura que ha sido utilizado por dos profesores y sus estudiantes en una situación de aprendizaje auténtica.

El resto del documento se organiza de la siguiente forma. En la sección II se describen las principales funciones de las aplicaciones de UL estudiadas. La arquitectura propuesta se presenta y describe en la sección III. En la sección IV se describe el prototipo basado en la arquitectura propuesta. Finalmente, en la sección V se discuten las conclusiones y el trabajo futuro.

II. ANÁLISIS DE APLICACIONES DE APRENDIZAJE UBICUO

El análisis de esta sección tiene como objetivo extraer los componentes que formarían parte de una arquitectura software que permita apoyar la creación de aplicaciones de UL para distintos dominios. Una parte de las aplicaciones de UL se limita a hacer uso de datos propios lo que puede dar lugar a los problemas mencionados en la sección I. SituLearn[2] es una aplicación representativa de este tipo ya que permite compartir recursos educativos entre docentes dentro de su propia plataforma, pero actúa como un silo de datos para el resto de aplicaciones de UL. En cambio, MeLOD [7], EducaWood[3] y CHEST [9] obtienen parte de sus datos de repositorios abiertos. Además, permiten la reutilización de sus recursos puesto que los publican como LOD. A continuación, se analizan las principales características de las aplicaciones mencionadas y que se muestran resumidas en la Tabla I.

SituLearn es un conjunto de aplicaciones de código abierto. SituLearn-Editor (herramienta de autoría) permite la creación de situaciones de UL de distintos dominios (como historia o geografía) compuestas por ST y tareas de aprendizaje. Opcinalmente, estas situaciones de aprendizaje pueden ser compartidas por los usuarios del SituLearn-Editor. Las tareas de aprendizaje disponibles son preguntas de opción múltiple (*Multiple Choice Question*, MCQ), MCQ con imagen y texto (una sola palabra). Los tres tipos de tareas permiten ofrecer realimentación a los estudiantes. Los datos generados por los profesores se almacenan en su base de datos privada. Los estudiantes pueden acceder a estos recursos y realizar las situaciones de UL a través de SituLearn-Player. Mientras los estudiantes completan las tareas sus profesores pueden hacer un seguimiento de su progreso utilizando SituLearn-Monitor.

MeLOD es una aplicación para apoyar situaciones de UL que tengan lugar al visitar una ciudad. Los datos de las ST se recuperan de repositorios LOD (como DBpedia). Los usuarios pueden personalizar los tipos de ST que desean recibir. Además, pueden votar (mediante valoraciones y comentarios) los ST que visitan. Las preferencias y los datos generados por los usuarios al utilizar MeLOD (por ejemplo, al visitar un ST) se almacenan como LOD. Usando su panel de control se pueden analizar estos datos.

EducaWood es una aplicación diseñada para su uso en educación forestal. Los usuarios pueden crear anotaciones de árboles mediante el uso de formularios. Aparte de estas anotaciones, EducaWood recupera y muestra información de repositorios LOD externos (como Wikidata). Las anotaciones pueden ser completadas de forma colaborativa por varios usuarios (por ejemplo, un usuario indica las coordenadas del árbol y otro determina su especie y adjunta una fotografía). EducaWood solo admite la tarea de anotar árboles por el momento. Los estudiantes pueden descargar la información disponible en EducaWood.

CHEST pretende apoyar a los profesores en la creación de situaciones de UL del ámbito del patrimonio cultural. Los datos sobre las ST se recuperan de fuentes de OD (OpenStreetMap) y de repositorios LOD (como Wikidata). Los profesores pueden crear nuevas ST (basándose en datos que CHEST recupera de LOD). Además de las ST, los profesores pueden crear distintos tipos de tareas de aprendizaje (fotografía, preguntas de opción múltiple, texto, etc.) asociadas a estos lugares. Por último, los profesores pueden utilizar CHEST para crear itinerarios compuestos por ST y tareas de aprendizaje. Todos estos datos se almacenan como LOD. Los estudiantes pueden seguir los itinerarios a través de la misma aplicación.

[1] https://www.wikidata.org/wiki/Wikidata:Statistics
[2] https://situlearn.univ-lemans.fr/
[3] https://educawood.gsic.uva.es/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
RESUMEN DE LAS CARACTERÍSTICAS PRINCIPALES DE LAS APLICACIONES DE APRENDIZAJE UBICUO ANALIZADAS.

	SituLearn	**MeLOD**	**EducaWood**	**CHEST**
Dominio	Ciencias de la Tierra, Historia, Geografía, Educación Física	Patrimonio cultural	Educación forestal	Patrimonio cultural
Autoría del usuario	ST, tareas, itinerarios		ST	ST, tareas, itinerarios
Tipos de tareas	MCQ, imagen con MCQ, texto	Valoración y comentarios	Anotación de árboles	MCQ, texto, fotografía (múltiple)
Fuentes de datos externas		DBpedia, Europeana, GeoName	Wikidata, IFN	OpenStreetMap, Wikidata, DBpedia
Publicación de datos	Datos privados	LOD	LOD	LOD
Monitorización	Sí	Sí		

III. HACIA UNA ARQUITECTURA SOFTWARE PARA EL APRENDIZAJE UBICUO

Los principales componentes de la arquitectura software propuesta para posibilitar la puesta en marcha de aplicaciones de UL en múltiples dominios utilizando OD se muestran en la Fig. 1. Estos componentes se derivan del análisis realizado en la sección II y de la experiencia adquirida diseñando e implementando aplicaciones de UL.

Los componentes **gestor de ST**, **gestor de tareas de aprendizaje** y **gestor de itinerarios** serían los responsables de responder a las solicitudes relacionadas con los recursos de aprendizaje. Estos módulos responden al hecho de que en todas las aplicaciones revisadas en la sección II se definen ST. A estas ST se les suele asociar tareas de aprendizaje. Dentro de estos módulos se deberá gestionar una memoria caché que permita reducir el número de consultas que se realicen a repositorios externos. Estas memorias también buscan reducir los tiempos de respuesta. En el módulo del gestor de ST se deberán incorporar adaptadores para que sea posible recuperar información de múltiples repositorios (locales o externos y que puedan gestionar OD y LOD). Además, tanto CHEST como SituLearn ofrecen a los profesores la posibilidad de agregar situaciones de UL complejas mediante la combinación de tareas y ST. Este módulo deberá hacer consultas a los gestores de ST y tareas de aprendizaje para conseguir sus recursos. Dado que la aplicación de UL deberá permitir que los estudiantes completen las tareas propuestas (como en SituLearn, MeLOD, EducaWood y CHEST), el gestor de tareas de aprendizaje deberá administrar sus respuestas. Estos datos se deberán incorporar al perfil del estudiante en la base de datos privada de la aplicación.

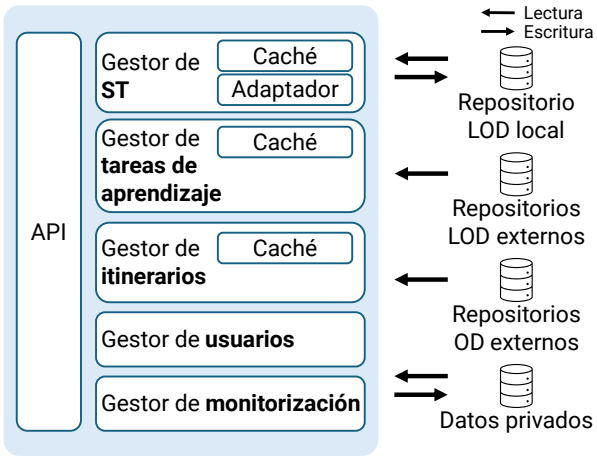


Fig. 1. Arquitectura propuesta para aplicaciones de UL.

El **gestor de usuarios** controla la identificación de los usuarios y los roles que pueden asumir. Con estos roles se controlarán las acciones que puede realizar cada usuario. Por ejemplo, mientras que en CHEST y SituLearn solamente los usuarios con rol de docente pueden crear nuevos recursos de aprendizaje, en EducaWood todos los usuarios registrados pueden añadir nueva información.

El **gestor de monitorización** es el responsable de supervisar las acciones realizadas por los usuarios. Puede ser utilizado por los profesores para hacer un seguimiento de las acciones que realizan sus estudiantes. Así, podrán consultar las tareas que los estudiantes han completado, los lugares que han visitado, sus preferencias, etc.

El medio para acceder a estos componentes será a través de una API. Gracias a ello, cada componente de esta arquitectura se puede implementar, parcial o completamente, en distintas partes de la aplicación software. Por ejemplo, los adaptadores de las ST solo serán necesarios en la aplicación software donde se realicen peticiones a las fuentes de información externa.

IV. NEST: UN EJEMPLO DE IMPLEMENTACIÓN DE LA ARQUITECTURA

NEST[4] es un prototipo diseñado para ayudar a profesores de todo el mundo a poner en marcha situaciones de UL en el dominio del aprendizaje forestal en medios urbanos y vida activa. Esta aplicación se basa en la arquitectura propuesta en la sección III. NEST hace uso del gestor de ST para recuperar ST de repositorios de OD (como OpenStreetMap) y LOD (Wikidata y su propio repositorio). Además, los profesores pueden crear otras ST que no se encuentren en la aplicación, tareas de aprendizaje e itinerarios haciendo uso, de forma transparente, del gestor correspondiente en cada caso. Los estudiantes pueden seguir los itinerarios diseñados por los profesores y resolver las tareas que les propongan. El prototipo NEST fue utilizado por dos profesores. Estos profesores crearon 49 ST, 146 tareas y 5 itinerarios. En una entrevista realizada a los profesores una vez que finalizaron la actividad con

[4]https://nest.gsic.uva.es

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

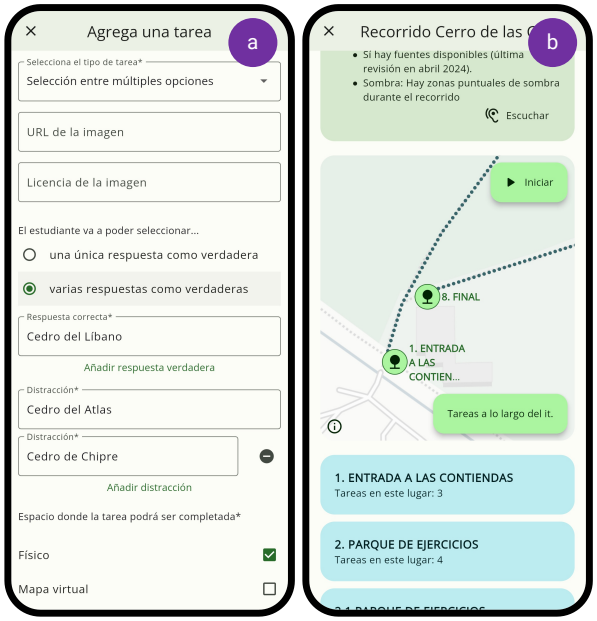


Fig. 2. Capturas de pantalla de NEST. (a) Creación de una tarea de aprendizaje. (b) Información de un itinerario.

NEST indicaron que «*NEST es intuitiva*» y «*la creación de los recursos de aprendizaje ubicuo es sencilla*». Según estos profesores, NEST «*fomenta la autonomía de los participantes*» y destacan su «*sistema colaborativo*». 22 de sus estudiantes completaron más de 1000 tareas.

NEST fue diseñada como una aplicación cliente–servidor. El cliente (Fig. 2) fue desarrollado utilizando Flutter[5] para poder crear aplicaciones con un único código para múltiples plataformas (como web, Android e iOS). Sus interfaces se basan en mapas y formularios. Dependiendo del rol del usuario (profesor, estudiante o usuario sin registrar) podrá realizar unas acciones u otras. El servidor de NEST fue desarrollado usando Node.js[6] y, entre otras funciones, se encarga de gestionar las consultas a los diferentes repositorios LOD. Tanto en el cliente como en el servidor de NEST contienen partes de los gestores de la arquitectura propuesta.

V. CONCLUSIONES Y TRABAJO FUTURO

La arquitectura software que se ha propuesto es un primer paso hacia la creación de una aplicación genérica de apoyo al UL que podría personalizarse para diferentes dominios educativos. Esta aplicación podría hacer uso de múltiples repositorios de OD contribuyendo a reducir el esfuerzo de los docentes al crear y poner en marcha situaciones de UL. Con aplicaciones construidas sobre esta arquitectura, los docentes podrán crear tareas de aprendizaje vinculadas a ST y combinarlas en itinerarios (siendo apoyados en el uso de LOD). Además, podrán monitorizar a los estudiantes mientras realizan sus actividades.

[5]https://flutter.dev/
[6]https://nodejs.org/

Sin embargo, es necesario dar unos pasos previos para conocer diferentes aspectos de la actividad que los docentes desean desarrollar. Además, es necesario definir un modelo de datos que respalde este tipo de situaciones y que pueda ser utilizado en múltiples dominios. En este modelo de datos deben aparecer los conceptos que se han considerado en la arquitectura propuesta. Todos estos esfuerzos se deberán traducir en un proceso que defina los pasos y actores necesarios para realizar la personalización entre profesores e ingenieros de las aplicaciones de UL que utilicen LOD implementadas a partir de la arquitectura propuesta. En [10] se presentó una primera aproximación del proceso formado por varias etapas donde (etapa 1) diferentes ingenieros preparan adaptadores para utilizar datos de múltiples repositorios, (etapa 2) informan a los docentes para que puedan crear tareas de UL y (etapa 3) los docentes ponen en marcha sus situaciones de UL.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente apoyado por los proyectos PID2023-146692OB-C32 y LOD.For.Trees (TED2021-130667B-I00) financiados por MICIU/AEI/10.13039/501100011033 y FEDER, UE y por NextGenerationEU y MCIN/AEI. Pablo García Zarza ha sido financiado con cargo a la convocatoria de contratos predoctorales UVa 2022, cofinanciada por el Banco Santander.

REFERENCIAS

[1] ITU-T, "Standards for technology-enabled learning," ITU-T Technology Watch Report, 2012. [Online]. Available: https://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000180001PDFE.pdf
[2] ITU-T, "Service framework to support web objects based ubiquitous self-directed learning," ITU-T Recommendation Y.2241, 2017.
[3] ITU-T, "Smart farming education service based on u-learning enironment," ITU-T Recommendation Y.2246, 2021.
[4] G.-J. Hwang, C.-C. Tsai, and S. J. H. Yang, "Criteria, strategies and research issues of context-aware ubiquitous learning," *Source: Journal of Educational Technology & Society*, vol. 11, pp. 81–91, 2008.
[5] J. A. Muñoz-Cristóbal, L. P. Prieto, J. I. Asensio-Pérez, A. Martínez-Monés, I. M. Jorrín-Abellán, and Y. Dimitriadis, "Deploying learning designs across physical and web spaces: Making pervasive learning affordable for teachers," *Pervasive and Mobile Computing*, vol. 14, pp. 31–46, 10 2014.
[6] B. Bomsdorf, "Adaptation of Learning Spaces: Supporting Ubiquitous Learning in Higher Distance Education," in *Mobile Computing and Ambient Intelligence: The Challenge of Multimedia*, ser. Dagstuhl Seminar Proceedings (DagSemProc), vol. 5181. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2005, pp. 1–13.
[7] G. Fulantelli, D. Taibi, and M. Arrigo, "A semantic approach to mobile learning analytics," in *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, ser. TEEM '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 287–292.
[8] T. Berners-Lee, "Linked data - design issues," 7 2006. [Online]. Available: https://www.w3.org/DesignIssues/LinkedData
[9] P. García-Zarza, M. L. Bote-Lorenzo, G. Vega-Gorgojo, and J. I. Asensio-Pérez, "CHEST: A Linked Open Data-based Application to Annotate and Carry Out Learning Tasks About Cultural Heritage," in *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption*. Springer International Publishing, 2022, pp. 441–447.
[10] P. García-Zarza, "Supporting Teachers in the Generation of Ubiquitous Learning Situations Across Multiple Domains and Spaces Based on Linked Open Data," in *Doctoral Consortium of the 18th EC-TEL co-located with EC-TEL 2023*, 2023, pp. 62–71. [Online]. Available: https://ceur-ws.org/Vol-3539/paper10.pdf

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Mejora del posicionamiento Wi-Fi RTT mediante clustering y desviación estándar

Nestor Gonzalez Diaz, Enrica Zola, Israel Martin-Escalona
Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya (UPC)
C. Jordi Girona 1-3, 08034 Barcelona.
(nestor.gonzalez.diaz, enrica.zola, israel.martin)@upc.edu

Se propone un sistema de localización indoor para Wi-Fi RTT basado en fingerprinting, dirigido a mejorar la escalabilidad sin sacrificar la precisión. Wi-Fi RTT ofrece medidas de tiempo de vuelo precisas para posicionamiento, pero genera tráfico específico de localización, restando recursos a la red Wi-Fi y limitando por tanto la escalabilidad del sistema. Para mitigarlo se propone aplicar clustering, de modo que cada terminal interrogue solo un subconjunto óptimo de AP, distribuyendo el tráfico de localización. Además, añadiendo la desviación estándar de las muestras RTT como característica podemos mejorar la precisión de la estimación.

***Palabras Clave*—Wi-Fi RTT, clustering, fingerprinting, escalabilidad, precisión**

I. DESCRIPCIÓN Y RESULTADOS

El sistema de posicionamiento propuesto [1] emplea fingerprinting Wi-Fi Round-Trip Time (RTT) en dos fases. En la etapa offline se recopilaron, en múltiples puntos de referencia (RPs), (i) distancias derivadas del RTT obtenido mediante Fine Timing Measurements (FTMs), (ii) su desviación estándar (σ), y (iii) medidas del indicador de la potencia de señal recibida (RSSI) empleando 7 puntos de acceso (APs) compatibles con Wi-Fi RTT. Sobre las medidas RSSI se aplicó K-Means para agrupar las medidas en k clústeres, cada uno de ellos integrando su propio clasificador KNN, y considerando las tres medidas disponibles. Para cada uno de estos clústeres se genera una lista priorizada de APs de acuerdo a su relevancia, empleando para ello el algoritmo *Extra Trees*. Durante la fase online, el terminal infiere primero su clúster utilizando el RSSI, que se obtiene de forma pasiva, sin inyectar tráfico en la red. Una vez fijado el clúster, se determina la posición solicitando tantos RTTs como se requiera, de acuerdo a la lista de AP previamente generada. Eso permite, eventualmente, reducir el tráfico FTM.

A. Principales resultados

En [1], se evaluó la aportación del clustering y el uso de σ a la escalabilidad y la precisión del sistema de localización, planteando escenarios con diferentes densidades de APs. El uso de σ junto al RTT permite mejorar la exactitud, especialmente en escenarios con un número limitado de APs. Usando la información a partir de un único AP se pudo alcanzar un RMSE (raíz del error cuadrático medio) cercano al metro. Por otro lado, la aplicación de clustering permitió una mejor distribución de las tramas FTM entre los APs, resultando en un overhead inferior por AP y permitiendo así mejorar la escalabilidad del sistema de localización. El escenario óptimo se alcanzó con 5 clústeres, lo que supone alrededor de 1 AP/clúster. Bajo estas condiciones, el sistema alcanzó precisiones submétricas (RMSE $\approx$ 0,5 m) con una capacidad de localizar hasta 107 usuarios de forma simultánea, es decir, unas 3,5 veces más que usando multilateración (30 usuarios).

El sistema propuesto aprovecha la infraestructura Wi-Fi existente y evita hardware adicional, resultando idóneo para hospitales, aeropuertos o fábricas donde los APs son escasos pero la demanda de localización es alta. Líneas de investigación futuras contemplan evaluar una gama más amplia de dispositivos para atajar el error de estimación del RTT entre fabricantes, así como optimizar el clustering y la selección de características para disminuir la complejidad computacional.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Generalitat de Catalunya con la subvención 2021-SGR-00594.

REFERENCIAS

[1] N. Gonzalez Diaz, E. Zola, and I. Martin-Escalona, "On the Integration of Standard Deviation and Clustering to Promote Scalable and Precise Wi-Fi Round-Trip Time Positioning," Technologies, vol. 12, no. 10, article 172, 2024, DOI: 10.3390/technologies12100172.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Autoorganización de servicios inteligentes basada en Grafos

Sandra Machado-Zurita, Luis Cruz-Piris, Enrique de la Hoz, Ivan Marsa-Maestre
Departamento de Automática,
Universidad de Alcalá
Campus Universitario. Ctra. Madrid-Barcelona, Km. 33,600. 28805 Alcalá de Henares (Madrid).
sandra.machado@uah.es, luis.cruz@uah.es, enrique.delahoz@uah.es, ivan.marsa@uah.es

Resumen

La autoorganización es un fenómeno que aparece de forma espontánea en la naturaleza. Aplicada a redes de comunicaciones tiene entre sus objetivos minimizar la sobrecarga debida a la configuración masiva de dispositivos, desarrollar protocolos que faciliten el funcionamiento de la red, o generar nuevos tipos de red con un enfoque principalmente descentralizado. Presenta retos como la necesidad de lograr coordinación local y entre sus componentes. Las estructuras basadas en grafos proporcionan nuevas características frente a enfoques tradicionales. En este trabajo se presentan los primeros trabajos realizados en la tesis doctoral titulada: Autoorganización de servicios inteligentes basada en Grafos. Se muestra el sistema implementado para automatizar la generación automática de simulaciones de redes de comunicación, que permita la generación masiva de datos para poder ser utilizado en futuras fases del proyecto. Además, se ha realizado una primera aproximación que ha permitido validar tanto los datos generados, como su utilidad para realizar predicciones sobre el funcionamiento de la red utilizando técnicas de machine learning.

***Palabras Clave*—autoorganización, servicios inteligentes, grafos, redes autoorganizadas, GNN.**

I. INTRODUCCIÓN

La naturaleza ofrece ejemplos de autoorganización: bancos de peces, rutas óptimas de hormigas o la sincronización de luciérnagas [1]. Los sistemas autoorganizados se definen por estructura y funcionalidad: la estructura describe el rol de cada entidad y sus interacciones (animal, dispositivo o agente), y la funcionalidad determina el objetivo global. Por ejemplo, los bancos de peces siguen reglas simples locales para protegerse de depredadores mediante interacción grupal.

Las redes autoorganizadas (SON por sus siglas en inglés Self-organized Networking) en 5G mejoran la capacidad de tráfico, reducen latencias y aumentan la fiabilidad [2]. Están compuestas por autooptimización (caché, encaminamiento, balanceo), autoconfiguración (ajuste automático de parámetros) y autoreparación (detección y recuperación de fallos). Su éxito depende de algoritmos de optimización distribuidos, cruciales en entornos heterogéneos donde cada usuario y servicio requieren gestión local.

Al aplicar optimización distribuida a escenarios reales surgen desafíos [3]:

- La gran escala del problema impide la búsqueda exhaustiva de óptimos.
- Las interdependencias entre elementos dificultan alcanzar el objetivo global o social.
- La presencia de agentes autónomos con intereses propios puede reducir la eficiencia de las soluciones sociales.

Estos sistemas requieren coordinación local limitada, permitiendo soluciones distribuidas de comunicación, computación o reparto energético. Son escalables, adaptables y resilientes [4]. Durante el desarrollo de esta tesis se busca extender la autoorganización a todas las capas del sistema, decidiendo cuándo recolectar datos, cómo procesarlos y cómo transmitirlos, para garantizar la funcionalidad requerida por el servicio.

En la Sección II se recoge brevemente el estado del arte sobre trabajos previos sobre esta temática. En la Sección III mostraremos los primeros desarrollos realizados para obtener una herramienta capaz de automatizar la generación y simulación de escenarios de redes inalámbricas, que permitan la generación masiva de datos. También probaremos cómo estos datos pueden ser utilizados en sistemas conocidos de machine learning. En la Sección IV se realiza una primera prueba de concepto donde se muestran los resultados iniciales. Por último, en la Sección V, estableceremos cómo estos desarrollos aportan una base para las siguientes fases de la tesis.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

II. REDES AUTOORGANIZADAS APLICADAS A LA OPTIMIZACIÓN

El estado del arte proporciona técnicas para abordar estos desafíos de forma parcial [5], [6], [7] pero ningún enfoque los soluciona completamente hasta la fecha. Las técnicas de optimización centralizada han demostrado ser buenas para encontrar soluciones eficientes en problemas de gran escala, a menudo usando técnicas de búsqueda basadas en metaheurísticas. Sin embargo, no trabajan bien con el comportamiento autónomo y competitivo, y por lo tanto necesitan mecanismos que aseguren la adherencia de los individuos a la solución óptima. La optimización distribuida, la teoría de juegos y las técnicas de negociación automática, por el contrario, manejan adecuadamente los conflictos de intereses, pero no escalan bien con el tamaño del problema y las interdependencias.

En redes de comunicación [8], se relaciona la asignación de canales Wi-Fi con el coloreado de grafos, proponiendo técnicas como DSATUR, coloreados on-line y T-colouring; se implementa DSATUR considerando canales ortogonales y sin incluir clientes. El problema general de coloreado con interferencia mínima y canales limitados es NP-duro [9]. En Inteligencia Artificial, los grafos permiten modelar relaciones ocultas, como en redes neuronales basadas en grafos (GNN, por sus siglas en inglés Graph Neural Network), útiles para extraer conocimiento estructurado de datos complejos.

En [10] se propone un marco basado en una red neuronal convolucional de grafos secuencial con compuertas diseñado para la detección de botnets en entornos de redes IoT. Al aprovechar las capacidades de las GNNs para representar el tráfico de red como estructuras de grafos, y manejar la dinámica secuencial de los ataques de botnets, se mejora la precisión y eficiencia en la detección de ataques. El modelo que proponen utiliza una estructura de grafo con múltiples aristas y marcas de tiempo que representa los flujos de tráfico de red proporcionados por los atributos de las aristas, lo cual permite discernir la dinámica temporal de la red de bots. Esta característica única permite al modelo descubrir patrones ocultos y dependencias en los datos, lo que conduce a una detección más precisa y oportuna de actividades sofisticadas de botnets.

Otra investigación donde se aborda el tema, enfocado a soluciones para IoT en el campo de la ciberseguridad, es [11]. Con el objetivo de detectar intrusiones en infraestructuras de comunicación del IoT, se modela la red como un grafo dinámico que evoluciona con el tráfico de red en tiempo real. Esta propuesta incorpora una ponderación adaptativa de los enlaces basada en métricas como la frecuencia de comunicación, el volumen de datos transmitidos y el tipo de protocolo, lo que permite diferenciar comportamientos normales de interacciones potencialmente maliciosas. El modelo optimiza la representación de las relaciones entre dispositivos mediante un mecanismo de agregación de características estructurales y contextuales, mejorando la capacidad del sistema para detectar ataques distribuidos como Denegación de Servicio (DoS por sus siglas en inglés Denial of Service), suplantación de identidad (Spoofing) y ataques Hombre en el Medio (MiTM por sus siglas en inglés Man-in-the-Middle). En pruebas con escenarios de ataques distribuidos, el modelo alcanzó un puntaje F1 de 0.95 y un AUC-ROC de 0.98, superando a los métodos tradicionales basados en aprendizaje automático por más del 7 %. Además, el sistema mostró una reducción del 40 % en la tasa de falsos positivos en comparación con los sistemas de detección de intrusiones (IDS por sus siglas en inglés Intrusion Detection Systems) basados en firmas, lo que lo hace más viable para entornos operativos donde la precisión es crítica. En cuanto a eficiencia computacional, el modelo demostró ser escalable, manteniendo un tiempo de inferencia de 2.5 ms por muestra en grafos de hasta 10000 nodos, lo que permite su implementación en tiempo real sin afectar la latencia del sistema.

III. AUTOORGANIZACIÓN DE SERVICIOS INTELIGENTES BASADA EN GRAFOS

A. Planificación y objetivos

Se plantea investigar el uso de nuevas estructuras de datos basadas en grafos como fundamento para el desarrollo de técnicas inteligentes autoorganizadas, capaces de gestionar y configurar sus dispositivos de forma autónoma. Por tanto, como primer objetivo, se propondrá un modelo que sea lo suficientemente expresivo como para capturar la estructura y dinámica de los sistemas inteligentes.

Posteriormente, se integrará el modelo basado en grafos creado como base para la toma de decisiones en servicios inteligentes autoorganizados. Una vez que el modelo basado en grafos haya capturado adecuadamente las características estructurales y funcionales del sistema, será utilizado como herramienta para las técnicas de inteligencia artificial distribuidas que guiarán el comportamiento autónomo de los dispositivos. Se explorará la viabilidad de aplicar enfoques de optimización tanto centralizados como distribuidos, junto con mecanismos de negociación automática, con el fin de coordinar la toma de decisiones y garantizar una autoorganización eficaz del sistema. Para validar las propuestas se utilizarán escenarios sintéticos y/o reales que permitan demostrar la robustez de las propuestas.

Este documento se elabora en el momento inicial de la presente investigación, por lo que es un trabajo en proceso donde actualmente se está trabajando en la consecución del primer objetivo, primeras herramientas desarrolladas, y resultados iniciales.

B. Sistema de generación y simulación automática de redes inalámbricas.

Es fundamental disponer de herramientas que permitan validar y evaluar los datos que serán utilizados en las etapas posteriores de esta tesis. La dificultad de modificar escenarios reales para comprobar situaciones límites hace necesario el uso de emuladores o simuladores que permitan replicar el comportamiento de los sistemas reales en entorno controlados. Con este propósito se ha desarrollado

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

el sistema mostrado en la Fig. 1, el cual ilustra las fases iniciales del trabajo de investigación. A través de un sistema gestor (desarrollado en Python) se interactúa con cada módulo para permitir, con un conjunto de instrucciones iniciales limitadas (representación de la red, número de puntos de acceso, número de clientes, parámetros de configuración, etc.) ejecutar el proceso completo hasta la obtención y validación de los resultados.

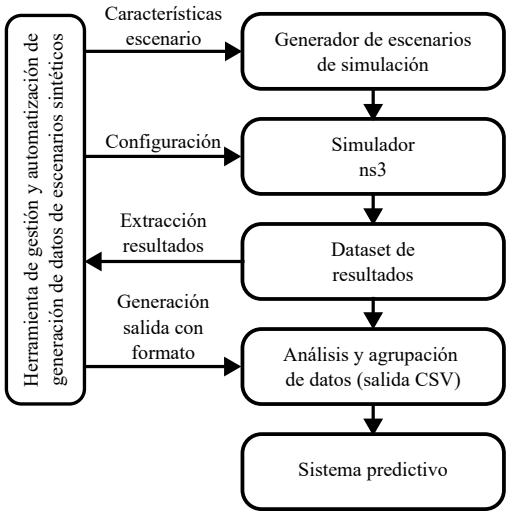


Figura 1. Esquema inicial desarrollado.

Las principales funciones cada módulo son:

- Generador de escenario de simulación: en base a un conjunto limitado de datos, se generan los grafos iniciales que contienen las características básicas del escenario (APs, STAs, posiciones, enlaces, etc.).
- Simulador ns3: a la entrada del simulador se utiliza una definición del escenario en el lenguaje C++. Leyendo los grafos del módulo anterior, se han implementados las funciones necesarias utilizando la biblioteca de Python Ninja para rellenar esta descripción, generar el archivo C++ del escenario, compilar y ejecutar la simulación, y almacenar la salida del simulador.
- Dataset de resultados: el sistema cuenta con una base de datos NoSQL y un directorio de archivo donde se almacenan los resultados obtenidos en cada simulación. A través de consultas a esta base de datos es posible filtrar y extraer la información requerida por la herramienta gestora (Python).
- Generación salida con formato: dependiendo del uso que se le vaya a dar a estos datos, se han desarrollado diferentes script que extraigan y conformen los archivos utilizados en posteriores fases.

Por último, en el ejemplo desarrollo para este artículo, se ha utilizado un sistema predictivo basado en SVM (*support-vector machines*) que recibe los datos adaptados a las necesidades de la herramienta o sistema utilizado. En este caso, dado que se ha empleado el *toolbox* de optimización de Matlab, el formato CSV era el más indicado.

IV. CASO DE USO: REDES WIFI

Una red inalámbrica basada en WiFi (que opere en modo infraestructura) consta de dos elementos principales: Puntos de Acceso (APs) y estaciones (usuarios). Los APs actúan como routers que conectan a los clientes para brindar acceso a Internet. El estándar IEEE 802.11 define las especificaciones de la capa MAC (Media Access Control) y la capa física (PHY) para redes WLAN en las bandas de 2.4, 5 y 60 GHz [12]. El rendimiento depende de la intensidad de la señal transmitida, y los enlaces entre APs y clientes pueden ser de conexión o de interferencia, en función de factores como SNR, potencia o canal de operación.

A. *Experimentación y prueba de concepto*

Utilizando la herramienta mostrada en la Fig. 1, se ha generado un escenario sintético con 26 APs y 500 STAs, que ejecuta 103 simulaciones por grafo: 3 con todos los APs en un mismo canal (1, 6 o 11), 50 con canales ortogonales aleatorios (1, 6 y 11) y 50 con canales entre 1 y 11. Para todos ellos se obtiene y almacenan métricas como Throughput, SNR, Delay, Jitter y paquetes transmitidos/recibidos.

A partir de una descripción general, la herramienta genera de forma automática un conjunto de 17 tipos de grafos personalizados con 26 APs y distintas cantidades de usuarios. Para cada tipo se construyen 10 variantes con posiciones aleatorias, alcanzando un total de 170 escenarios distintos.

Al finalizar el proceso, tenemos un Dataset con datos de un conjunto amplio de variaciones sobre las configuraciones de los APs del sistema y las posiciones de los STAs. Utilizando Matlab y un modelo supervisado de Máquina de Vectores Soporte (SVM) se plantea ser capaces de predecir el Throughput de cada AP al variar las configuraciones de la red inalámbrica.

Los resultados globales obtenidos fueron: RMSE = 1.8492 Mb/s, MAE = 1.2676 Mb/s, R^2 = 0.7922 y correlación = 0.8907. Las gráficas por AP se muestran en Fig. 2 y en la Fig. 3. Estos resultados demuestran errores bajos y una fuerte correlación entre los Throughputs reales y los predichos.

V. CONCLUSIONES

En este trabajo se muestran los primeros pasos realizados en el desarrollo de la tesis doctoral mencionada. Las pruebas y resultados obtenidos en las secciones II y IV demuestran el correcto funcionamiento de la herramienta desarrollada para automatizar la generación de datos sintéticos. Un proceso que antes consistía en tener que programar en C++ un escenario completo, se ha reducido a indicar un conjunto limitado de características del mismo. De igual forma, la herramienta automatiza la ejecución de las simulaciones y la recogida de datos, haciendo que esta tarea se pueda paralelizar y ejecutar en segundo plano.

Además, utilizando un ejemplo sencillo, se muestra como técnicas como SVM permiten estirar el comportamiento de la red sin necesidad de realizar nuevas simulaciones. Esto servirá como apoyo a la siguientes fases de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

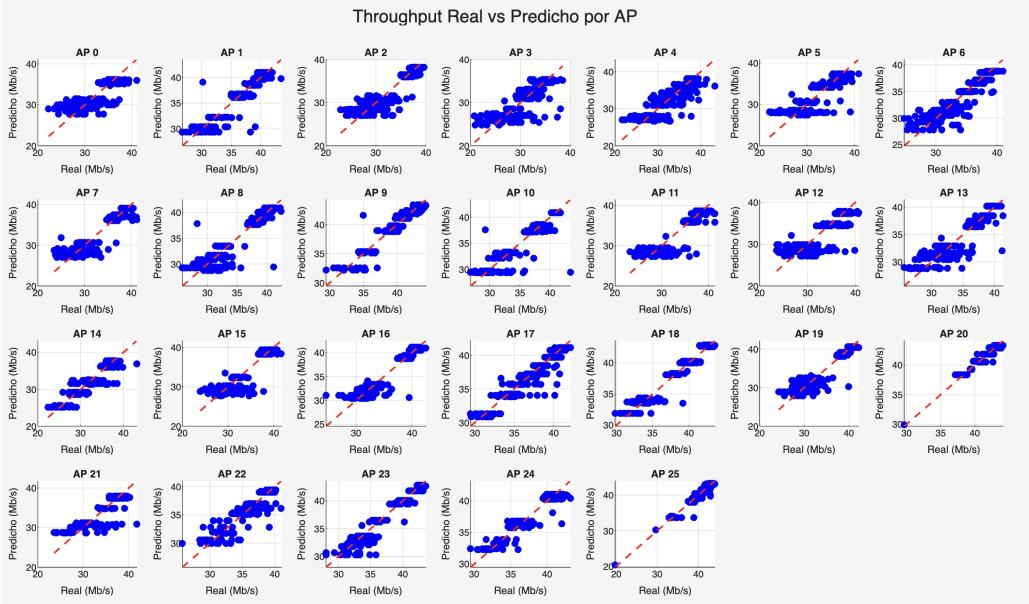


Figura 2. Comparación de Troughputs para cada AP.

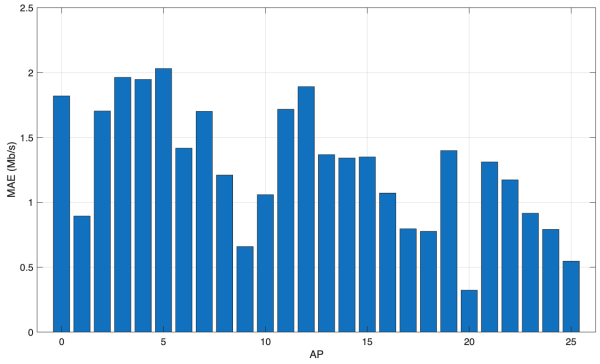


Figura 3. Error medio de Troughputs predichos para cada AP.

la investigación para futuros desarrollos basados en GNNs y otras técnicas que utilicen las características propias de las estructuras de datos basadas en grafos.

AGRADECIMIENTOS

Este artículo es parte del proyecto de investigación SBPLY/23/180225/000160, financiado por la UE a través del FEDER y por la JCCM a través de INNOCAM.

REFERENCIAS

[1] C. Prehofer and C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 78–85, Jul. 2005, conference Name: IEEE Communications Magazine. [Online]. Available: https://ieeexplore.ieee.org/document/1470824

[2] S. Ghosh, T. Dagiuklas, M. Iqbal, and X. Wang, "A Cognitive Routing Framework for Reliable Communication in IoT for Industry 5.0," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5446–5457, Aug. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9674838/

[3] Y. Wang, S. Wang, and L. Wu, "Distributed optimization approaches for emerging power systems operation: A review," *Electric Power Systems Research*, vol. 144, pp. 127–135, Mar. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779616305132

[4] F. Ahmed, "Self-Organization: A Perspective on Applications in the Internet of Things," in *Natural Computing for Unsupervised Learning*, X. Li and K.-C. Wong, Eds. Cham: Springer International Publishing, 2019, pp. 51–64. [Online]. Available: https://doi.org/10.1007/978-3-319-98566-4_3

[5] S. Ahmad and A. H. Mir, "Scalability, Consistency, Reliability and Security in SDN Controllers: A Survey of Diverse SDN Controllers," *Journal of Network and Systems Management*, vol. 29, no. 1, p. 9, Nov. 2020. [Online]. Available: https://doi.org/10.1007/s10922-020-09575-4

[6] G. D'Agostino and A. Scala, Eds., *Networks of Networks: The Last Frontier of Complexity*, ser. Understanding Complex Systems. Cham: Springer International Publishing, 2014. [Online]. Available: https://link.springer.com/10.1007/978-3-319-03518-5

[7] I. Rausch, J. Nauta, P. Simoens, and Y. Khaluf, "Modeling the influence of social feedback on altruism using multi-agent systems," in *Artificial Life Conference Proceedings 32*. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , 2020, pp. 727–735.

[8] J. Riihijarvi, M. Petrova, and P. Mahonen, "Frequency allocation for WLANs using graph colouring techniques," in *Second Annual Conference on Wireless On-demand Network Systems and Services*, Jan. 2005, pp. 216–222. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1383423

[9] D. Orden, I. Marsa-Maestre, J. M. Gimenez-Guzman, E. de la Hoz, and A. Álvarez Suárez, "Spectrum graph coloring to improve Wi-Fi channel assignment in a real-world scenario via edge contraction," *Discrete Applied Mathematics*, vol. 263, pp. 234–243, Jun. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166218X18306528

[10] T. Altaf, X. Wang, W. Ni, G. Yu, R. P. Liu, and R. Braun, "Gnn-based network traffic analysis for the detection of sequential attacks in iot," *Electronics*, vol. 13, no. 12, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/13/12/2274

[11] W. Villegas-Ch, J. Govea, A. Maldonado Navarro, and P. Palacios Játiva, "Intrusion Detection in IoT Networks Using Dynamic Graph Modeling and Graph-Based Neural Networks," *IEEE Access*, vol. 13, pp. 65 356–65 375, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10960408

[12] J. Salazar, "Redes inalámbricas," *Techpedia. Recuperado de: https://core. ac. uk/download/pdf/81581109. pdf*, 2016.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Virtualización de arquitecturas de voto electrónico

Jordi Rodríguez-Coello, Julián Fernández-Navajas, Álvaro Larrey-Gómez, Guillermo Azuara-Guillén, José Luis Salazar-Riaño, María Canales-Compés.
DIEC y DIIS, I3A (Instituto Universitario de Investigación en Ingeniería de Aragón).
Universidad de Zaragoza
Edificio Ada Byron, Campus Río Ebro, Calle María de Luna 1, 50018, Zaragoza, España.
{navajas,alarrey,gazuara,jsalazar,mcanales}@unizar.es

El objetivo de este artículo es mostrar el proceso de virtualizado mediante varias técnicas, de dos sistemas de voto electrónico a distancia. Al virtualizarlos se garantiza una infraestructura más modular, adaptable y fácil de gestionar, permitiendo la separación de servicios, la mejora en seguridad y la ejecución eficiente en distintos entornos sin afectar la estabilidad del sistema. Los dos sistemas están basados en arquitecturas y filosofías distintas: La primera, que llamaremos centralizada, es un servidor de votaciones con una base de datos asociada. La segunda arquitectura, que llamaremos distribuida, consta de dos servidores que deben comunicar entre sí, el servidor de censos y el gestor de urnas. Se virtualizan en máquinas virtuales y también como contenedores, a los que también se les añadirá una capa de orquestado (Kubernetes).

***Palabras Clave*- máquina virtual, contenedor virtual, voto electrónico.**

I. Introducción

En la actualidad, existe una amplia variedad de propuestas para sistemas de votación electrónica a distancia. La característica principal de estos sistemas es que la participación se lleva a cabo con herramientas informáticas, en ubicaciones remotas y la información se transmite por redes telemáticas.

A pesar de que el voto electrónico representa un desafío en términos de seguridad y confianza, el uso de técnicas informáticas en su implementación posibilita la integración de conceptos telemáticos, criptográficos y de teoría de la información, lo que permite introducir características innovadoras. En los últimos 20 años, han surgido numerosas propuestas vinculadas a la votación electrónica, destacando en la actualidad las basadas en tecnología blockchain, cifrado homomórfico y firmas en anillo [1].

En general, las diferentes funcionalidades de un sistema, entre las que se encuentran los sistemas de voto electrónico, pueden desplegarse en varias máquinas tanto reales como virtuales. El uso de máquinas virtuales ha sido un paso clave para replicar y analizar diferentes entornos de desarrollo. Este enfoque permite evaluar configuraciones, realizar pruebas de compatibilidad y modificar el comportamiento de servicios sin afectar al entorno productivo. Además, en situaciones donde puede darse un gran volumen de accesos al sistema, es necesario estar alerta, de forma continua, ante cualquier variación en la situación de demanda de recursos para poderlos gestionar. Es en esta situación cuando el uso de máquinas y contenedores virtuales [2,3] permite agilizar su funcionamiento, aumentando o disminuyendo los recursos disponibles. Para el uso de máquinas virtuales [4], existen diferentes herramientas como VMware (herramienta comercial con soporte empresarial), VirtualBox (de código abierto y orientada a entornos de desarrollo), QEMU (de código abierto y más flexible que las anteriores) o Proxmox (de código abierto y que combina máquinas y contenedores virtuales). Estas herramientas permiten la gestión de máquinas virtuales desplegadas manualmente en diferentes ubicaciones.

Sin embargo, las máquinas virtuales presentan el sistema operativo completo, lo que resulta, en muchos casos, costoso computacionalmente. Por este motivo surgen los contenedores virtuales que comparten el kernel del sistema anfitrión y son mucho más ligeros que las máquinas virtuales. Con el auge de los servicios en la nube, el contenerizado se ha consolidado como una solución esencial para lograr portabilidad, escalabilidad y eficiencia en el desarrollo y despliegue de aplicaciones. Una de las alternativas más extendidas para el uso de contenedores virtuales es Docker. Plataformas como Docker han simplificado este proceso, optimizando entornos nativos de la nube y arquitecturas híbridas. En comparación con la virtualización por máquinas, los contenedores destacan por su flexibilidad y facilidad de distribución, permitiendo instalar aplicaciones en diversas plataformas sin preocuparse por incompatibilidades con el sistema anfitrión. Docker permite, además, la posibilidad de orquestación, que es la capacidad de automatizar la gestión, el despliegue, escalabilidad, monitorización y administración de los recursos de los contenedores en un entorno distribuido.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Existen, para Docker, diferentes alternativas de orquestación, como son Docker Compose [5] o Kubernetes [6].

El artículo, que es un trabajo en proceso, se estructura de la siguiente manera: en el apartado II se presentan las aplicaciones a virtualizar, a continuación, en el apartado III se explican los métodos de virtualización utilizados y en el apartado IV se presentan las conclusiones y líneas futuras de investigación.

II. Arquitecturas de trabajo

Se van a manejar dos arquitecturas de trabajo diferentes para el voto electrónico. La primera de ellas, que llamaremos centralizada, descrita en la Fig. 1, está basada en una prueba de concepto descrita en [7,8]. Consta de dos partes diferenciables: Un servidor de votaciones con una base de datos asociada, y varios clientes con software específico, utilizados tanto por los organizadores de la votación como por los electores. La base de los clientes es un software criptográfico que se comunica con un navegador local mediante WebSocket y éste se comunica con el servidor web de votaciones. Será el servidor web el que se conecte con una base de datos que almacena las votaciones y sus respectivas urnas. Es necesario que el S.O. del cliente soporte el software criptográfico, que puede ser descargado desde el propio servidor web.

La segunda arquitectura, que llamaremos distribuida, descrita en la Fig. 2, consta de cuatro partes diferenciadas: Cliente para organizadores (con funciones de gestión de censos y urnas), Cliente para participantes (con funciones de generación de claves y firma de claves), Servidor de censos y Gestor de urnas. La comunicación entre los diferentes elementos se hace a través de comunicaciones TCP/IP. El Cliente para organizadores sólo se comunica con el servidor de censos, mientras que el Cliente para participantes lo hace tanto con el Servidor de censos como con el Gestor de urnas. Una característica importante es que estos clientes están desarrollados en Python, lo que da muchas facilidades para que se ejecuten en diferentes sistemas operativos.

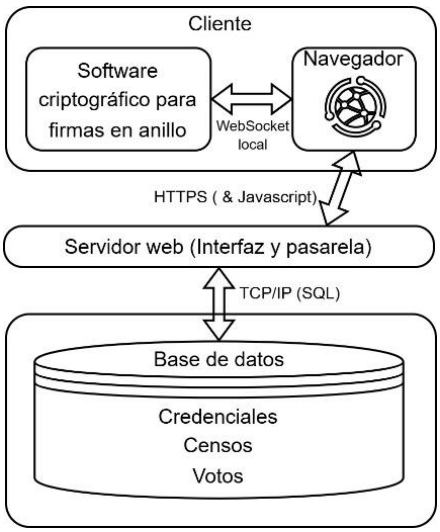


Fig. 1. Diagrama de la arquitectura centralizada

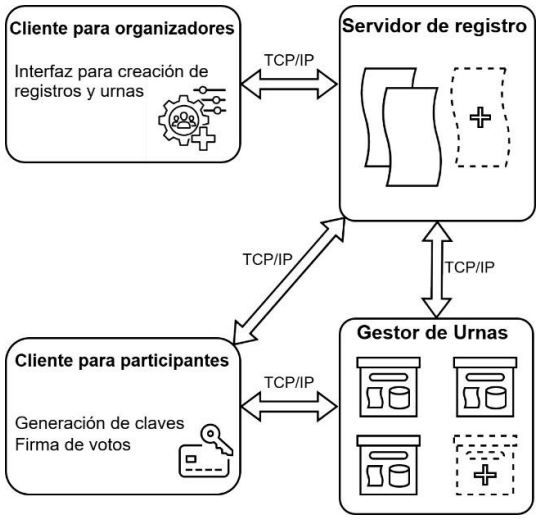


Fig. 2. Diagrama de la arquitectura distribuida

Hay que destacar que cuando se despliega una votación, toda la información relacionada a la votación debe estar contenida en un fichero que denominaremos "urna" y contendrá la información del censo permitido y los certificados. Este fichero deberá ser transferido desde el servidor de censos al servidor de urnas. Una vez terminada la votación, se bloquea la urna y toda la información debe permanecer inalterable. Esta última funcionalidad no está contemplada en el presente trabajo, quedando para futuros desarrollos.

III. Virtualización

Las dos arquitecturas descritas en el apartado anterior pueden desplegarse en diferentes máquinas que pueden ser reales o virtuales. A continuación, se describen las distintas técnicas de virtualizado utilizadas:

A. Virtualización de la aplicación centralizada mediante máquinas virtuales

Este es un paso clave para replicar y analizar el correcto funcionamiento del despliegue como paso previo a su contenerizado. Este enfoque permite evaluar configuraciones, realizar pruebas de compatibilidad y modificar el comportamiento de los servicios para que sean más compatibles con entornos virtualizados. Para esta virtualización se utiliza Proxmox [1], ya que permite gestionar máquinas virtuales y contenedores desde una interfaz web intuitiva, con herramientas avanzadas de gestión de almacenamiento, redes y backup.

Se opta por utilizar de base una distribución de Ubuntu 22.04 debido a que se considera estable y plenamente compatible con el ecosistema Docker, necesario para pasos posteriores de contenerizado.

También se crea una máquina virtual separada para la base de datos (MySQL), en lugar de alojarla junto al servidor web. Esta configuración mejora la flexibilidad y garantiza mayor resiliencia, al mantener la base de datos fuera de los contenedores y protegida ante posibles fallos.

Una vez completado el proceso, la máquina virtual que contiene el servidor puede ser clonada y distribuida.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

B. Implementación centralizada en contenedores

Una vez validado que la aplicación puede ser virtualizada, se pasa al contenerizado. Para ello se usará el motor de virtualizado Docker. Se parte de una imagen (contenedor) base de **Ubuntu 22.04**, sobre la cual se instalan todas las dependencias necesarias para el funcionamiento, pruebas y desarrollo del sistema. Las tablas I y II muestran la amplia gama de paquetes necesarios para que la aplicación centralizada tenga funcionalidad completa:

Tabla I
TABLA DE DEPENDENCIAS

Uso	Paquetes
Herramientas básicas	nano, sudo, curl
Herramientas de red	net-tools, iproute2, ping, tcpdump
Servidor SSH	openssh-server y openssh-client
Herramientas para Python	Python 3.9
Base de datos	MySQL y cliente
Servidor	Servidor Apache
Otros	tar, ufw, xauth

Además, son necesarios varios paquetes de Python, y un módulo para Ubuntu que permite a Apache interpretar aplicaciones Python mediante **WSGI**, esencial para la ejecución de criptografía avanzada desde el servidor web:

Tabla II
TABLA DE DEPENDENCIAS PYTHON

Uso	Paquetes
librerías de seguridad y cifrado	asn1tools, cryptography, rsa
permite a Python conectarse a bases de datos MySQL externas.	mysql-connector-python
módulo necesario para desplegar aplicaciones Python con Apache.	mod_wsgi

También se generan y habilitan los módulos necesarios de Apache, se configura el servicio **SSH** y se copia toda la aplicación al contenedor mediante **COPY**. Es importante asignar correctamente los permisos a Apache y **WSGI** para su correcta interacción.

La configuración de Apache, tanto para servir archivos estáticos como para definir las rutas de ProxyPass y WSGI, se realiza dentro del propio Dockerfile para mantener el control explícito de los pasos de despliegue:

Esta configuración incluye:

- Alias y nombre de servidor
- Rutas de archivos estáticos
- Definición de proxies inversos
- Asociación de scripts **WSGI**
- Definición de logs

Por último, se define un punto de entrada al contenedor, en este caso un script de Python encargado de arrancar el servidor web. Este script se define como comando de inicio usando **CMD**, lo que permite que, en caso necesario, se sobrescriba al lanzar el contenedor con docker run, proporcionando argumentos o comandos alternativos. Las funciones que implementa son:

- Verifica la conexión con la base de datos.
- Inicia los servicios de Apache2 y SSH.
- Ejecuta tail -f /dev/null para mantener un proceso en foreground y evitar que el contenedor se detenga inmediatamente.

Esta primera arquitectura ha sido correctamente contenerizada, lo que permite su despliegue de forma más ligera y eficiente que mediante máquinas virtuales. Además de facilitar su uso en entornos de producción, ofrece una gran versatilidad para entornos de prueba. La amplia gama de paquetes a instalar evidencia la compatibilidad y madurez de esta tecnología de contenerizado.

C. Implementación distribuida desarrollada mediante contenedores

Para la arquitectura no se utilizan máquinas virtuales, sino que se pasa directamente a crear los contenedores, uno para el servidor de registro y otro para el gestor de urnas. Se crean dos imágenes con dos Dockerfile distintos. Como imagen base se utiliza nuevamente **Ubuntu 22.04.**

Se instalan los paquetes fundamentales que requiere la aplicación, cuyas dependencias se detallan en la Tabla III:

Tabla III
TABLA DE DEPENDENCIAS DEL SISTEMA DISTRIBUIDO

Uso	Paquetes
Base	Python 3.12
comunicaciones	FastAPI, Uvicorn, Pydantic, pyOpenSSL
Base de datos	Aiosqlite, databases, SQLite
Criptografía	pycryptodome

El código de las aplicaciones se encuentra en archivos de Python (.py) que actúan como sus propios puntos de entrada, pueden ser lanzados directamente con un CMD en el Dockerfile.

Además, es necesario incluir la base de datos. Sin embargo, esta forma de introducir la base de datos **no es recomendable en entornos productivos**, debido a los siguientes motivos:

- Si el contenedor se elimina o reinicia, **se perderían todos los cambios realizados** en la base de datos.
- **No se podría compartir la base de datos entre varios contenedores**, lo que limita la escalabilidad y la interoperabilidad si no se utilizan técnica de NFS.

La exposición de puertos en cada Dockerfile es esencial para permitir el acceso a los servicios desde la red del host o entre contenedores, especialmente en aplicaciones web asíncronas. El comando **CMD** se utiliza para definir el proceso principal del contenedor, indicando el puerto en el que debe escuchar la aplicación y exponiéndola en todas las interfaces de red (**0.0.0.0**), lo que asegura la correcta recepción de peticiones externas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

IV. CONCLUSIÓN Y LÍNEAS FUTURAS

En el presente trabajo en proceso se virtualizan dos arquitecturas de voto electrónico, tanto mediante máquinas virtuales como con contenedores. Una vez realizada la prueba de concepto se continuará con un análisis de la eficiencia (recursos utilizados, escalado, disponibilidad, resiliencia) y la seguridad.

Como evolución de la segunda implementación, se propone desplegar la aplicación en un entorno Kubernetes, aprovechando su capacidad de orquestación automática, escalado, alta disponibilidad y resiliencia ante fallos. El objetivo principal es garantizar que los servicios puedan distribuirse de forma flexible en diferentes nodos del clúster, permitiendo balancear la carga de trabajo entre los distintos pods para mantener el rendimiento incluso en escenarios de alta demanda.

Para ello, se organizarán los distintos componentes de la aplicación en pods independientes, agrupados en Deployments o despliegues que faciliten su escalado automático y gestión de actualizaciones. Además, se utilizarán servicios de tipo NodePort o ClusterIP para exponerlos de forma controlada y permitir el balanceo de carga interno de Kubernetes.

Dado que la aplicación utiliza SQLite como base de datos, la cual almacena sus datos en un fichero local, **se plantea una mejora esencial**: desacoplar el almacenamiento de datos del ciclo de vida de los contenedores. Para ello, se integrará un sistema de almacenamiento compartido mediante **NFS** (Network File System), montado en los pods a través de Persistent Volume Claims (**PVCs**). Esto permitirá que todos los pods accedan de forma segura y consistente a las bases de datos asociadas a cada votación, evitando la pérdida de datos en caso de reinicio o eliminación de los mismos, y facilitando la persistencia del servicio. Otra posible solución sería montar un volumen externo, permitiendo que los datos persistan fuera del ciclo de vida del contenedor y que puedan ser accedidos por distintos servicios o réplicas. Este planteamiento asegura un entorno más profesional, estable y escalable, alineado con las buenas prácticas de despliegue de microservicios en contenedores.

Para evaluar la **eficiencia de los contenedores y el comportamiento de la aplicación web bajo carga**, se proponen una serie de pruebas que permitirán comparar el **consumo de CPU**, la **capacidad de respuesta** y la **gestión de peticiones concurrentes** durante procesos de votación simulados. Estas pruebas recrearán situaciones de alta concurrencia, donde múltiples usuarios acceden simultáneamente al sistema.

Entre las herramientas recomendadas para realizar estas pruebas se incluyen: **Apache Benchmark (ab)**: para lanzar múltiples peticiones HTTP concurrentes y medir el tiempo de respuesta y el rendimiento de la aplicación web. **Sysbench**: aunque más habitual para bases de datos como MySQL, puede adaptarse para realizar pruebas de estrés en sistemas SQLite indirectamente o complementar la carga del sistema. **docker stats**: para monitorizar en tiempo real el consumo de **CPU, memoria y recursos de red** de los contenedores durante las pruebas.

Estas pruebas permitirán **identificar cuellos de botella**, verificar la **escalabilidad** de los despliegues en Kubernetes y ajustar la configuración de los pods y servicios para garantizar un funcionamiento óptimo en escenarios de alta demanda.

Como los conceptos de eficiencia y seguridad suelen ser antagónicos, todas estas pruebas de análisis de eficiencia deben completarse con un análisis de la estabilidad de la seguridad. Al igual que existen metodologías de análisis de servicios básicos de seguridad (confidencialidad, autenticidad e integridad), deberemos aplicar metodologías especializadas en sistemas de e-voting, cerrando las brechas que pudieran haber surgido tras la inclusión de la arquitectura descentralizada, ya sea mejorando los protocolos criptográficos, o limitando el impacto de la nueva arquitectura en cuanto a complejidad de computación.

AGRADECIMIENTOS

Los investigadores pertenecen al I3A (Instituto Universitario de Investigación en Ingeniería de Aragón, Universidad de Zaragoza). Se ha realizado en el marco de los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation), Proyecto Estratégico C077.23, desarrollado bajo el amparo del convenio entre INCIBE y la Universidad de Zaragoza. NEWLAN (PID2022-136476OB-100) del gobierno de España (financiado por MI-CIU/AEI/10.13039/ 501100011033/ FEDER EU, ERDF/EU). Departamento de Educación, Ciencia y Universidades del Gobierno de Aragón: Grupo CeNIT (Ref T31_23R).

REFERENCIAS

[1] Barelli, Riccardo & D'Onghia, Mario & Longari, Stefano, "Towards Secure Electronic Voting: a Survey on E-Voting Systems and Attacks.", IEEE Access, Vol 11, pp. 1-31, 2023. DOI: 10.1109/ACCESS.2025.3569334.

[2] ¿Qué es Docker y cómo funciona? https://www.redhat.com/es/topics/containers/what-is-docker. Consultado en 12/06/2025

[3] K3s - Lightweight Kubernetes. https://docs.k3s.io/ Consultado en 12/06/2025

[4] atlantic, «Proxmox: Qué es y por qué es la mejor alternativa a VMware», Atlantic International Technology. Accedido: 15 de abril de 2025. [En línea]. Disponible en: https://www.atlantic.es/blog/proxmox/

[5] The Kubernetes Networking Guide. CNI. Flannel https://www.tkng.io/cni/flannel/ Consultado en 12/06/2025

[6] Docker Compose https://docs.docker.com/compose/ Consultado en 12/06/2025

[7] Martínez, José & Salazar, José & Piles, Joan & Saldana, Jose & Casadesus, Luis & Ruíz-Mas, José & Navajas, Julian. (2014). An eVoting System Based on Ring Signatures. Network Protocols and Algorithms.. Network Protocols and Algorithms. 6. 38-54. 10.5296/npa.v6i2.5390.

[8] José Luis Salazar, Joan Josep Piles, José Ruiz-Mas, José María Moreno-Jiménez, "Security approaches in e-cognocracy", Computer Standards & Interfaces, Volume 32, Issues 5–6, Pages 256-265, ISSN

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Bienestar digital: Detección y bloqueo de flujos IP en redes Wi-Fi

Zineb Helali Amoura, Roberto Sauras Alcober, Daniel Andrés Pérez, Julián Fernández-Navajas, María Canales-Compés, José Ruiz-Mas.

Departamento Ing. Electrónica y Comunicaciones, Instituto Universitario de Investigación en Ingeniería de Aragón.
Universidad de Zaragoza
Edificio Ada Byron, Campus Río Ebro, Calle María de Luna 1, 50018, Zaragoza, España.
Email autor de contacto, {navajas,mcanales,jruiz}@unizar.es

El objetivo de este artículo es presentar diferentes alternativas para la monitorización y gestión de flujos de tráfico IP desde dispositivos móviles, con el fin de identificar patrones de uso que puedan derivar en hábitos perjudiciales. Se proponen dos esquemas de despliegue para redes Wi-Fi domésticas, ambos basados en la sustitución del punto de acceso (AP) tradicional por un equipo configurado para la captura de tráfico. La primera alternativa consiste en emplear un AP comercial modificado, mientras que la segunda utiliza una plataforma de bajo coste de propósito general, como Raspberry Pi. Es un requisito fundamental que la sustitución sea transparente para los dispositivos del usuario, garantizando su funcionamiento normal. Una vez interceptado el tráfico, se emplean Wireshark y Zeek para su análisis e inspección. Posteriormente, se utilizan Iptables y Open vSwitch (OvS) para aplicar políticas de control y gestión sobre los paquetes, basadas en diferentes criterios.

***Palabras Clave*- Calidad de Experiencia, Monitorización de Red, Redes Definidas por Software.**

I. Introducción

En la actualidad, el uso de dispositivos conectados a Internet (teléfonos móviles, tabletas y ordenadores) se ha extendido de forma masiva y desde edades muy tempranas. Diversos estudios han alertado sobre los efectos que un uso prolongado y no supervisado de las pantallas puede tener sobre el desarrollo neurológico, el comportamiento social y el bienestar general, especialmente en población infantil y adolescente [1,2]. Este fenómeno ha dado lugar a una creciente preocupación por parte de familias, educadores e instituciones sobre cómo se utiliza la conectividad, qué tipo de contenido se consume y cuánto tiempo se permanece en línea.

Dejando al margen consideraciones éticas, resulta pertinente disponer de herramientas capaces de observar y analizar en tiempo real el comportamiento de flujos IP en redes de uso común. Se trata de utilizar sistemas para detectar amenazas de seguridad o ataques maliciosos, para adaptarlo a un sistema que identifica patrones de uso que puedan derivar en hábitos perjudiciales o poco saludables para los usuarios de internet. Estas situaciones, sin ser necesariamente maliciosas, pueden comprometer el equilibrio digital necesario para un desarrollo saludable de la vida personal, académica y social.

Aunque ya existen soluciones de control parental que permiten analizar, en cada dispositivo, el tráfico generado, y la consiguiente limitación para los caso no deseados, las restricciones impuestas por los sistemas operativos móviles actuales, especialmente en Android e iOS, hacen que esto no sea posible sin acceso privilegiado. Además, el *rooteo* de dispositivos es cada vez menos viable por motivos técnicos y de seguridad dado que los sistemas, tanto Android como iOS, restringen significativamente el acceso de terceras aplicaciones a funciones de supervisión o control, dificultando una intervención externa y personalizada. Por otra parte, muchos usuarios no tienen conocimientos técnicos suficientes para configurar correctamente estas herramientas, o bien los menores pueden encontrar formas de eludirlas.

Por ello, este trabajo está basado en la observación del tráfico desde la propia red, lo que permite inferir patrones de uso como se hace en trabajos de detección de intrusiones [3]. Esta solución debe ser replicable y poder desplegarse en plataformas de bajo coste, como una Raspberry Pi, facilitando así su aplicación en entornos reales.

A pesar de que existen muchas tecnologías que permiten conectar los equipos, en los dispositivos móviles, las más usadas son tecnologías Wi-Fi y móviles. En un entorno de laboratorio es posible trabajar con ambas y en el presente trabajo se ha elegido trabajar sobre conexiones Wi-Fi dejando las redes móviles como una futura línea de investigación.

Una vez fijada la tecnología Wi-Fi, se debe proceder al análisis de los flujos de tráfico, utilizando una herramienta de captura de tráfico y unos *script* para poder identificar automáticamente de qué tipo de tráfico se trata, la aplicación específica a la que se conecta y si ha sido el usuario el que ha interactuado con la aplicación o es el servidor el que ha iniciado el flujo de tráfico y el usuario no está usando su dispositivo. Con esta información se deben extraer distintos datos sobre los patrones de uso de un usuario y disponer de la opción de bloquear distintos tipos de conexiones en función de un horario u otros criterios. Los datos de patrones de uso pueden ser de gran ayuda no solo para realizar un

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

estudio general de la sociedad y su interacción con los dispositivos móviles, sino que también podría ayudar a psicólogos u otro tipo de especialistas en su trabajo a la hora de entender a un paciente.

La estructura del artículo es tal que en el apartado II se presenta las arquitecturas de trabajo, en el apartado III se explican los métodos de virtualización utilizados y en el IV se exponen conclusiones y líneas futuras de investigación.

II. ARQUITECTURAS DE TRABAJO

En el presente trabajo nos vamos a centrar en estudiar las conexiones realizadas desde un dispositivo móvil conectado a una red doméstica.

A. Alternativas de despliegue de soluciones.

Se han valorado distintas formas de capturar el tráfico generado por los dispositivos móviles. En la Figura 1 se muestran las diferentes alternativas de conexión a servicios de Internet como Youtube, Instagram, Spotify, Whatsapp, etc. La que está numerada con el valor 1, identifica un camino para el flujo IP a través de la red móvil. La etiquetada como 2 es una conexión a través de la Wi-Fi proporcionada por el router doméstico. Estas dos posibilidades son las habituales para las conexiones de los dispositivos móviles. Las alternativas numeradas con el 3 y 4 son las propuestas en este trabajo. La 3 consiste en sustituir la conexión Wi-Fi proporcionada por el router doméstico por otra que sea proporcionada por un AP que se ha configurado a tal efecto en el laboratorio. La alternativa 4 hace lo mismo que la 3 pero en lugar de utilizar un AP comercial, utiliza una plataforma de bajo coste de propósito general (como puede ser una Raspberry). Estas alternativas deben ser sencillas de implementar, que no supongan modificar en exceso la red existente y que las estaciones móviles no detecten cambio alguno en su conexión.

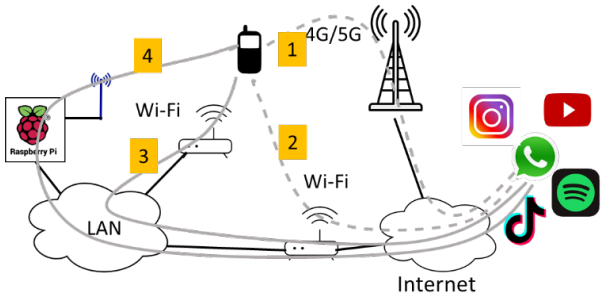


Fig. 1. Alternativas presentes y propuestas

B. Herramientas de monitorización de tráfico IP.

Otro aspecto fundamental es qué sistema de monitorización del tráfico IP para accesos Wi-Fi se va a utilizar. De las múltiples opciones posibles nos vamos a quedar con dos herramientas ampliamente conocidas: Wireshark [4] y Zeek [5]. Se seleccionaron estas herramientas por su capacidad para llevar a cabo análisis detallados del tráfico de red y su capacidad para obtener los datos necesarios para cumplir los objetivos de este trabajo. Wireshark es una herramienta de captura de tráfico IP, paquete a paquete, que ofrece una amplia gama de funciones de filtrado que permiten especificar qué tipos de tráfico y protocolos se desean capturar y analizar. Zeek tiene un enfoque orientado a eventos. Es una herramienta automatizada que nos permite recolectar el tráfico de red y mediante unos *script*, lo analiza en tiempo real, creando una lista de eventos que reflejan la actividad y que son almacenados en diferentes archivos log. Para el desarrollo de los *script*, Zeek utiliza su propio lenguaje denominado Zeek-Script, permitiendo un análisis automático y personalizado del tráfico y que facilita la detección de posibles irregularidades o tendencias de comportamiento en la red.

Wireshark será de gran ayuda para que mediante una visualización de los paquetes recibidos y los protocolos empleados en la red saber lo que deberíamos hacer de forma automática. Esto nos dará la capacidad de descubrir conexiones particulares que nos llamen la atención, particularmente durante una captura manual en la que tenemos control total sobre las acciones realizadas desde un dispositivo móvil.

Después, Zeek se empleará para automatizar los conocimientos obtenidos en capturas previas. Zeek nos dará la capacidad de incluir los filtros que necesitamos y nos ofrecerá datos más detallados y precisos sobre el tráfico de red que ha sido capturado.

C. Filtrado de paquetes y control de tráfico.

Una vez monitorizado el tráfico y se haya identificado aquel sobre el cual debemos actuar, se procede al filtrado y control del mismo. La identificación de tráfico en un proceso complejo, cuya realización se deberá analizar en trabajos futuros, dado que el presente es en proceso y no final. Estas técnicas de filtrado permiten decidir qué paquetes pueden circular y cómo se deben gestionar, basándose en diferentes criterios. El filtrado de paquetes consiste en examinar cada paquete de red para determinar si debe ser aceptado, rechazado o modificado, en función de reglas predefinidas. Estas reglas pueden basarse en direcciones IP, puertos, protocolos, estados de conexión u otros atributos del tráfico. El control del tráfico, por su parte, amplía el filtrado con acciones adicionales, como la priorización o limitación del ancho de banda, lo que permite gestionar el rendimiento de la red y garantizar la calidad de servicio (QoS).

Se va a utilizar dos mecanismos: Iptables y OvS [6]. Iptables, opera en el núcleo del sistema operativo. Permite definir reglas que controlan el tráfico entrante y saliente, basándose en la información contenida en las cabeceras de los paquetes (por ejemplo, dirección IP de origen o destino, puerto de origen o destino, protocolo). Este tipo de filtrado es eficiente para tareas básicas, pero no tiene visibilidad sobre los datos de las aplicaciones. OvS (Open vSwitch) es una herramienta que construye la conmutación en switch virtuales que operan desde los niveles MAC hasta los de aplicación. Ofrece un mayor grado de flexibilidad y control directo sobre el tráfico gracias a su funcionamiento en diferentes niveles de la arquitectura. En este trabajo, OvS se emplea para aplicar políticas de control de tráfico de manera precisa y dinámica, actuando sobre flujos específicos de la red Wi-Fi monitorizada. OvS puede ser configurado individualmente o bien a través de un controlador externo, por ejemplo Ryu.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

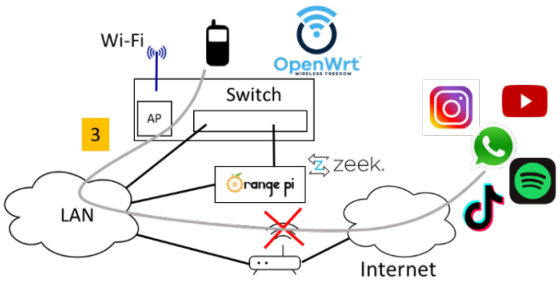


Fig. 2. Despliegue con AP comercial

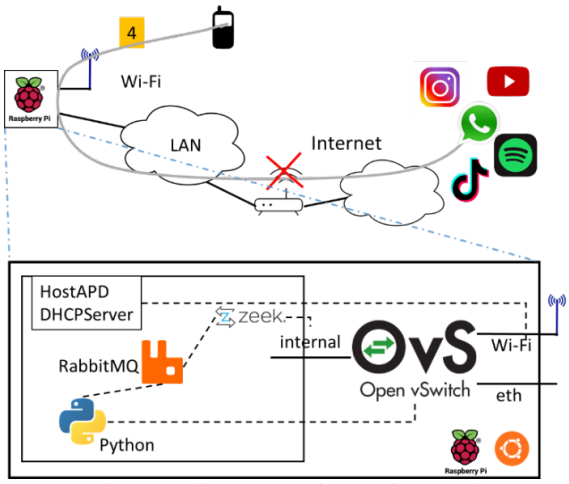


Fig. 3. Despliegue con AP de propósito general

III. Despliegue con AP comercial

Habitualmente, los dispositivos acceden a Internet mediante la conexión a un *router* principal que dispone de un AP. Como se presenta en la figura 2, en este despliegue se desactivará la red Wi-Fi del router y se conectará a la LAN un nuevo punto de acceso con el mismo ESSID, al cual deberán conectarse los dispositivos. Este AP (con S.O. OpenWRT e Iptables) define un puerto espejo para la conexión Wi-Fi, que estará conectado a una OrangePi con Zeek para inspeccionar todo el tráfico que lo atraviese. Se han desarrollado un conjunto de *script* de Zeek que identificará las aplicaciones en uso, limitará el acceso a las mismas de los dispositivos registrados en una lista, y todo ello, dentro de un rango horario determinado. Estos script pueden, además, bloquear las conexiones mediante reglas de Iptables que se ejecutan en el AP, para lo que será necesario enviarlas por *ssh*.

Para desarrollar convenientemente los *script* de Zeek, se ha realizado una etapa previa de captura mediante Tcpdump en el equipo Orange_pi y posterior monitorización manual utilizando Wireshark. Las capturas se han tomado en un entorno controlado de laboratorio en el cual se limita el uso a un único servicio (se ha hecho en particular para Youtube y Spotify). Este análisis inicial de las capturas, ha permitido programar la automatización del proceso.

Durante el proceso de análisis se planteó la posibilidad de emplear un dispositivo móvil *rooteado* o realizar las acciones desde un ordenador con el fin de capturar todo el tráfico de forma más controlada. Sin embargo, esta opción se descartó ya que no reflejaría fielmente el entorno real de uso de un dispositivo móvil, que constantemente está generando conexiones a distintos servicios, actualizaciones de aplicaciones y sincronizaciones en segundo plano. Así, se decidió capturar el tráfico en un escenario más cercano al día a día de un usuario, donde los datos recogidos reflejan de manera más precisa las conexiones que se producen de forma natural mientras se utilizan aplicaciones como Spotify o Youtube.

Uno de los aspectos más comprometidos es la identificación de los servicios a los que se conectan los equipos móviles. Básicamente se hace mediante dirección IP, puesto que los servidores de las aplicaciones están en un rango conocido de IP públicas. Sin embargo estas direcciones pueden cambiar de un día para otro. Este es el motivo por el cual se ha definido un fichero externo que han de leer los *script* de zeek. Existe un fichero para cada aplicación que puede ser modificado en tiempo de ejecución sin que afecte al funcionamiento del sistema.

IV. Despliegue con AP de propósito general

Como se muestra en la figura 3, en este despliegue se sustituye la red Wi-Fi del router por una Raspberry Pi con S.O. Ubuntu, manteniendo el mismo ESSID. En este equipo, el servicio Hostapd crea el punto de acceso para los dispositivos móviles y el servidor isc-dhcp-server les asigna direcciones IP dinámicamente manteniendo un registro, facilitando así su posterior identificación.

Una vez establecida la red Wi-Fi gestionada por el sistema, el elemento que permite capturar y controlar el tráfico es OvS (directamente o a través de un controlador como Ryu), que actúa como puente virtual entre todas las interfaces implicadas. En esta arquitectura, OvS conecta la interfaz inalámbrica desde la que emite el punto de acceso, la interfaz de salida hacia Internet y una interfaz virtual interna denominada Internal. Esta última es clave, ya que el sistema se configura para que todo el tráfico que entra por la red Wi-Fi y se dirige hacia Internet sea duplicado y enviado también a Internal. De este modo, Zeek puede escuchar pasivamente en esa interfaz y capturar todo el tráfico generado por los dispositivos móviles sin interferir en el flujo original.

Gracias a esta duplicación, Zeek analiza en tiempo real las conexiones establecidas por los dispositivos, identificando patrones característicos del uso de servicios como Spotify o Youtube mediante *scripts* personalizados. Cuando se detecta un evento relevante, Zeek genera un log estructurado que es enviado como mensaje a través del sistema de mensajería RabbitMQ [8]. Este mensaje llega al módulo de decisión desarrollado en Python, que evalúa si se cumplen las condiciones necesarias para permitir o bloquear el tráfico. En función de esa lógica, el módulo puede actuar directamente sobre OvS (o el controlador Ryu) para introducir reglas que bloqueen la conexión correspondiente,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

por ejemplo, impidiendo el acceso a una IP o puerto determinado.

Este diseño permite separar de forma clara la captura del tráfico (gestionada por Zeek), el análisis y la toma de decisiones (a cargo del módulo Python), y el control de red (realizado mediante OvS y Ryu).

V. CONCLUSIONES Y LÍNEAS FUTURAS

Se ha conseguido desarrollar una pareja de despliegues sobre escenarios domésticos Wi-Fi que permiten el análisis de los flujos de tráfico, utilizando una herramienta de captura de tráfico como es Zeek y unos *script* (tanto en Zeek como en Python) para poder identificar automáticamente de qué tipo de tráfico se trata y la aplicación específica a la que se conecta.

Durante el análisis de los dos despliegues se ha podido constatar varios aspectos relevantes que afectan a la variabilidad del tráfico:

En primer lugar, los protocolos utilizados por las diferentes aplicaciones y servicios han variado en el transcurso del tiempo. Por ejemplo, lo que se transmitía mediante TCP se ha modificado transmitiendo mediante UDP y usando el protocolo QUIC a nivel superior. Esto ha hecho que se redefinan los *script* de Zeek y que se haga de tal forma que para un futuro sea adaptable fácilmente.

Para este análisis, se realizaron capturas del tráfico mientras se repetían acciones concretas como la reproducción de canciones o la visualización de vídeos. Mediante la herramienta Wireshark se registraron todos los paquetes generados durante estas actividades [7], lo que permitió obtener una visión detallada de las conexiones establecidas y de la actividad de la red. Posteriormente, se procedió a un análisis manual minucioso: se revisaron cuidadosamente los paquetes capturados, identificando patrones consistentes que se repetían, tales como dominios asociados a los servicios analizados (audio-fa.scdn.co para Spotify o tiktokcdn.com para TikTok), así como los rangos de direcciones IP involucrados y las características generales del tráfico. Sin embargo las conclusiones no ha sido todo lo satisfactoria que se esperaban, por lo cual se ha decidido continuar con el análisis para obtener resultados más fiables.

Además, se ha constatado que muchas direcciones IP cambian constantemente, lo que dificulta una identificación precisa si se basara únicamente en ellas. No obstante, se observó que, aunque las direcciones IP pueden variar, suelen mantenerse dentro de ciertos rangos de direcciones asignados a las CDNs de cada servicio. Para reforzar la fiabilidad de la detección, se plantea combinar la identificación de estos rangos de IP con el análisis de los dominios específicos (SNI/DNS) que aparecen durante las conexiones. Este enfoque aporta una mayor certeza de que la acción que se está analizando pertenece a un servicio concreto y no a otro que pueda compartir la misma infraestructura de distribución de contenidos. Este análisis se puede utilizar para obtener información complementaria de la localización de servidores en función de la ubicación geográfica de los clientes o de la franja horaria de acceso.

Otro aspecto importante es ver cómo se han desarrollado elementos de inteligencia tanto en los *script* de Zeek como en Python. Esto hace que el sistema disponga de varias alternativas que pueden ser comparadas en diferentes situaciones. Incluso, para un futuro, se podrá aplicar inteligencia artificial sobre el propio Python. Esto es algo que supone la continuación de la actual línea de investigación.

Algo que todavía no se ha conseguido es saber si el flujo IP detectado ha sido producido porque el usuario es el que ha interactuado con la aplicación o el usuario no está usando su dispositivo y el flujo lo provoca la aplicación de manera desatendida. Como se ha dicho en la introducción, esta distinción resulta fundamental porque se deben extraer distintos datos sobre los patrones de uso de un usuario que pueden ser de gran ayuda no solo para realizar un estudio general de la sociedad y su interacción con los dispositivos móviles, sino que también podría ayudar a psicólogos u otro tipo de especialistas en su trabajo a la hora de entender a un paciente.

AGRADECIMIENTOS

Los investigadores pertenecen al I3A (Instituto Universitario de Investigación en Ingeniería de Aragón, Universidad de Zaragoza). Esta investigación ha sido financiada por NE-WLAN (PID2022-136476OB-100) del gobierno de España (financiado por MI-CIU/AEI/ 10.13039/501100011033/FE-DER EU, ERDF/EU) y por el Departamento de Educación, Ciencia y Universidades del Gobierno de Aragón: Grupo Ce-NIT (Ref T31_23R).

REFERENCIAS

[1] Nasaescu, E., Marín, L., Llorent, V. J., Ortega, R., & Zych, I. (2018). Abuse of technology in adolescence and its relation to social and emotional competencies, emotions in online communication, and bullying. Computers in Human Behavior, 88, 114-120.

[2] Jean M. Twenge y W. Keith Campbell. ((Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study)). Preventive Medicine Reports 12 (2018), pgs. 271-283. doi:10.1016/ j.pmedr.2018.10.003. url: https://doi. org/10.1016/j.pmedr.2018.10.003.

[3] Rodríguez, María; Alesanco, Álvaro; Mehavilla, Lorena; García, José. Evaluation of Machine Learning Techniques for Traffic Flow-Based Intrusion Detection. SENSORS. 22 - 23, pp. 9326 [30 pp.]. 2022. ISSN 1424-8220 DOI: 10.3390/s22239326

[4] Wireshark Foundation. (n.d.). Documentation. Wireshark. Retrieved June 25, 2024, from https://www.wireshark.org/docs/

[5] Zeek Documentation Team. (n.d.). Quick start guide. Zeek Documentation. Retrieved June 25, 2024, from https://docs.Zeek.org/en/master/quickstart.html (visitado 09-06-2025)

[6] Azwad Hasan, Md Omar Faisal y Sariful Haque Khan. ((Exploring SDN Based Firewall and NAPT: A Comparative Analysis with Iptables and OVS in Mininet)). ResearchGate (2024). url: https://www.researchgate.net/publication/384062702_Exploring_SDN_Based_Firewall_and_NAPT_A_Comparative_Analysis_with_Iptables_and_OVS_in_Mininet.

[7] Ruchi Tuli. ((Analyzing Network Performance Parameters Using Wireshark)). International Journal of Network Security & Its Applications (IJNSA) 15.1 (2023), págs. 1-10. doi:10.5121/ijnsa.2023. 15101. url: https://aircconline.com/ijnsa/V15N1/15123ijnsa01.pdf.

[8] RabbitMQ. AMQP Concepts Guide. Guía conceptual que explica los intercambios, colas, enlaces y otras nociones clave de AMQP en RabbitMQ. 2024. url: https://www.rabbitmq.com/tutorials/amqp-concepts (visitado 09-06-2025).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Pronóstico del rendimiento citrícola con IA híbrida, imágenes captadas por drones y variables climáticas

Guillem Montalbán-Faet, Andoni Salcedo-Navarro, Santiago Felici-Castell, Juan J. Pérez-Solano
Antonio Soriano-Asensi, Enrique Navarro-Camba, Miguel García-Pineda, Jaume Segura-Garcia
Departamento de Informática, Universitat de València
Avda de la Universitat, s/n - 46100 Burjassot.
Email: guillem.montalban@uv.es, andoni.salcedo@uv.es, santiago.felici@uv.es, juan.j.perez@uv.es
antonio.soriano-asensi@uv.es, enrique.navarro@uv.es, miguel.garcia-pineda@uv.es, jaume.segura@uv.es

Este trabajo presenta un enfoque híbrido para la predicción del rendimiento citrícola, combinando técnicas de aprendizaje automático, visión por computador mediante drones y análisis de variables meteorológicas. Se recopilaron datos climáticos históricos de estaciones locales y se modelaron en diferentes formatos (mensual, resumido y anual con eventos extremos) para evaluar su poder predictivo. Se aplicaron diversos algoritmos, desde regresión lineal hasta modelos avanzados como Random Forest, destacándose este último por su precisión. Asimismo, se desarrolló un procedimiento de detección y conteo de naranjas mediante imágenes captadas por drones (DJI Mavic 3M), a partir del entrenamiento con redes neuronales YOLOv8 y YOLOv11, y mejoradas mediante la técnica SAHI. Los resultados mostraron una precisión superior al 80% en la estimación por árbol. El estudio confirma que eventos meteorológicos críticos como heladas, lluvias intensas o rachas de viento tienen un impacto significativo sobre la producción. La combinación de estos datos con técnicas de IA permite anticipar riesgos y optimizar la planificación agrícola. Este enfoque es escalable a otros cultivos y regiones, y contribuye al desarrollo de sistemas inteligentes para una agricultura sostenible.

***Palabras Clave*—Drones, ML, YOLO, Predicción de producción citrícola, factores climáticos.**

I. INTRODUCCIÓN

España produjo 7.2 Mt de cítricos dulces en 2024, liderando la exportación europea; casi la mitad procede de la Comunidad Valenciana [1], [2]. Esta cadena de valor moviliza a más de 100 000 trabajadores estacionales y exige previsiones de cosecha con varios meses de antelación para optimizar contratos y logística.

El *aforo* tradicional, esto es, la extrapolación de conteos manuales combinada con estadísticas climáticas, presenta (i) cobertura espacial limitada, (ii) variabilidad interobservador y (iii) escasa capacidad para capturar interacciones no lineales clima–fenología. La teledetección satelital ha mejorado la caracterización de la cubierta vegetal, aunque sigue sin ofrecer conteos explícitos de frutos [3]. La irrupción de los drones equipados con cámaras subcentimétricas y detectores en tiempo real, como YOLO, ha reducido drásticamente el coste de captura y procesamiento de imágenes [4]. Sin embargo, la integración de señales visuales con variables meteorológicas para predecir el rendimiento a escala de parcela permanece como un reto poco abordado [5].

Paralelamente, la agricultura de precisión se está beneficiando de tecnologías emergentes como la Inteligencia Artificial (IA), el Internet de las Cosas (IoT) y los vehículos aéreos no tripulados (UAV). Los sistemas IoT, mediante redes de sensores inalámbricos, permiten monitorizar en tiempo real variables ambientales y favorecen decisiones basadas en datos que reducen desperdicios y mejoran la sostenibilidad. Los UAV, dotados de cámaras multiespectrales y térmicas, proporcionan observaciones aéreas de alta resolución que complementan los datos a ras de suelo. Esta sinergia posibilita inspecciones autónomas a gran escala, cartografiando la distribución de frutos, detectando anomalías y estimando rendimientos con una precisión inédita.

La agricultura, especialmente el cultivo de frutales como el naranjo, está fuertemente influenciada por las condiciones climáticas. En regiones mediterráneas, la producción de naranjas constituye una actividad económica fundamental y cualquier herramienta que permita predecir con antelación la producción puede facilitar decisiones clave sobre logística, comercialización y gestión del riesgo. En este trabajo planteamos un *framework* híbrido que fusiona detección de frutos mediante YOLOv11 +

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

SAHI [6] con predictores climáticos diarios para estimar el rendimiento de naranjos. La propuesta combina la potencia de la visión por computador con la información agrometeorológica [7], capturando tanto la variabilidad espacial de los conteos como la influencia temporal del clima. El sistema es replicable en cooperativas que dispongan de estaciones meteorológicas cercanas y drones de bajo coste, por lo que ofrece una vía práctica para mejorar la planificación de cosecha, la contratación laboral y la logística en el sector citrícola.

El artículo se estructura de la siguiente manera: la Sección II revisa la literatura relacionada y resalta sus principales diferencias; la Sección III examina el impacto de las variables meteorológicas en la producción de naranjas; la Sección IV describe el procedimiento de conteo de frutos mediante imágenes capturadas por drones y redes neuronales; la Sección V presenta y analiza los resultados obtenidos con los modelos entrenados; la Sección VI introduce el framework híbrido propuesto para optimizar la producción de cítricos; por último, la Sección VII recoge las conclusiones y plantea líneas de investigación futura.

II. TRABAJOS PREVIOS

La literatura reciente sobre estimación del rendimiento citrícola puede agruparse en tres líneas principales: (i) conteo visual de frutos con visión por computador, (ii) modelos climáticos puros y (iii) enfoques multimodales que combinan varias fuentes de datos.

Conteo visual de frutos

Los detectores basados en CNN han disparado la precisión y velocidad del conteo. Daiaeddine et al. emplean YOLOv7 sobre imágenes UAV y alcanzan $R^2 = 0.98$ entre frutos detectados y conteos manuales, pero el modelo ignora por completo los factores meteorológicos y se limita a 20 árboles de muestreo [8]. Trabajos como AG–YOLO [9], introducen mecanismos de atención para mitigar la oclusión, mientras que [10] integran CNN y datos multiespectrales para cartografiar densidades frutales a escala de huerto. Todas estas propuestas se centran en el *conteo estático* y no generan predicciones temporales de rendimiento.

Modelos climáticos

Por otro lado, hay trabajos que abordan el problema utilizando únicamente variables meteorológicas. Un ejemplo es el ANN de Almady et al., que modela el rendimiento de cinco cultivares egipcios con temperatura, lluvia y unidades térmicas acumuladas, alcanzando un MAPE del 5.4% [11]. Si bien capturan la dinámica fenológica, carecen de información sobre la carga real de frutos y, por tanto, no pueden corregir desequilibrios inducidos por poda, plagas u oclusiones canópicas.

Enfoques multimodales

Los intentos de fusión de fuentes son aún escasos.

- Ma et al. combinan imágenes RGB y datos de campo para un modelo CNN de predicción de rendimiento; aunque mejoran la fiabilidad, no usan detección por instancias ni contemplan eventos climáticos extremos [12].
- Suarez et al. pronostican el rendimiento de bloques citrícolas entre 2 y 9 meses antes de la cosecha usando series temporales Sentinel-2 y clima; sacrifican resolución espacial (bloque frente a parcela) y carecen de conteo explícito de frutos [13].
- El dataset *CropNet* introduce imágenes Sentinel-2 y meteorología para predicciones a escala de condado en EE. UU.; ofrece un valioso marco de pruebas, pero no incluye anotaciones de frutos individuales ni aborda el nivel de detalle parcelario requerido por las cooperativas. [14].

Diferencias con nuestra propuesta

Nuestro trabajo se aparta de estas líneas al:

1) Emplear **YOLOv11** junto con **SAHI** para un conteo por instancias robusto a oclusiones y variabilidad de tamaño, superando los detectores YOLOv7/8 usados por la mayoría de estudios citados.
2) Integrar **predictores climáticos diarios**, incluidos *eventos extremos* (olas de calor, lluvias torrenciales), capturando no linealidades clima–fenología ausentes en los modelos puramente visuales o climáticos.
3) Ofrecer un **framework híbrido replicable** con drones de "bajo coste" y estaciones meteorológicas convencionales, pensado para cooperativas que necesitan previsiones con meses de margen para planificar mano de obra y logística.
4) Entregar predicciones a *escala de parcela*, un nivel de resolución clave para contratos agrarios, frente al árbol individual (poco escalable) o al bloque/condado (demasiado agregado) de trabajos previos.

En resumen, mientras los estudios existentes tienden a aislar la componente visual o la meteorológica, o bien trabajan a escalas espaciales que no satisfacen las necesidades operativas, nuestra propuesta fusiona ambas dimensiones en un sistema práctico y de bajo coste, que cubre el vacío entre investigación y aplicación industrial.

III. IMPACTO DE VARIABLES METEOROLÓGICAS EN LA PRODUCCIÓN DE NARANJAS

En esta sección, se investiga el uso de datos meteorológicos como predictores de la producción de naranjas. El enfoque combina el análisis de datos, la ingeniería de características y el aprendizaje automático, con el objetivo de identificar patrones y relaciones que permitan mejorar la estimación de la producción agrícola de la finca "La Campaneta" situada en el municipio de Corbera (Valencia), ver Figura 1.

A. Metodología

1) Origen y procesamiento de los datos: Se han utilizado datos de estaciones meteorológicas de AEMET y AVAMET (Polinyà del Xúquer). Estas fuentes proporcionan información precisa sobre una amplia gama de variables climáticas: precipitaciones, temperaturas, viento,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

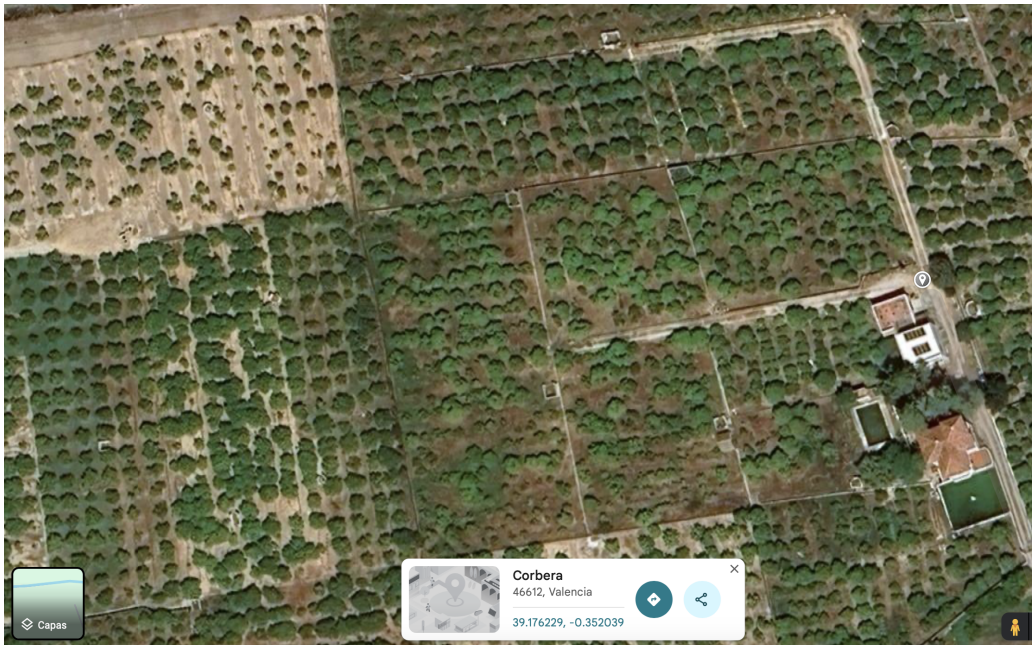


Fig. 1. Parcela situada en el municipio de Corbera (Valencia), donde se han realizado las pruebas

humedad y cobertura nubosa. Los parámetros disponibles en el dataset están recogidos en la Tabla I.

Se recopilaron más de 25 variables mensuales, incluyendo valores extremos (temperaturas máximas y mínimas absolutas), acumulados (precipitación mensual), y conteos de eventos (número de días con viento fuerte, heladas, etc.). La estructura original de los ficheros fue adaptada: se separaron campos con información combinada, se normalizaron unidades y se completaron valores faltantes mediante interpolación y referencias cruzadas con datos de estaciones cercanas.

2) Fases fenológicas y sensibilidad climática: El ciclo productivo del naranjo se puede dividir en:

- **Floración (marzo-mayo):** periodo de alta sensibilidad a vientos y lluvias intensas.
- **Cuajado y engorde (junio-septiembre):** el fruto es susceptible a olas de calor y sequías.
- **Maduración y cosecha (octubre-diciembre):** los excesos de precipitación o viento pueden reducir la calidad del fruto.

Se identificaron tres tipos de eventos meteorológicos críticos: heladas, lluvias torrenciales y rachas de viento. Estas variables se parametrizaron mediante umbrales climáticos significativos (por ejemplo, ($\geq$ 30) mm de lluvia en un día) y se integraron al dataset en forma de variables binarias o normalizadas.

3) Estrategias de modelado del dataset: Se evaluaron distintos esquemas de representación del dataset:

1) **Dataset mensual ampliado:** para este caso se ha escogido el subconjunto de variables compuesto por ['hr', 'nw_55', 'tm_min', 'nt_30', 'np_100', 'nw_91', 'e', 'np_300', 'p_mes', 'w_med', 'nt_00', 'tm_mes', 'tm_max', 'fecha'], que recoge los valores mensuales para humedad relativa, precipitación, tensión de vapor, velocidad del viento y temperatura, además de variables que hacen referencia a eventos relevantes para la producción frutícola, cómo los días de mucho calor (nt_30), heladas (nt_00) o lluvias torrenciales (np_100 y np_300). Cada mes del año recibe las variables meteorológicas correspondientes y se le asigna el mismo valor de producción anual. Este enfoque genera 48 observaciones a partir de 4 años.
2) **Dataset mensual resumido:** se discrimina más información que en el conjunto ampliado para comprobar si existe redundancia que pueda afectar en la estimación de la producción de forma negativa. En este caso se tiene en cuenta los valores mensuales de precipitación, temperatura y velocidad de viento y las lluvias torrenciales más fuertes (np_300) y los días con heladas.
3) **Dataset anual con eventos:** este enfoque híbrido se construye a partir de promedios anuales y se centra en variables meteorológicas que, por su relevancia agronómica, se consideran determinantes en la producción de naranjas. Se seleccionan variables representativas de temperatura media, velocidad del viento y precipitación total, complementadas con eventos extremos expresados mediante variables derivadas. Estas se construyen a partir de datos normalizados y ponderaciones informadas por conocimiento técnico y observación empírica.
El período más crítico para la formación del fruto es la primavera (marzo, abril y mayo), por lo que se da especial atención a los eventos meteorológicos registrados durante estos meses. Las variables más

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
DESCRIPCIÓN DE VARIABLES METEOROLÓGICAS MENSUALES

Variable	Descripción
fecha	Fecha en formato "YYYY-M"
hr	Humedad relativa media [%]
n_des	Número de días despejados [Nº de días]
e	Tensión de vapor media [décimas de hPa]
n_cub	Número de días cubiertos [Nº de días]
n_nub	Número de días nubosos [Nº de días]
np_001	Número de días de precipitación apreciable ($\geq$ 0.1 mm) [Nº de días]
np_010	Número de días de precipitación $\geq$ 1 mm [Nº de días]
np_100	Número de días de precipitación $\geq$ 10 mm [Nº de días]
np_300	Número de días de precipitación $\geq$ 30 mm [Nº de días]
p_max	Precipitación máxima diaria [mm]
p_mes	Precipitación total mensual [mm]
nw_55	Número de días con viento $\geq$ 55 km/h [Nº de días]
nw_91	Número de días con viento $\geq$ 91 km/h [Nº de días]
w_rec	Recorrido medio diario [km]
w_med	Velocidad media (media de observaciones 07, 13 y 18 UTC) [km/h]
w_racha	Dirección y velocidad de la racha máxima [dirección en decenas de grado; velocidad en m/s]
nt_00	Número de días con temperatura mínima $\leq$ 0 ºC [Nº de días]
nt_30	Número de días con temperatura máxima $\geq$ 30 ºC [Nº de días]
ta_max	Temperatura máxima absoluta [ºC]
ts_min	Temperatura mínima más alta [ºC]
ta_min	Temperatura mínima absoluta [ºC]
ti_max	Temperatura máxima más baja [ºC]
tm_mes	Temperatura media [ºC]
tm_min	Temperatura media de las mínimas [ºC]
tm_max	Temperatura media de las máximas [ºC]

relevantes se parametrizan del siguiente modo:

- **Lluvias torrenciales:** en todos los años analizados hay al menos un día con precipitaciones superiores a 100 mm y a 300 mm. Para resumir esta información, se normalizan las variables np_100 y np_300, y se define la siguiente combinación ponderada:

  $$\text{lluvia_torrencial} = 0.3\times\text{np100_n}+0.7\times\text{np300_n}$$

- **Heladas:** se utiliza la variable nt_00, que representa el número de días con temperatura mínima menor o igual a 0 ºC. Esta se normaliza para obtener:

  $$\text{heladas} = \text{nt00_n}$$

- **Rachas de viento:** las rachas superiores a 50 km/h afectan negativamente tanto a la floración como al fruto joven. Como la variable nw_91 presenta valores nulos, se descarta, y se emplea únicamente la variable nw_55 normalizada:

  $$\text{rachas_viento} = \text{nw55_n}$$

Este diseño permite capturar tanto las condiciones climáticas generales del año como los impactos puntuales de eventos extremos en momentos críticos del ciclo del cultivo, aportando sensibilidad agronómica sin renunciar a la simplicidad de un modelo anual.

Toda esta metodología se recoge de forma gráfica en la Figura 2, que resume las fases clave desde la recopilación de datos hasta la predicción mediante modelos de aprendizaje automático.

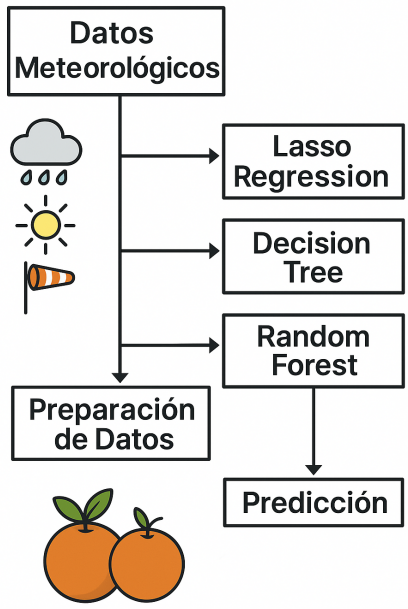


Fig. 2. Metodología para la predicción de la producción de naranjas mediante modelos de aprendizaje automático

B. Modelos Aplicados

Los modelos implementados en las primeras pruebas han sido:

- **Regresión Lineal (LR):** modelo paramétrico clásico que estima la relación lineal entre variables independientes y dependientes. Sirve como línea base.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- **Regresión Polinómica (PR):** extiende LR al incluir términos cuadráticos e interacciones, permitiendo capturar relaciones no lineales suaves. Es más flexible pero puede sobreajustar con pocos datos.
- **Random Forest (RF):** conjunto de árboles de decisión construido sobre subconjuntos aleatorios del dataset. Es robusto frente a valores atípicos, maneja bien datos no lineales y ofrece interpretabilidad mediante la importancia de variables.

Cada modelo se entrenó utilizando validación cruzada *k-fold* y se evaluó con RMSE (*Root Mean Square Error*) y error porcentual respecto a la producción.

Adicionalmente, se realizó un análisis complementario con cada uno de los algoritmos para identificar qué grupo de variables climáticas aportaba mayor poder predictivo por sí solo. Para ello, se crearon versiones simplificadas de los datasets donde se evaluaron de manera independiente las variables de viento, temperatura y precipitación. Los resultados revelaron que, de forma consistente, todos los algoritmos obtenían mejores resultados cuando se entrenaban exclusivamente con variables de viento, temperatura y lluvia, en ese orden de importancia general. Este hallazgo refuerza la hipótesis de que estos tres factores climáticos son especialmente relevantes para explicar la variabilidad en la producción anual de naranjas.

IV. DETECCIÓN Y CONTEO DE NARANJAS MEDIANTE IMÁGENES CAPTADAS CON DRONES

Para la observación de naranjas sobre el propio campo, hemos utilizado un dron DJI Mavic 3M para la obtención de imágenes y para su detección y conteo hemos usado dos redes neuronales para realizar una comparativa, como son YOLOv8 [5] y YOLOv11 [15], [16].

Las características de la cámara que lleva nuestro dron DJI Mavic 3M son:

- Especificaciones de la cámara RGB:
  1) Posee un campo de visión (FoV) de 84º con una distancia focal equivalente a 24 mm.
  2) El tamaño de las imágenes tomadas es de 5280 x 3956 píxeles.
  3) Su resolución es en H.263 y 4K de 3840×2160 a 30 fps, y en Full HD es de FHD: 1920×1080 a 30 fps.
  4) Los formatos de imagen que soporta son: JPEG/DNG para imágenes y MP4 para video.
- Especificaciones de la cámara multiespectral:
  1) Posee un FoV de 73.91º con una distancia focal equivalente a 25 mm.
  2) Bandas de cámara multiespectral (Verde (G): $560 \pm 16nm$; Rojo (R): $650 \pm 16nm$; Borde rojo (RE): $730 \pm 16nm$; Infrarrojo cercano (NIR): $860 \pm 26nm$)
  3) Su resolución es en H.263 y 4K de 3840×2160 a 30 fps, y en Full HD es de FHD: 1920×1080 a 30 fps.
  4) 4. Los formatos de imagen que soporta son: JPEG/PNG para imágenes y MP4 para video

En el caso de las redes neuronales usadas, el acrónimo ”YOLO” hace referencia a sus siglas en inglés ”You Only Look Once”, que indica que su entrenamiento con unos pocos casos en cada imagen, mediante un etiquetado adecuado, puede producir detecciones muy precisas. YOLOv8 destaca por su excelente combinación de precisión, velocidad y estabilidad. Posteriormente, con los avances en detección de objetos, surgió la posibilidad de emplear la versión de YOLOv11 y realizar la comparativa de rendimiento frente al mismo dataset de entrenamiento. Ambos algoritmos se han importado directamente desde el directorio proporcionado por Ultralytics y se ha cargado una versión preentrenada, seleccionada para ahorrar tiempo y recursos.

Un aspecto destacado de esta sección es el uso de una herramienta para aumentar la precisión del modelo. Dado el gran tamaño de las imágenes capturadas por el dron, se ha optado por emplear la técnica SAHI (Slicing Aided Hyper Inference) [6], que permite una detección más precisa mediante una ventana deslizante, la cual segmenta de la imagen grande en detecciones más reducidas en tamaño. Esto permite a YOLO analizar detalles más finos que podrían perderse en imágenes de gran tamaño y resolución.

Para la evaluación de su rendimiento se han usado las métricas clave, la precisión ($Accuracy$), la recuperación ($Recall$), la precisión positiva ($Precision$) y la puntuación F1 ($F1 - Score$), que se miden a partir de la evaluación de la matriz de confusión en la que se obtienen el número de verdaderos positivos (True Positives o TP), el número de verdaderos negativos (True Negatives o TN), el número de falsos positivos (False Positives o FP) y el número de falsos negativos (False Negatives o FN).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

V. RESULTADOS

A. Efecto de la meteorología sobre la producción de naranjas

A continuación se presentan los resultados obtenidos en la evaluación de los modelos sobre los tres datasets definidos: mensual ampliado (Tabla II), mensual reducido (Tabla III) y anual con eventos (Tabla IV). Las métricas utilizadas son el error cuadrático medio (RMSE) promedio y su correspondiente porcentaje respecto al total de producción.

Estos resultados confirman que el dataset mensual reducido logra una buena compensación entre complejidad y rendimiento, manteniendo un RMSE bajo con modelos avanzados. El modelo Random Forest se destaca como el más consistente en cuanto a rendimiento y estabilidad en todos los conjuntos.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla II
RESULTADOS SOBRE EL DATASET MENSUAL AMPLIADO

Modelo	RMSE promedio	Error %
Regresión Lineal (LR)	3255.44 kg	37.29%
Regresión Polinómica (PR)	27795.69 kg	318.39%
Random Forest (RF)	706.13 kg	8.09%

Tabla III
RESULTADOS SOBRE EL DATASET MENSUAL REDUCIDO

Modelo	RMSE promedio	Error %
Regresión Lineal (LR)	2984.28 kg	34.18%
Regresión Polinómica (PR)	6075.90 kg	69.60%
Random Forest (RF)	623.76 kg	7.15%

B. Detección de naranjas mediante YOLO

Para el trabajo, con YOLOv8 y YOLOv11 con la técnica SAHI, se ha empleado un tamaño de batch de 16 y se ha entrenado durante 350 épocas. Las imágenes se han reescalado a 1760x1318 píxeles para su etiquetado y entrenamiento. Tras el entrenamiento, se seleccionó el modelo con el mejor rendimiento para realizar una validación, con el objetivo de medir su precisión y evaluar su robustez. Finalmente, se llevarán a cabo pruebas con imágenes en condiciones reales para observar el desempeño del modelo y verificar su capacidad para detectar objetos con precisión.

La tabla V muestra las métricas de evaluación de la precisión en la detección de las imágenes obtenidas por el dron. Estas evaluaciones se han realizado usando SAHI. En ella se observa que, tras un etiquetado manual adecuado del dataset de 350 imágenes, los valores de accuracy obtenidos superan el 80%, usando YOLOv11.

En la figura 3 se muestran dos inferencias del modelo YOLOv11 entrenado sin la técnica SAHI (a) y con la técnica SAHI (b). Se observa claramente un mayor número de casos cuando se usa SAHI.

C. Discusión

Este estudio demuestra que la combinación de ingeniería de características y aprendizaje automático puede ofrecer herramientas precisas para predecir la producción agrícola. A pesar de la limitación del número de años analizados, el uso de datos mensuales amplía el tamaño de muestra efectiva y permite al modelo aprender mejor. Además, el enfoque basado en eventos extremos tiene valor explicativo: por ejemplo, se observó que los años con mayor frecuencia de lluvias intensas durante la primavera coincidieron con menores niveles de producción. Una limitación del enfoque es la homogeneidad artificial creada al replicar la producción anual en cada mes. Sin embargo, en ausencia de datos de producción mensuales, esta estrategia resulta razonable.

Por otro lado, el uso de técnicas de IA combinadas con detección por imágenes (mediante el uso de drones y YOLO) permite que se pueda realizar una valoración promedio de la producción por árbol y así realizar una estimación de manera más precisa de la producción de un campo.

Tabla IV
RESULTADOS SOBRE EL DATASET ANUAL CON EVENTOS

Modelo	RMSE promedio	Error %
Regresión Lineal (LR)	3819.94 kg	43.76%
Regresión Polinómica (PR)	3986.71 kg	45.67%
Random Forest (RF)	3697.84 kg	42.36%

Fig. 3. Detección de naranjas con YOLOv11. En (a) no se ha usado SAHI y en (b) sí se ha usado SAHI.

VI. FRAMEWORK HÍBRIDO PARA EL PRONOSTICO DE PRODUCCIÓN CITRÍCOLA

Esta sección describe el diseño modular del framework y los flujos de procesamiento de datos que permiten generar una estimación mensual de rendimiento citrícola.

A. Visión general

La Figura 4 resume el pipeline end-to-end. Dos flujos de entrada, meteorología diaria y conteo de frutos aéreo, se procesan en paralelo y convergen en una capa de fusión para alimentar un modelo híbrido de regresión basado en *Random Forest*.

De acuerdo, con los datos obtenidos en las secciones previas, nuestro framework actuará con los siguientes detalles para cada módulo:

a) Imágenes UAV: Captura a $0.7\,\mathrm{cm\,px^{-1}}$ con vuelos a $12\,\mathrm{m}$; tiling de $640\,\mathrm{px}$ y *data augmentation* balancean oclusión y variabilidad lumínica.

b) Meteorología: Se derivan 25 variables incluyendo índices de estrés hídrico y extremos; escalado robusto.

c) Detección YOLOv11+SAHI: 120 epochs, batch 16, *mAP@0.5*=0.84.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

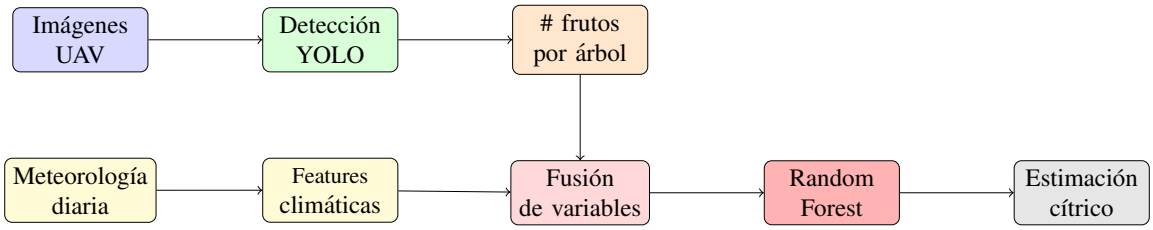


Fig. 4. Pipeline del framework híbrido propuesto.

Tabla V
MÉTRICAS DE EVALUACIÓN EN LA DETECCIÓN DE NARANJAS.

Modelo	Accuracy	Recall	Precision	F1-Score
Yolov8	0.74	0.78	0.77	0.77
Yolov11	0.81	0.88	0.81	0.84

d) Fusión: Concatenación de 28 atributos (25 clima + 3 visión); entrenamiento de RF con 300 árboles y profundidad 10.

B. Pseudocódigo del flujo completo

El Algoritmo 1 resume los pasos principales desde la adquisición de datos hasta la predicción mensual. El flujo se compone de cinco etapas principales:

1) **Extracción de características climáticas**. A partir de las series D, se calculan estadísticos descriptivos y índices fisiológicos (*GDD*, estrés hídrico) que capturan el impacto del clima sobre la fenología citrícola.
2) **Detección y conteo de frutos**. Cada imagen se subdivide para preservar resolución efectiva; YOLOv11 detecta frutos y SAHI combina detecciones solapadas, generando un conteo robusto C.
3) **Agregación por árbol**. El número total de frutos se normaliza por el recuento de árboles visibles para obtener un indicador de carga productiva y su variabilidad (σ_C).
4) **Fusión de variables**. Se concatena el vector de clima F_{clima} con los tres descriptores de visión F_{vision}; este vector de 28 dimensiones alimenta al regresor.
5) **Inferencia con Random Forest**. Cada árbol produce una estimación individual y la media ponderada entrega la salida final $\hat{Y}$. La naturaleza ensemblada mejora la generalización frente a la multicolinealidad.

Algorithm 1: Estimación de rendimiento citrícola

Entrada: D: series meteorológicas diarias; I: conjunto de imágenes UAV; $\mathcal{W}_{yolo}$: pesos YOLO; $\mathcal{W}_{rf}$: pesos Random Forest
Salida : $\hat{Y}$: producción citrícola estimada (kg/ha)

$F_{clima} \leftarrow$ EXTRACTFEATURES(D) // 25 variables
$C \leftarrow 0$ // contador de frutos
foreach *imagen* $i \in I$ **do**
  $T \leftarrow$ TILE(i)
  foreach *tile* $t \in T$ **do**
    $B \leftarrow$ YOLODETECT$(t, \mathcal{W}_{yolo})$ // bounding boxes
    $C \leftarrow C + |B|$
$F_{vision} \leftarrow \{C/|\text{árboles}|,\ \text{densidad},\ \sigma_C\}$
$F \leftarrow F_{clima} \cup F_{vision}$ // vector de 28 features
$\hat{Y} \leftarrow$ RANDOMFORESTPREDICT$(F, \mathcal{W}_{rf})$
return $\hat{Y}$

VII. CONCLUSIONES

El estudio evidencia que combinar un conteo visual de alta resolución con predictores agrometeorológicos diarios permite prever el rendimiento citrícola con precisión operativa. Al analizar cuatro campañas sucesivas, se observa que los episodios extremos de lluvia, heladas y viento intenso son los principales responsables de las caídas de producción, ya que afectan etapas críticas como la floración, el cuajado y la maduración. Incorporar estos eventos en el modelo mejora notablemente su capacidad explicativa y facilita la toma de decisiones para mitigar daños. Además, trabajar con series meteorológicas mensuales, en lugar de promedios anuales, capta mejor las fluctuaciones climáticas y reduce de manera significativa el error de predicción.

En términos de desempeño algorítmico, el enfoque híbrido resulta particularmente sólido: el Random Forest ofrece un equilibrio óptimo entre robustez frente a valores atípicos, interpretabilidad y un bajo RMSE, superando con holgura a las alternativas lineales y polinómicas. Por su parte, el flujo visual basado en la combinación de YOLOv11 y SAHI alcanza una accuracy superior al 80%, lo que refuerza la fiabilidad de las estimaciones del número de frutos. Este marco tecnológico, apoyado en drones de bajo coste y estaciones meteorológicas convencionales, se presenta, por tanto, como una solución escalable y transferible a otros cultivos y regiones con escasa necesidad de adaptación.

Con todo, persisten ciertas limitaciones: el horizonte temporal de datos aún es corto para capturar anomalías climáticas infrecuentes; la homogeneidad de la variable respuesta a lo largo del año introduce colinealidad que podría resolverse con aforos parciales; y la validación se ha realizado en una única parcela, lo que restringe la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

generalización espacial. De ahí que las próximas líneas de trabajo contemplen ampliar las series históricas, incorporar huertos con distintas orientaciones, explorar modelos de boosting y redes neuronales temporales más profundas, fusionar señales multisensoriales (NDVI, NDRE, humedad de suelo) y, por último, valorar el retorno económico derivado de la reducción de incertidumbre y del ahorro logístico que posibilita esta plataforma predictiva.

AGRADECIMIENTOS

Los autores agradecen a la Agencia Estatal de Investigación (AEI) y al Fondo Europeo de Desarrollo Regional (FEDER) la financiación de esta investigación dentro de los proyectos con referencias de subvención TED2021-131040B-C33 (Agricultura 6.0) financiado en el programa "Proyectos Estratégicos para la Transición Ecológica y Digital 2021" por MCIN/AEI/10. 13039/501100011033 y la Unión Europea NextGenerationEU/PRTR, así como al proyecto PID2021-126823OB-I00 financiado por MCIN/AEI/10. 13039/501100011033 y por la "Unión Europea NextGenerationEU/PRTR". También a la Generalitat Valenciana, por la financiación del proyecto DRONIA dentro de la convocatoria Retech-IA GVA (ref. INREIA/2024/164). Los autores también quieren agradecer a la empresa IDS Topografía por el apoyo y ayuda prestada.

REFERENCIAS

[1] Food and Agriculture Organization of the United Nations FAOSTAT: Crops and livestock products 2024. (https://www.fao.org/faostat,2024), Consultado: 15 mayo 2025

[2] European Commission, Directorate-General for Agriculture EU Citrus Market Report 2024. (https://agriculture.ec.europa.eu,2024), Accedido: 20 mayo 2025

[3] Huang, L. & Zhang, W. Remote sensing-based estimation of citrus yield in subtropical orchards. *Remote Sensing*. **15**, 1053 (2023)

[4] Jocher, G., A. Chaurasia, and J. Qiu. "Ultralytics YOLO (Version 8.0. 0)[Computer Software]. 2023." Available online: Available online: https://github. com/ultralytics/ultralytics (accessed on 3 June 2025).

[5] D. Reis, J. Kupec, J. Hong, and A. Daoudi. "Real-time flying object detection with YOLOv8". arXiv:2305.09972 [cs.CV] 2024. [Online]. Available: https://arxiv.org/abs/2305.09972

[6] SAHI. https://github.com/obss/sahi. [Online]. [Accessed: 15/05/2025].

[7] V. B. Moreto, G. d. S. Rolim, B. G. Zacarin, A. P. Vanin, L. M. de Souza, and R. R. Latado, "Agrometeorological models for forecasting the qualitative attributes of "València" oranges," Theoretical and Applied Climatology, vol. 130, p. 847–864, 2017.

[8] Daiaeddine MJ, Badrouss S, El Harti A, Bachaoui EM, Biniz M, Mouncif H. "UAV imagery, advanced deep learning, and YOLOv7 object detection model in enhancing citrus yield estimation". Foods and Raw Materials. 2025;13(2):242–253. https://doi.org/10.21603/2308-4057-2025-2-650

[9] Y. Lin, Z. Huang, Y. Liang, Y. Liu, W. Jiang, "AG-YOLO: A Rapid Citrus Fruit Detection Algorithm with Global Context Fusion". Agriculture 2024, 14, 114. https://doi.org/10.3390/agriculture14010114

[10] M. Zhang, Sh. Wang, "Agricultural Unmanned Systems: Empowering Agriculture with Automation", Agronomy, 2024 14, 6, 1203. [Online] https://www.mdpi.com/2073-4395/14/6/1203, doi: 10.3390/agronomy14061203

[11] S.S. Almady, M. Abdel-Sattar, S.M. Al-Sager, S.A. Al-Hamed, A.M. Aboukarima, "Employing an Artificial Neural Network Model to Predict Citrus Yield Based on Climate Factors". Agronomy 2024, 14, 1548. [Online] https://doi.org/10.3390/agronomy14071548

[12] J. Ma, Y. Wu, B. Liu, W. Zhang, B. Wang, Z. Chen, G. Wang, A. Guo, "Wheat Yield Prediction Using Unmanned Aerial Vehicle RGB-Imagery-Based Convolutional Neural Network and Limited Training Samples". Remote Sens. 2023, 15, 5444. [Online] https://doi.org/10.3390/rs15235444

[13] L.A. Suarez, A. Robson, J. Brinkhoff, "Early-Season Forecasting of Citrus Block-Yield Using Time Series Remote Sensing and Machine Learning: A Case Study in Australian Orchards". International Journal of Applied Earth Observation and Geoinformation, 2023, 122, 103434. [Online] https://www.sciencedirect.com/science/article/pii/S1569843223002583

[14] F. Lin, K. Guillot, S. Crawford, Y. Zhang, X. Yuan, N.F. Tzeng, "An Open and Large-Scale Dataset for Multi-Modal Climate Change-aware Crop Yield Predictions". In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), August 25–29, 2024, Barcelona, Spain. ACM. [Online] https://doi.org/10.1145/3637528.3671536

[15] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding. "Yolov10: Real-time end-to-end object detection". arXiv:2405.14458 [cs.CV] 2024. [Online]. Available: https://arxiv.org/abs/2405.14458

[16] M. Al Rabbani Alif, "YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications in Intelligent Transportation Systems". arXiv:2410.22898 [cs.CV] 2024. [Online]. Available: https://arxiv.org/abs/2410.22898

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Metodología docente para el análisis de tráfico de red mediante sistemas inteligentes

V. Carneiro, D. Fernández, M. Álvarez, F.J. Nóvoa
Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC),
Universidade da Coruña
15071. A Coruña (Spain)
{victor.carneiro,diego.fernandez,marco.antonio.agonzalez,francisco.javier.novoa}@udc.es

Este trabajo presenta una metodología estructurada y adaptativa para desarrollar sistemas inteligentes de detección de anomalías en tráfico de red, basada en la adaptación del modelo CRISP-DM al ámbito de la ciberseguridad. La propuesta abarca todo el ciclo de vida del proyecto, desde la recopilación y preparación de datos —registros, flujos y paquetes de red— hasta la validación de modelos de aprendizaje automático y procesamiento de lenguaje natural, con énfasis en explicabilidad y robustez. Se ofrece como guía replicable y flexible para TFG y TFM, adaptable a diversos contextos de análisis y defensa activa. Su objetivo es facilitar proyectos académicos técnicamente rigurosos, promoviendo coherencia, reproducibilidad y competencias prácticas alineadas con los estándares actuales.

I. Introducción

La creciente sofisticación y volumen del tráfico de red ha motivado el desarrollo de mecanismos automáticos de análisis orientados a detectar anomalías y comportamientos maliciosos. En este contexto, los modelos de inteligencia artificial (IA), tanto supervisados como no supervisados, han demostrado un potencial significativo para la clasificación eficiente de patrones irregulares en entornos complejos y heterogéneos, especialmente en nuevos ataques o ataques de día cero [1]. El presente trabajo se enmarca en este ámbito, proponiendo una metodología estructurada y adaptativa, basada en el modelo CRISP-DM [2], para la realización de trabajos de fin de grado (TFG) o fin de máster (TFM) que involucren el análisis de tráfico de red mediante técnicas de IA. Esta metodología aborda de manera integral el tratamiento de datos provenientes de capturas de paquetes (*PCAPs*), flujos o secuencias de red y registros del sistema, desde su análisis inicial hasta el despliegue de modelos funcionales, considerando tanto la calidad del proceso como su rendimiento operativo.

La motivación principal de este trabajo radica en la necesidad de cubrir dicho vacío, proporcionando una referencia estructurada y operativa que facilite la labor tanto de estudiantes como de tutores. En un momento en que el uso de modelos de IA, incluyendo los grandes modelos de lenguaje (LLM), se ha disparado en entornos académicos, se vuelve imperativo disponer de una herramienta metodológica clara que oriente el uso riguroso y efectivo de datos de tráfico de red para fines de clasificación, detección de anomalías o defensa activa en ciberseguridad.

Este artículo realiza varias aportaciones en este ámbito. En primer lugar, se presenta una adaptación del modelo CRISP-DM al dominio específico del análisis de tráfico de red, introduciendo una descripción detallada de las fases necesarias para abordar un proyecto completo. En segundo lugar, se catalogan las principales tareas y decisiones que deben tomarse en cada fase, incluyendo ejemplos representativos. En tercer lugar, se identifican y analizan las métricas de evaluación más adecuadas para validar modelos de clasificación, con énfasis en métricas de explicabilidad cuando resulten pertinentes. Finalmente, se abordan aspectos clave para el despliegue efectivo de los modelos desarrollados, contemplando consideraciones de escalabilidad, robustez y rendimiento.

El objetivo general de este trabajo consiste en establecer una guía metodológica exhaustiva que sirva de base para la elaboración de TFG y TFM que impliquen el uso de datos de red en combinación con modelos de inteligencia artificial. Se pretende así dotar a los proyectos académicos de una mayor coherencia, reproducibilidad y calidad técnica, promoviendo al mismo tiempo el desarrollo de competencias prácticas alineadas con los estándares.

El resto del artículo se estructura de la siguiente manera. En la Sección II se presenta una revisión del trabajo relacionado, incluyendo aproximaciones metodológicas y técnicas existentes en el ámbito del análisis de tráfico de red mediante IA. La Sección III detalla la metodología propuesta, adaptando el modelo CRISP-DM al contexto específico tratado. En la Sección IV se expone un estudio de caso que ilustra la aplicación práctica de la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

metodología. Finalmente, en la Sección V se presentan las conclusiones.

II. Trabajo relacionado

La selección de una metodología adecuada es fundamental en cualquier proceso de minería de datos. El SAS Institute define este concepto como ”el proceso de hallar anomalías, patrones y correlaciones en grandes conjuntos de datos para predecir resultados” [3]. La extracción de conocimiento a partir de los datos se debería llevar a cabo mediante procesos sistemáticos y procedimientos con los pasos bien definidos. Una metodología es un acercamiento general que guía las técnicas y actividades que se usarán en un dominio específico. La metodología no se basa en ciertas tecnologías o herramientas. En su lugar, aporta un marco de trabajo para adquirir resultados mediante el uso de un amplio rango de métodos, procesos y heurísticas [4].

Las tres metodologías más empleadas en minería de datos son: Cross Industry Standard Process for Data Mining (CRISP-DM), Knowledge Discovery in Databases (KDD) [5] y Sample, Explore, Modify, Model, Assess (SEMMA) [6]. Según Gregory Piatetsky, la más usada es CRISP-DM, alcanzando un 42% de los proyectos en 2014 [4].

En el caso de SEMMA, fue propuesta por el SAS Institute, que define esta metodología como el proceso para seleccionar, explorar y realizar modelado aplicado a grandes cantidades significativas de datos que permiten descubrir patrones como herramientas de apoyo para el negocio. El SAS Institute ha comentado que SEMMA, más que un metodología es un conjunto de herramientas enfocadas a los aspectos de desarrollo propio de modelo de minería. Propone una aproximación técnica a la minería de datos.

Por su parte, KDD está enfocado en el descubrimiento de conocimiento, puesto que propone extraer patrones a partir de los datos, para que el usuario los pueda analizar.

Con respecto a CRISP-DM, ofrece un marco más general, ya que integra todas las tareas necesarias en los proyectos de minería de datos, desde la fase de comprensión del problema hasta la puesta en producción de la solución.

CRISP-DM fue propuesta en 1999 por un importante consorcio de empresas europeas con el fin de desarrollar una guía de referencia de libre distribución. Se trata de una metodología madura y extensamente implantada. Aunque la metodología se presenta como un ciclo, CRISP-DM es flexible, y las fases no siempre se ejecutan de un modo secuencial.

En la Figura 1 se aprecian las seis fases que componen CRISP-DM: Comprensión del Negocio, Comprensión de los Datos, Preparación de los Datos, Modelado, Evaluación y Despliegue.

III. Metodología propuesta

A continuación, se abordarán las fases de la metodología CRISP-DM adaptadas para el desarrollo de sistemas inteligentes de detección de anomalías en red.

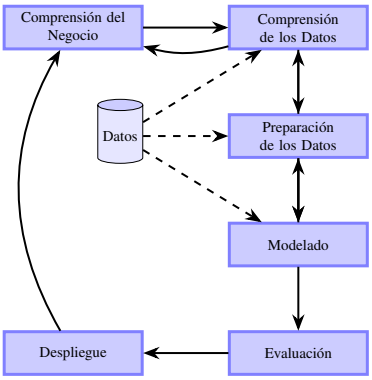


Fig. 1. Fases de CRISP-DM

A. Comprensión del negocio

Se identifican los objetivos de detección, que pueden incluir identificación de intrusiones, fallos de servicio, accesos irregulares, etc. Se delimitan los requisitos técnicos, restricciones y objetivos del sistema de análisis a desarrollar.

B. Comprensión de los datos

En esta etapa, se realiza una exploración sistemática de las fuentes de datos disponibles, evaluando tanto su estructura como su contenido, con el objetivo de identificar oportunidades y limitaciones inherentes al conjunto de datos. Las actividades esenciales de esta fase son:

- **Recolección de datos iniciales**. Es habitual emplear fuentes de datos públicas y reconocidas como CICIDS2017, UNSW-NB15 o MAWI, las cuales ofrecen un amplio espectro de escenarios realistas de ciberataques, tráfico benigno y comportamiento anómalo. En la Tabla I se muestran los principales datasets utilizados en este tipo de trabajos.
- **Descripción de los datos**. La descripción de estas fuentes incluye el análisis de la estructura de los logs, donde se detallan los campos registrados por flujo o sesión (direcciones IP, puertos, protocolos, duración, número de paquetes, etiquetas de clase, entre otros), así como la naturaleza de los eventos recopilados (por ejemplo, escaneo, denegación de servicio, inyección de comandos o acceso no autorizado).
- **Exploración de los datos**. Se trata de una tarea en la que se realiza una selección de atributos relevantes, que se tendrán en cuenta en la fase de preparación de los datos. Este proceso está guiado por criterios estadísticos, como puede ser la varianza, la entropía o la correlación, el análisis gráfico de las características importantes para la variable objetivo, y el conocimiento del dominio. Este paso no sólo permite reducir la complejidad del análisis posterior, sino también aumentar la interpretabilidad y eficiencia de los modelos predictivos. Por ejemplo, al representar gráficamente la duración de conexión frente al número de paquetes por flujo y desagregarlo por tipo de ataque, es posible observar características diferenciadoras entre tráfico legítimo y malicioso.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- **Verificación de la calidad de los datos**. Implica la comprobación de la completitud de los registros, la detección de valores atípicos o inconsistentes, y la evaluación de posibles sesgos de clase. Por ejemplo, en conjuntos de datos de ciberseguridad, suele observarse un fuerte desbalance entre clases benignas y maliciosas, lo cual debe ser identificado tempranamente para evitar distorsiones en la etapa de modelado.

Network	
KDD99	NSL-KDD
UNSW-NB15	NF-UNSW-NB15-v2
CSE-CIC-IDS2017	CSE-CIC-IDS2018
InSDN	UGR'16
LITNET-2020	Edge-IIoTset
UWF-ZeekData22	CESNET-TimeSeries24
WTMC-2021	CRiSIS-2022
NFS-2023	ZYELL-NCTU
NAB	SMD
MAWILab	CTU-13
CIC-IoT	Bot-IoT
Logs	
HDFS	Openstack Logs
Thunderbird Logs	BGL Logs
Linux Syslogs	Windows Event Logs
Proxifier Logs	Zookeeper Logs
Apache Logs	Spark Logs

Tabla I
DATASETS USADOS EN DETECCIÓN DE ANOMALÍAS.

C. Preparación de los datos

En esta etapa se tiene por objetivo la construcción del dataset final que se va a emplear en el modelado. Esta fase es crítica dentro del ciclo de vida del proyecto, ya que la calidad y adecuación de los datos impacta directamente en el rendimiento del modelo. La metodología CRISP-DM contempla esta etapa como un paso necesario antes de iniciar el modelado, asegurando que los datos estén en condiciones óptimas para su análisis posterior. Adquiere especial relevancia, ya que los datos pueden provenir de múltiples fuentes heterogéneas, como registros de tráfico de red, archivos de logs, sensores IDS/IPS (Intrusion Detection/Prevention Systems), entre otros. Para ello, es necesario realizar una serie de actividades, que se detallan a continuación:

- **Selección de atributos relevantes**. Este procedimiento consiste en determinar qué características del conjunto de datos son relevantes para alcanzar los objetivos del análisis, como la detección de intrusiones, la clasificación de tráfico malicioso o la prevención de amenazas. La selección se basa tanto en el conocimiento del dominio como en técnicas estadísticas o de aprendizaje automático. Por ejemplo, en un dataset de tráfico de red, atributos como el número de paquetes por sesión, el tiempo de duración de la conexión, el puerto de destino o el número de bytes transferidos pueden resultar más relevantes que campos triviales como el identificador de sesión. Esta selección puede apoyarse en análisis de correlación, técnicas de reducción de dimensionalidad (e.g. el Análisis de Componentes Principales o PCA), o mediante el uso de modelos base que permitan estimar la importancia de cada atributo (e.g. la Eliminación Recursiva de Características o RFE).
- **Limpieza del dataset**. En esta actividad se deben gestionar tanto los valores perdidos (missing values) como los valores atípicos (outliers), ya que pueden distorsionar los resultados del modelado. En escenarios reales de red, los datos pueden estar incompletos, duplicados o contener registros corruptos. Es habitual encontrar missing values debido a errores en la captura de tráfico o en la agregación de logs. Por ejemplo, puede faltar información sobre el protocolo utilizado en una sesión o el país de origen de una IP. Estos valores pueden imputarse mediante técnicas como la media, la mediana, la moda o mediante métodos más avanzados como la imputación multivariante.
  Asimismo, los outliers pueden corresponder tanto a errores como a comportamientos anómalos (i.e. un pico inusual en el número de paquetes en una sesión puede ser indicativo de un ataque de denegación de servicio). Pueden identificarse mediante visualizaciones (e.g. diagramas de cajas y bigotes), métricas estadísticas (e.g. Z-Score o rangos intercuartílicos) o técnicas específicas (e.g. Isolation Forest), y decidir si deben eliminarse o transformarse. En este contexto, los outliers no siempre deben eliminarse, ya que pueden representar comportamientos maliciosos, por lo que su tratamiento debe hacerse con especial precaución.
- **Transformación de variables**: Esta actividad incluye diversas operaciones que buscan adaptar los datos al formato requerido por los algoritmos de modelado. Entre las transformaciones más comunes se encuentran:
  - *Normalización o estandarización*: esencial cuando se utilizan modelos sensibles a la escala, como los basados en distancias (e.g. KNN, SVM). Por ejemplo, el tamaño de los paquetes, tiempo de respuesta, o número de conexiones por unidad de tiempo, para permitir comparaciones consistentes entre distintos registros.
  - *Codificación de variables categóricas*: los algoritmos de aprendizaje automático no pueden operar directamente con variables categóricas, por lo que deben codificarse. Técnicas como One-Hot Encoding, Label Encoding o hashing son ampliamente utilizadas. Por ejemplo, los protocolos de red (TCP, UDP, ICMP), las direcciones IP o los tipos de servicios.
  - *Agregación temporal*: útil para representar eventos por ventanas de tiempo (por ejemplo, número de conexiones entrantes por minuto), facilitando la detección de patrones de ataque persistente.
- **Integración de datos de diferentes fuentes**: En muchas ocasiones, los datos provienen de diversas fuentes, como firewalls, sondas de red, sistemas de detección de intrusiones, bases de inteligencia de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

amenazas, que deben ser combinadas para construir una visión unificada. La integración de estos datos es esencial para enriquecer el contexto del análisis y aumentar la capacidad de detección. Por ejemplo, un registro de conexión sospechosa extraído de un log de red puede enriquecerse con información geográfica de la IP de origen, así como con datos históricos de reputación de dicha dirección. Esta fusión requiere procesos de reconciliación de formatos, resolución de conflictos y sincronización temporal.

- **Selección de muestras**: Los problemas de ciberseguridad suelen presentar conjuntos de datos altamente desbalanceados, donde los ejemplos de tráfico malicioso representan un pequeño porcentaje frente al tráfico normal. Para mitigar esto, se emplean técnicas como *oversampling* (e.g. SMOTE), *undersampling*, o el uso de métricas robustas como la *F1-score* o el área bajo la curva ROC (*AUC-ROC*). También pueden usarse algoritmos adaptativos que penalizan más los errores en la clase minoritaria, como *Balanced Random Forest* o *cost-sensitive learning*. Además, la validación cruzada con estratificación permite mantener los porcentajes de las clases en todos los subconjuntos generados.

D. Modelado

En esta etapa se procede a la construcción y evaluación de los modelos predictivos, basados en el conjunto de datos previamente preparado. El modelado se orienta típicamente a tareas como la detección de intrusiones, clasificación de tráfico, identificación de comportamientos anómalos o predicción de amenazas. La metodología CRISP-DM establece que deben seleccionarse técnicas apropiadas según los objetivos del negocio, los requisitos computacionales y la naturaleza de los datos.

Para este tipo de problemas, se consideran tanto técnicas de aprendizaje automático tradicionales como enfoques avanzados basados en redes neuronales profundas:

- **Modelos clásicos**: Modelos como Random Forest, Support Vector Machines (SVM) o K-Nearest Neighbors (KNN) son ampliamente utilizados debido a su interpretabilidad y eficiencia. Por ejemplo, un modelo Random Forest puede aplicarse eficazmente para clasificar sesiones de red como benignas o maliciosas, a partir de características como duración, puertos usados o número de bytes transferidos. Las SVM son útiles en contextos donde existe una clara frontera entre clases, como la separación entre tráfico legítimo y tráfico relacionado con escaneos de red.
- **Deep Learning**: Modelos más complejos, como las redes LSTM (Long Short-Term Memory) o Transformers, se emplean cuando es necesario capturar la naturaleza secuencial o contextual de los datos. Por ejemplo, las LSTM permiten modelar secuencias temporales de eventos de red, lo cual es útil para identificar patrones de ataque persistente o APTs (Advanced Persistent Threats). Por otro lado, los Transformers, que han demostrado un desempeño sobresaliente en procesamiento de lenguaje natural, se están adaptando progresivamente a la ciberseguridad para tareas como la modelización contextual de logs, correlación de alertas, o detección de anomalías multieventos.

El proceso de modelado comprende además las siguientes actividades clave:

- **Entrenamiento y validación cruzada**: Para asegurar la robustez del modelo y evitar el sobreajuste (*overfitting*), se realiza una división del conjunto de datos en subconjuntos de entrenamiento, validación y prueba. Se aplica validación cruzada *k-fold*, que permite evaluar el desempeño del modelo en diferentes particiones y obtener métricas más generalizables. Esto es especialmente importante cuando se dispone de datasets de tamaño limitado, como suele ocurrir con ciertos registros de incidentes de seguridad.
- **Ajuste de hiperparámetros**: La optimización de hiperparámetros se realiza mediante técnicas como *Grid Search*, *Random Search* o aproximaciones bayesianas. Parámetros como la profundidad de los árboles en *Random Forest*, el *kernel* y el parámetro de margen en SVM, o el número de capas y unidades en redes neuronales, son ajustados para maximizar el rendimiento. En contextos de red, incluso pequeños ajustes pueden marcar la diferencia entre detectar una intrusión y pasarla por alto.

E. Evaluación

Esta etapa tiene como objetivo determinar en qué medida los modelos desarrollados cumplen con los objetivos de negocio y los criterios técnicos establecidos al inicio del proyecto. No sólo debe centrarse en el rendimiento cuantitativo del modelo, sino también en su capacidad para ser interpretado, validado por expertos humanos, y aplicado en contextos operativos reales.

1) Métricas de evaluación de la predicción: . Para cuantificar el rendimiento predictivo, se emplean diversas métricas clásicas de clasificación:

- **Exactitud (Accuracy)**: representa la proporción de predicciones correctas sobre el total de observaciones. Aunque útil como métrica general, puede ser engañosa en contextos con clases desbalanceadas, como la detección de ataques poco frecuentes en tráfico normal.
- **Precisión (Precision)**: mide la proporción de verdaderos positivos entre todas las predicciones positivas. Es especialmente relevante en ciberseguridad para evitar falsas alarmas (false positives), que pueden generar ruido operativo en los centros de monitorización (SOC).
- **Exhaustividad (Recall)**: indica la capacidad del modelo para identificar correctamente todos los eventos positivos (e.g. ataques detectados). Una alta exhaustividad es prioritaria en IDS, donde pasar por alto una amenaza puede tener consecuencias graves.
- **F1-score**: combina precisión y recall en una única métrica armónica, útil cuando se busca un equilibrio

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

entre ambas.

- **Área bajo la curva ROC (AUC-ROC)**: evalúa la capacidad del modelo para distinguir entre clases. Un valor cercano a 1 indica un excelente poder discriminativo. En entornos de red, esta métrica es útil para comparar modelos con umbrales de decisión ajustables.

2) Métricas de evaluación de los agrupamientos: a continuación, se destacan distintas métricas empleadas en aprendizaje no supervisado, con técnicas de *clustering* o agrupamiento:

- **Índice de Davies-Bouldin (DBI)**: mide la similitud promedio de cada agrupamiento con su agrupamiento más similar, donde la similitud es la relación entre las distancias dentro del agrupamiento y las distancias entre los agrupamientos.
- **Índice de Calinski-Harabasz**: también conocido como criterio de la razón de la varianza, se basa en la cohesión y en la separación. La primera indica que cada miembro de un clúster debe ser lo más cercano posible a los otros miembros del mismo clúster. Por su parte, la separación indica que los clústeres deben estar distantes entre sí, y presentar un mínimo acoplamiento. Para medir la separación se puede considerar el punto más cercano, el más distante o los centroides. En consecuencia, esta métrica mide la dispersión de las muestras a nivel inter-clúster e intra-clúster.
- **Silhouette**: buscar medir cómo de buenos son los clusters definidos. Para lograrlo, para todas las muestras calcula la distancia promedio de dicha muestra al resto de puntos en el mismo clúster, y la distancia promedio de dicha muestra a todos los demás puntos en el clúster más cercano.

3) Explicabilidad del modelo: Dada la creciente complejidad de los modelos modernos (especialmente en Deep Learning), se vuelve necesario incorporar técnicas de explicabilidad para facilitar la comprensión y validación de sus decisiones, especialmente en contextos sensibles como la respuesta a incidentes o auditorías de seguridad.

- **Local Interpretable Model-agnostic Explanations (LIME)**: permite explicar individualmente las predicciones de cualquier modelo clasificatorio generando aproximaciones locales interpretables. Por ejemplo, destacar que una alerta de tráfico malicioso se generó por una combinación de puerto inusual y número elevado de paquetes en un corto período.
- **SHapley Additive exPlanations (SHAP)**: basado en teoría de juegos, cuantifica la contribución de cada característica a la predicción del modelo. Esto permite construir interpretaciones globales (influencia general de cada variable) y locales (e.g. por qué se clasificó una sesión como ataque DDoS). En entornos de red, esto resulta crucial para justificar por qué un determinado patrón fue marcado como malicioso, facilitando la confianza del analista.
- **Fidelidad**: mide cuán bien se ajusta la explicación a lo que realmente hace el modelo subyacente. Una explicación con alta fidelidad replica con precisión las verdaderas decisiones del modelo, siendo especialmente útil para auditores o personal de cumplimiento.
- **Comprensibilidad**: se refiere a la facilidad con la que un ser humano puede entender la explicación ofrecida. En ciberseguridad, esto permite a los analistas tomar decisiones rápidas y efectivas frente a posibles amenazas, basándose en criterios transparentes.

F. Despliegue

La etapa de despliegue representa la transición del modelo analítico desde un entorno experimental hacia su integración en un sistema operativo o productivo. Esta fase es particularmente crítica, dado que las decisiones generadas por los modelos impactan directamente en la protección de infraestructuras, la detección temprana de amenazas y la respuesta ante incidentes.

El objetivo del despliegue es asegurar que los modelos desarrollados puedan ser utilizados de forma efectiva, sostenible y en tiempo adecuado dentro de los procesos operativos de seguridad de la organización. A continuación, se detallan los principales aspectos considerados:

- **Integración en entornos productivos**. El modelo debe integrarse dentro del ecosistema tecnológico existente, lo cual incluye sistemas de monitorización de red, Security Information and Event Management (SIEMs), firewalls y plataformas de respuesta a incidentes. La implementación puede realizarse a través de microservicios, contenedores (e.g. Docker), o mediante APIs RESTful que permitan consultas desde sistemas externos.
  Por ejemplo, un modelo de clasificación de tráfico puede desplegarse como un servicio que reciba flujos NetFlow en tiempo real, analice su comportamiento y devuelva etiquetas de riesgo (benigno, sospechoso, malicioso) que se visualicen directamente en la consola de un analista SOC.
- **Inferencia en tiempo real**. Uno de los requerimientos suele ser su capacidad de tomar decisiones con baja latencia. Por ello, el modelo debe estar optimizado para inferencia en tiempo real, incluso bajo condiciones de alta carga. Esto implica el uso de pipelines de datos eficientes (e.g. Apache Kafka o Flink), procesamiento en streaming y arquitecturas distribuidas. Un caso típico sería la detección de ataques DDoS, donde es necesario identificar patrones anómalos de tráfico en segundos para poder aplicar medidas de mitigación de forma proactiva.
- **Reentrenamiento y actualización continua**. Dado el carácter dinámico y evolutivo de las amenazas cibernéticas, los modelos desplegados deben incorporar mecanismos de actualización continua. Esto puede incluir:
  - Reentrenamiento periódico, utilizando datos recientes etiquetados por analistas.
  - Entrenamiento incremental, que permita adaptar el modelo sin necesidad de volver a entrenarlo completamente.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

  - Evaluación de deriva de concepto, para detectar cuándo el modelo comienza a degradarse debido a cambios en los patrones de tráfico o nuevas técnicas de ataque.

  Por ejemplo, un modelo que detecta *malware* basado en comportamiento puede requerir reentrenamiento mensual, a medida que aparecen nuevas variantes que el modelo original no reconoce.
- **Escalabilidad**. El sistema debe ser escalable tanto horizontal como verticalmente para garantizar un desempeño estable ante aumentos en el volumen de datos. Esto es particularmente importante en redes corporativas o proveedores de servicios donde se procesan millones de eventos por minuto. El uso de servicios en la nube, clústeres distribuidos o balanceadores de carga permite adaptar el sistema a escenarios de alta disponibilidad.
- **Visualización de resultados**. La capacidad de presentar los resultados del modelo de forma clara, concisa y accionable es esencial para los analistas de ciberseguridad. Esto implica el desarrollo de *dashboards* interactivos, alertas visuales y resúmenes ejecutivos. Herramientas como Kibana, Grafana o sistemas personalizados permiten mostrar:
  - Mapas de calor de eventos por geolocalización.
  - Líneas de tiempo de eventos anómalos.
  - Análisis detallado de predicciones con explicaciones generadas por SHAP o LIME.

  Por ejemplo, un *dashboard* puede mostrar en tiempo real los eventos de red clasificados como amenazas críticas, con enlaces directos a los logs originales y explicaciones de por qué se generó la alerta.

IV. Estudio de caso

Como ejemplo de aplicación, se implementa la metodología utilizando el conjunto de datos *InSDN* [7]. Se procesan los datos, se entrenan varios modelos y se analizan los resultados obtenidos tanto en rendimiento como en interpretabilidad.

A. Comprensión de los datos

El conjunto de datos InSDN refleja las complejidades y vulnerabilidades de las redes SDN. Consta de un total de 343.939 observaciones, siendo 68.424 de tráfico normal y 275.515 de ataque, lo cual supone aproximadamente una distribución 80%-20%.

El dataset está dividido en tres grupos en función del tipo de tráfico y de las máquinas objetivo: tráfico normal, tráfico de ataque hacia un servidor, y tráfico de ataque hacia una máquina concreta. Se usa la herramienta CICFlowMeter para crear flujos bidireccionales de tráfico de red a partir de archivos PCAP. Esto hace que el primer paquete que encontremos determine el sentido del flujo, hacia adelante (*forward*) o hacia atrás (*backward*). La herramienta permite extraer las características de los flujos existentes en el dataset, proporcionando como salida más de 80 características, que se pueden dividir en 8 grupos:

- Atributos identificadores de red
- Atributos basados en paquetes
- Atributos basados en bytes
- Atributos de tiempo entre llegadas
- Atributos de temporización de flujo
- Atributos de flags TCP
- Atributos descriptores de flujo
- Atributos descriptores de subflujo

Durante el proceso de localización de posibles anomalías en los datos, en primer lugar se comprueban las filas repetidas y las filas que siempre presentan el mismo valor, y se eliminan. En el caso de este conjunto de datos, la lista es considerablemente larga: *Fwd PSH Flags*, *Fwd URG Flags*, *CWE Flag Count*, *ECE Flag Cnt*, *Fwd Byts/b Avg*, *Fwd Pkts/b Avg*, *Fwd Blk Rate Avg*, *Bwd Byts/b Avg*, *Bwd Pkts/b Avg*, *Bwd Blk Rate Avg*, *Fwd Seg Size Min*.

Otro tipo de anomalía en los datos son los valores nulos o inexistentes. En este caso, sólo existe un valor nulo, el número de puerto en el protocolo *ICMP*, algo totalmente normal y que no representa ningún problema.

La siguiente comprobación es la validación de los rangos de valores de las características. Salvo las características categóricas, el resto deberían ser de tipo numérico y positivas. Sin embargo, existen valores negativos: *Init Fwd Win Byts*, *Init Bwd Win Byts*, *Flow IAT Min*, *Bwd IAT Min*, *Bwd IAT Tot*, *Bwd IAT Mean*, *Bwd IAT Max*, *Flow Duration*, *Flow IAT Max*, *Flow IAT Mean*, *Flow Pkts/s*. Entre ellas, destacan *Init Fwd Win Byts* e *Init Bwd Win Byts*, que representan el número de bytes en la ventana de control de flujo de TCP en ambos sentidos. En la primera de ellas, todos los valores son negativos e iguales (-1), mientras que en la segunda, un 45% de los datos son negativos y de valor invariable (-1).

Otro caso a destacar son los flags TCP. Se ha verificado que los valores distintos a 0 de las características relacionadas con los *flags* TCP sólo aparecen con este protocolo y no con UDP o ICMP. Sin embargo, aunque todos los valores con flujos TCP varían entre 0 y 1, en el conjunto de datos se detectaron incongruencias. Se analizaron los ficheros en formato PCAP de los que se obtuvieron las definiciones de los flujos y se detectó la presencia de flags que no aparecían en los flujos definidos. De hecho, se probaron distintas versiones de la herramienta CICFlowMeter, obteniendo con todas ellas el mismo problema.

A raíz de lo indicado, se decidió prescindir de estas características que representan contadores de los flags TCP: *FIN Flag Cnt*, *SYN Flag Cnt*, *RST Flag Cnt*, *PSH Flag Cnt*, *ACK Flag Cnt* y *URG Flag Cnt*.

B. Preparación de los datos

En la subfase de construcción de características dentro de la fase de preparación de los datos, se realizan las transformaciones necesarias en los datos para extraer de ellos más información de la que pueden representar a simple vista, y se transforman ciertas características para poder ser usadas por los diferentes algoritmos.

Entre todas las características cabe mencionar especialmente la característica objetivo o *target*. El atributo

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Label del dataset InSDN utiliza cadenas de caracteres para identificar el tráfico normal o el tipo de ataque en cuestión. Por tanto, se hace necesario convertir estos valores en datos que los algoritmos puedan emplear. Así, por ejemplo, si se trata de clasificación binaria, se definirá el valor 0 para el tráfico normal, y el valor 1 para cualquier tipo de ataque.

En el conjunto de datos de InSDN también existe el atributo *Timestamp*, que representa la marca temporal del flujo. Si bien puede suponer un grave problema de sobreajuste si se emplea sin modificar su codificación, puede ser muy útil para identificar patrones temporales que aparecen en distintos ataques. Para representar el carácter cíclico de este atributo (con distintas granularidades), es necesario transformarlo. Para esta transformación a menudo se emplean el seno y el coseno.

Otro de los cambios a realizar es sobre los puertos origen y destino, en este caso representados como números enteros en los atributos *Src Port* y *Dst Port*. De no ser tratados, estos atributos pueden conllevar un sobreajuste en el modelo. Una aproximación simple consiste en agrupar los puertos según las categorías de la IANA [8]: puertos bien conocidos (0 - 1023), registrados (1024 - 49151) y dinámicos o privados (49152 - 65535). Esta aproximación es muy genérica, por lo que puede ser interesante considerar los puertos de ciertos servicios para estudiar su comportamiento. En el caso de InSDN, se podrían utilizar los puertos 53 (DNS), 80 (HTTP), 137 (NetBIOS-dgm), 138 (NetBIOS-ssn), 443 (HTTPS) y 3306 (MySQL), por ejemplo. La técnica empleada para realizar la transformación suele ser *Label Encoding* [9]. Este método consiste en dar un valor numérico diferente a cada posible nivel del atributo categórico.

Con respecto a las direcciones IP existentes, los atributos *Src IP* y *Dst IP*) contienen las direcciones de origen y destino de los distintos flujos. A menudo, se parte de una cadena de caracteres que se necesita codificar de nuevo para que los algoritmos puedan identificar las diferentes redes. El mero hecho de prescindir de los bits de host en las direcciones evita un posible sobreajuste y consigue el objetivo mencionado de permitir agrupar los dispositivos que pertenecen a la misma red, especialmente útil cuando se trata del análisis del tráfico de una organización con direccionamiento privado. Para el direccionamiento público, se puede generalizar con un valor como "wan", o hacer uso de *blacklists* externas y optar por aproximaciones algo más elaboradas. Para lograr esta nueva codificación se puede emplear la técnica de *OneHot Encoding* [9]. *Label Encoding* tiene una desventaja, pues los valores numéricos pueden ser malinterpretados por los algoritmos creyendo que existe una jerarquía u orden entre ellos. *OneHot Encoding* soluciona este problema, ya que con este método cada categoría es convertida en un atributo, a la que se le asigna un valor 0 o 1. Cada muestra tendrá un 1 como valor para el atributo que se corresponde con la categoría a la que pertenece, y 0 en el resto. Pero esto tiene también un problema, y es el elevado número de columnas que se añaden, lo que puede perjudicar a la hora de construir los modelos, y provocar que se opte por la alternativa de *Label Encoding*.

C. Modelado

Durante esta fase, pueden construirse diversos modelos supervisados de clasificación con el objetivo de detectar tráfico malicioso en el conjunto de datos *InSDN*. Entre las opciones más recomendables se encuentran algoritmos representativos de distintas familias del aprendizaje automático, como *Decision Tree*, *Random Forest*, *SVM*, regresión logística y algoritmos de gradiente aumentado (*XGBoost*). Estos modelos se pueden implementar mediante la librería *scikit-learn*, salvo en el caso de *XGBoost*, que se puede manejar con su paquete específico.

La partición del conjunto de datos podría realizarse mediante validación cruzada estratificada de k particiones (e.g. $k = 10$), garantizando una distribución representativa de las clases. Para evitar el sobreajuste y maximizar el rendimiento predictivo, convendría realizar un ajuste de hiperparámetros utilizando una búsqueda en rejilla (*GridSearchCV*) con la métrica *F1-score* como criterio de optimización, especialmente adecuado en contextos de desbalanceo de clases.

Algunos de los hiperparámetros que se podrían considerar para optimización son:

- **Random Forest:** *n_estimators*, *max_depth*, *min_samples_split*.
- **SVM:** parámetro de penalización *C*, tipo de núcleo (*kernel*) y parámetro *gamma*.
- **XGBoost:** *learning_rate*, *n_estimators*, *max_depth*, *colsample_bytree*.

Previo al entrenamiento, sería recomendable estandarizar los datos con técnicas como *StandardScaler*, en particular para modelos sensibles a la escala (SVM, regresión logística). En el caso de modelos basados en árboles, este paso podría omitirse sin afectar negativamente al rendimiento.

D. Evaluación

La evaluación del rendimiento de los modelos puede realizarse mediante múltiples métricas, como las descritas en la sección E, para captar tanto la capacidad de detección como la precisión general, especialmente en situaciones de desbalanceo de clases.

Además, podría evaluarse la interpretabilidad del modelo analizando la importancia de las características (e.g. a través del atributo *feature_importances_* en modelos con árboles). Características como *Flow Duration*, *Total Fwd Packets* o el puerto de destino codificado podrían destacar por su capacidad discriminativa.

También se podría incorporar un análisis temporal de robustez, aplicando los modelos sobre subconjuntos de datos generados en diferentes franjas horarias o días. Esta evaluación permitiría comprobar si los modelos conservan su capacidad predictiva ante cambios en la distribución temporal de los flujos.

En términos de visualización, se podrían utilizar matrices de confusión, curvas ROC y gráficos de importancia

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de variables para respaldar las conclusiones cuantitativas con una interpretación visual clara.

E. Despliegue

Python es uno de los lenguajes de programación más populares, y se está convirtiendo en una de las mejores opciones para el desarrollo de proyectos de inteligencia artificial.

Esta popularidad se debe a la amplia selección de librerías y *frameworks* que posee. Además, este lenguaje destaca por su código conciso, legible y por la agilidad que ofrece su sintaxis simple. Una de las librerías más comunes para el empleo de algoritmos de aprendizaje máquina es *Scikit-Learn*.

Por otra parte, PySpark es un API de Python para usar en Apache Spark (que permite un despliegue distribuido), facilitando el desarrollo de programas eficientes gracias a las librerías que permite utilizar, como puede ser Spark-ML. Entre las grandes ventajas de PySpark está el soporte de uso con otros lenguajes como Scala, Java y R, facilidad de trabajo con Resilient Distributed Datasets (RDDs) o con DataFrames (para simplificar la construcción de pipelines) y su gran rapidez en comparación con otros frameworks de procesamiento de datos.

Sin embargo, la implementación de los algoritmos en Scikit-learn y Spark-ML presenta sus diferencias, incluyendo no sólo cambios en eficiencia, sino también en los parámetros de los algoritmos, entre otros. Esto implica que conclusiones de la fase de modelado y evaluación en un entorno centralizado, pueden no ser ciertas en el entorno distribuido en el que tiene lugar el despliegue.

V. Conclusiones

La metodología propuesta ofrece, en primer lugar, una guía detallada para el tratamiento completo de datos de red, abarcando desde su captura y limpieza hasta su explotación analítica mediante modelos de clasificación. Este enfoque garantiza una mayor coherencia metodológica en proyectos académicos que tradicionalmente han carecido de una estructura clara y homogénea.

En segundo lugar, se constata que la aplicación de prácticas rigurosas en la fase de preparación de los datos —incluyendo la eliminación de atributos inconsistentes, la transformación semántica de variables categóricas y el tratamiento del desbalanceo de clases— resulta determinante para el rendimiento final del sistema. La elección de codificaciones adecuadas y la incorporación de variables derivadas como las transformaciones temporales o categorizaciones de puertos, permite una representación más rica y significativa del comportamiento de red.

En tercer lugar, los resultados obtenidos a partir del estudio de caso evidencian que los modelos como *Random Forest* y *XGBoost*, constituyen alternativas altamente efectivas, ofreciendo no sólo altos valores de precisión y F1-score, sino también una buena interpretabilidad en términos de importancia de características. La validación cruzada es una técnica imprescindible para obtener estimaciones robustas del rendimiento, y su integración en *pipelines* reproducibles refuerza la trazabilidad del proceso.

Asimismo, la evaluación a través de métricas múltiples —incluyendo AUC-ROC y análisis de errores— proporciona una visión más completa del comportamiento del modelo y permite identificar puntos críticos de mejora. Se destaca también la importancia de considerar la robustez temporal y la capacidad de generalización del modelo ante flujos de red con dinámicas cambiantes.

En conjunto, la metodología propuesta permite establecer un marco de trabajo coherente, reproducible y escalable para el desarrollo de proyectos académicos en ciberseguridad. Su aplicación no sólo facilita la adquisición de competencias técnicas avanzadas en el tratamiento de datos y construcción de modelos de IA, sino que también promueve la adopción de buenas prácticas alineadas con estándares del ámbito profesional.

Agradecimientos

Este trabajo ha sido realizado en el CITIC, en el marco del proyecto PID2023-150794OB-I00, financiado por el Ministerio de Ciencia, Innovación y Universidades (MICIU) y la Agencia Estatal de Investigación (AEI) /10.13039/501100011033, y cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), Unión Europea. El CITIC, acreditado como centro de excelencia del Sistema Universitario Gallego y miembro de la Red CIGUS, cuenta además con el apoyo de la Consellería de Educación, Ciencia, Universidades e Formación Profesional de la Xunta de Galicia, cofinanciado por la Unión Europea a través del programa FEDER Galicia 2021–2027 (Ref. ED431G 2023/01).

Referencias

[1] S. Ness, V. Eswarakrishnan, H. Sridharan, V. Shinde, N. V. P. Janapareddy, and V. Dhanawat, "Anomaly detection in network traffic using advanced machine learning techniques," *IEEE Access*, vol. 1, 2025. [Online]. Available: https://doi.org/10.1109/access.2025.3526988

[2] C. Schröer, F. Kruse, and J. M. Gómez, "A Systematic Literature Review on Applying CRISP-DM Process Model," *Procedia Computer Science*, vol. 181, pp. 526–534, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050921002416

[3] SAS Institute, "Data mining: Qué es y cómo funciona," disponible en: https://www.sas.com/es_es/insights/analytics/data-mining.html. Último acceso: 10-06-2025.

[4] F. Foroughi and P. Luksch, "Data science methodology for cybersecurity projects," arXiv, Preprint, mar 2018, arXiv preprint. Disponible en: https://arxiv.org/abs/1803.04219.

[5] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI Magazine*, vol. 17, no. 3, pp. 37–54, 1996.

[6] SAS Institute Inc., *SAS Enterprise Miner 5.3: User's Guide*, SAS Institute Inc., Cary, NC, 2003, describes the SEMMA methodology: Sample, Explore, Modify, Model, Assess.

[7] M. S. Elsayed, N.-A. Le-Khac, and A. D. Jurcut, "Insdn: A novel sdn intrusion dataset," *IEEE Access*, vol. 8, pp. 165 263–165 284, 2020.

[8] IANA, "Service name and transport protocol port number registry." junio 2025, disponible en: https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml. Último acceso: 10-06-2025.

[9] S. Raschka, *Python Machine Learning*. Birmingham, UK: Packt Publishing Ltd, 2015.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Evaluación de sistemas de bajo coste para medir consumo energético de algoritmos criptográficos en microcontroladores

Álvaro Larrey-Gómez[1,4,5], Javier Martínez-Cesteros[1,3], Guillermo Azuara Guillén[2,4,5], Carlos Medrano-Sánchez[2,3,6].

[1]Departamento de Ingeniería Informática e Ingeniería de Sistemas.
[2]Departamento de Ingeniería Electrónica y Comunicaciones.
[3]Grupo EduQTech, [4]Grupo CeNIT.
[5]Instituto de Investigación en Ingeniería de Aragón (I3A).
[6]Instituto de Investigación Sanitaria de Aragón.

Escuela Universitaria Politécnica de Teruel, Universidad de Zaragoza.
C/Atarazana, nº2, 44003 Teruel
{alarrey, javimzcs, gazuara, ctmedra}@unizar.es

El uso de algoritmos criptográficos en microcontroladores para aumentar la seguridad plantea desafíos relacionados con la optimización del consumo energético, especialmente en aplicaciones alimentadas por baterías. Este trabajo en proceso presenta un plan de evaluación comparativa de sistemas de bajo coste disponibles en el mercado para la medición del consumo energético en aplicaciones basadas en microcontroladores. Se llevarán a cabo pruebas en distintos escenarios, desde cargas resistivas estáticas a casos reales de ejecución de algoritmos criptográficos. Se analizará la precisión, reproducibilidad y capacidad de distinguir consumos, comparando resultados frente a un instrumento de referencia. Como resultados preliminares, se ha realizado un estado del arte identificando cuatro medidores de bajo coste. Se han realizado pruebas con dos de ellos, Power Monitor3 y ACS712-5B, tanto con cargas fijas como con algoritmos criptográficos distintos (RSA y AES con diferentes configuraciones) ejecutados en un STM32, comparando los resultados con las mediciones de un Fluke115.

***Palabras Clave*- Ciberseguridad, microcontrolador, consumo energético, criptografía.**

I. Introducción

El consumo energético es un factor determinante cuando se implementan algoritmos criptográficos en microcontroladores, especialmente en sistemas que operan con recursos limitados y alimentación mediante baterías. Para caracterizar y optimizar dicho consumo, existen en el mercado herramientas de medición con distintos niveles de precisión y coste, utilizadas tanto en entornos industriales como en laboratorios de investigación y desarrollo.

Muchos de los dispositivos comerciales tienen un coste muy elevado y/o tienen una funcionalidad muy específica, lo que puede ser inasumible en algunas ocasiones o simplemente no se requiere tanta precisión. Estas soluciones son normalmente empleadas para realizar ataques de canal lateral (side-channel) e ingeniería inversa como [1] o también vemos sistemas orientados a la medición del consumo en otros aspectos, como la transmisión de datos [2]. Podemos destacar la plataforma open-source Chip Whisperer [3], que aun siendo relativamente económica (algunos cientos de euros), puede resultar excesiva para estudios que únicamente requieran medir consumos sin tanta precisión y a frecuencias no tan elevadas.

La contribución de este trabajo en proceso consiste en validar la utilidad de medidores de consumo de bajo coste para optimizar la implementación de criptografía en microcontroladores, entre otros. Se han diseñado pruebas que modelan distintos escenarios de consumo. Se concluirá con una comparativa de aquellos medidores recomendados basándose en dichas pruebas. Este análisis comparativo aporta valor práctico significativo para desarrolladores e investigadores que buscan soluciones económicas y confiables para comparar el consumo energético de diferentes implementaciones de algoritmos criptográficos en microcontroladores.

En los siguientes apartados se presenta un breve estado del arte no exhaustivo respecto a medición de consumo en los sistemas mencionados, la metodología de pruebas a seguir para los distintos sensores a comparar, algunos resultados preliminares y el trabajo futuro previsto, respectivamente.

II. Estado del arte

De acuerdo con lo expuesto con anterioridad, la medición precisa del consumo energético en

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

microcontroladores es esencial para optimizar el diseño de dispositivos portátiles y maximizar su autonomía, especialmente cuando se integran algoritmos criptográficos que incrementan el consumo energético.

Poder comparar el consumo de varios algoritmos de cifrado puede ser un factor a tener en cuenta a la hora de elegir entre posibles candidatos que cumplan otros requerimientos relativos a la seguridad, el tiempo de procesado o la necesidad de recursos.

La técnica más extendida para este fin consiste en medir simultáneamente la corriente y la caída de tensión sobre el dispositivo, permitiendo calcular la potencia instantánea consumida. Este método se basa en principios físicos simples y proporciona resultados directos y fiables.

Existen otras técnicas de estimación de estas magnitudes físicas que se apoyan en alguna característica específica del dispositivo de estudio, habitualmente la frecuencia de alimentación en corriente alterna [4, 5] o el tiempo de autonomía en sistemas con baterías.

No obstante, el modelo más utilizado al tratar con microcontroladores sin periféricos es considerar una única carga en corriente continua, mayoritariamente resistiva y variable en el tiempo dependiendo de las operaciones realizadas por dicho microcontrolador. En dicho modelo la potencia se calcula como el producto de la caída de tensión en los bornes de alimentación del microcontrolador y la corriente que los atraviesa, lo que únicamente requiere de dos mediciones en el mismo instante de tiempo.

Aunque nuevas tecnologías de fabricación permitan crear sensores de corriente con un abanico más amplio de principios físicos [6], estos ya se conocen y emplean hace tiempo [7].

Hay varios tipos de sensores basados en distintas propiedades físicas, entre estos los más usuales son los basados en resistencia de Shunt, efecto Hall [8], TMR (Tunneling Magneto-Resistante) [9] y GMR (Giant Magneto-resistance) [10].

Las técnicas de medición que utilizan sensores magnéticos no son intrusivas, ya que no requieren de una modificación eléctrica de la carga a estudiar y permiten realizar lecturas precisas, pero por contra requieren de un proceso previo de caracterización del sensor, colocación precisa y estable cerca del conductor y aislarlo de posibles interferencias electromagnéticas. Además, la linealidad y precisión de estos sensores pueden verse afectadas por factores como la histéresis o el acoplamiento inductivo, lo que obliga a los fabricantes a incorporar sistemas de corrección y calibración en los propios dispositivos. Es por esto que su uso no es tan extendido como otras alternativas más robustas. En su lugar la mayoría de los sensores y multímetros accesibles en el mercado hacen uso de otra técnica de medición de corriente: mediante resistencia de Shunt.

La medición con resistencia de Shunt sigue siendo la técnica más utilizada en aplicaciones de bajo voltaje y corriente, debido a su simplicidad, linealidad y bajo coste. Una resistencia de Shunt se consigue colocando en serie con la carga objeto de estudio una resistencia con un valor conocido y relativamente pequeño. La caída de tensión en el componente es directamente proporcional a la corriente que lo atraviesa, permitiendo lecturas precisas y rápidas en un amplio rango de valores. La principal desventaja de esta técnica es que la resistencia debe estar bien caracterizada y en algunos casos puede ocupar mucho espacio dependiendo de la potencia del sistema en cuestión. Sin embargo, esta desventaja al utilizar resistencia Shunt en aplicaciones de electrónica digital con voltajes y corrientes relativamente pequeñas no supone un problema, pues lo habitual es utilizar resistencia Shunt en formato SMD (montaje superficial) de entre 1 y 0,01Ω, buscando minimizar la potencia disipada manteniendo una caída de tensión detectable por un conversor analógico-digital.

Es frecuente el uso de circuitos de medición ad-hoc como herramienta fiable para conseguir resultados en áreas relacionadas. Por ejemplo, en [11] se usa una resistencia Shunt como referencia para comparar la eficacia de tres sensores magnéticos en la medición de corrientes.

La plataforma open-source Chip Whisperer [3] y similares utilizan una resistencia Shunt junto con una FPGA de altas prestaciones. Esto permite análisis de consumo de alta resolución, sincronización con el reloj del microcontrolador y más, orientados a ataques de canal lateral y estudios de ingeniería inversa.

En [12] se detalla el diseño y puesta en marcha de un medidor para después realizar pruebas de consumo a un sistema operativo ligero, como TinyOS. El circuito de medición demuestra completa linealidad y, por tanto, facilidad de calibración.

Otro ejemplo de uso de resistencia Shunt se propone en [13], donde se modela tanto el consumo de ciertas operaciones de cálculo como el de los periféricos conectados al microcontrolador, midiendo el consumo total mientras realiza distintas tareas.

III. METODOLOGÍA

Los medidores de corriente que se compararán, con distinto rango de coste y/o tecnología de medición, mostrados en la Tabla I son: USB Power Monitor V3, Shunt INA219, Hall ACS712 y Power Profiler Kit II. Como referencia se utilizará un polímetro Fluke 115 calibrado.

Tabla I
CARACTERÍSTICAS DE INTERÉS PROPORCIONADAS POR EL FABRICANTE

Medidor	Rango	Resolución	Error	Freq.
Fluke115	6A- 10A	1mA-0mA	1,0% (I) +3cuentas*	4Hz
Power Monitor3	3V-20V 4A	0,1V	<±1% (V) <±1% (I)	3Hz
INA219	±2,5A	1mA**	<±0,2%(I)	500kHz
ACS712-05B	±5A 185mV/A	ADC externo	Vruido <±1,5%(V)	80kHz
Power ProfilerII	200nA 1A	0,1mA 1mA	- -	100kHz

(V) Medición de voltaje, (I) Medición de corriente.
* Fluke 115 aplica 6000 cuentas.
** Depende del valor de la resistencia, normalmente 0.1Ω.

El medidor Power Monitor V3 fabricado por PortaPow se conecta fácilmente con cables USB y muestra las mediciones a través de una pantalla que será necesario anotar manualmente. Además de voltaje y amperios este

dispositivo muestra también una métrica acumulativa en mAh.

El INA219 es un circuito integrado fabricado por Texas Instruments que transmite las mediciones por bus I2C. Requiere de montaje, aunque lo habitual es encontrarlo ya soldado en un pequeño circuito con todos los componentes necesarios para su funcionamiento, entre estos una resistencia de Shunt de 0.1Ω.

En el caso del sensor de efecto Hall ACS712, es necesario utilizar un conversor analógico-digital para medir la salida en mV asociada al campo magnético medido, lo que implica conocer parámetros adicionales proporcionados por el fabricante, Allegro Microsystems. La sensibilidad del ACS712-05B es lineal (185mV/A) y supuesta invariante con la temperatura (rango -40ºC a 25ºC, 0.054mV/A/ºC).

La plataforma Power Profiler Kit II [14] de Nordic Semiconductor transfiere datos y alimentación con un PC por USB y a su vez energiza el dispositivo de medida, con una tensión regulable entre 0.8 y 5V y picos de corriente de hasta 1A. Nordic proporciona software para configurar la herramienta y visualizar resultados. También se pueden utilizar sus ocho entradas digitales como analizador lógico o disparador, sincronizando las mediciones con el microcontrolador objeto de estudio.

Las pruebas comparativas con instrumentos de referencia se realizarán bajo diversas condiciones. Por un lado, con cargas resistivas fijas (consumo estático) y, por otro lado, con una aplicación concreta, el envío de datos encriptados de un sensor de presión distribuido [15], con varias opciones de algoritmos de cifrado de datos [16] para compararlos. Este tipo de sensores se utilizan en el ámbito médico y deportivo, por lo que el cifrado de los datos resulta de interés. Los resultados se analizarán en términos de error con respecto a los instrumentos de referencia en cada uno de los escenarios considerados, aplicándose también test estadísticos (t-test) para comprobar la significación de las diferencias.

Otro aspecto bajo consideración es el de la reproducibilidad de los resultados. En dispositivos de bajo coste es más habitual encontrar diferencias entre unidades, incluso que existan variantes (copias) de distintos fabricantes, por lo que se establece también un conjunto de medidas para compararlos entre sí.

Otra prueba interesante es la capacidad de los medidores para distinguir entre distintos algoritmos de cifrado de datos. Se estudiarán los datos agrupándolos por dispositivo de medida. Después se analizarán las diferencias entre el consumo para varios tipos y configuraciones de cifrado. Este análisis es importante porque pudiera suceder que el equipo económico no fuese fiable en términos absolutos en sus medidas, pero que sí pueda distinguir diferentes niveles de consumo en función de los algoritmos que se estén procesando. Los medidores económicos podrían ser útiles si sus medidas relativas fuesen fiables para determinar qué algoritmos o configuraciones de uso en un microcontrolador consumen menos. Se utilizarán también test estadísticos para comprobar las diferencias de las medidas de un mismo equipo en diferentes situaciones de cifrado. En las pruebas preliminares se ejecutan 4 variantes de algoritmos de cifrado: RSA con clave de 2048 bits y AES con clave de 256 bits en los modos Electronic codebook (ECB), Cipher block chaining (CBC), Cipher feedback 128 bits-block (CFB). Se cifra un texto codificado como UTF-8 de 128 caracteres, enviado por serie a un ordenador encargado de recibirlo y descifrarlo. Este proceso se repite cada 500ms.

Finalmente, se harán pruebas con distintos tipos de transitorios, con cargas periódicas a distintas frecuencias y otras con picos de consumo no periódicos, para conocer las limitaciones derivadas de la frecuencia de muestreo de cada medidor. Estas se realizarán en última instancia una vez se descarten medidores en las fases previas.

IV. Resultados

Hasta ahora se han realizado pruebas con el medidor Power Monitor V3 y ACS712 frente a la referencia (Fluke 115), tanto en medidas con carga fija en Tabla II, como con un microcontrolador (STM32F103C8T6) y distintos algoritmos de cifrado en Tabla III. El sensor ACS712, al requerir de un muestreador externo, es muy susceptible a cambios de montaje (cableado, distancia del circuito de carga y muestreador, etc). Tampoco es capaz de medir la tensión en la carga y es muy susceptible a variaciones de la tensión de alimentación.

Tabla II
Intensidades a través de carga fija 100.6 ±5%Ω, promedio y desviación estándar de treinta pruebas (una por minuto)

Medidor	Tensión (V)	Intensidad (A)
Fluke115 (Referencia)	4,986 (±0,042)	0,049 (±0,002)
Power Monitor3	5,104 (±0,080)	0,046 (±0,005)
ACS712	----*	0,055 (±0,009)

* ACS712 es un sensor únicamente de intensidad, no mide tensión.

Tabla III
Intensidades consumidas durante el cifrado y envío de la información. Promedio de treinta pruebas (una por minuto)

Medidor	AES-ECB	AES-CBC	AES-CFB128	RSA-2048
Fluke115*	0,027A	0,027A	0,026A	0,028A
Power Monitor3*	0,02A	0,02A	0,02A	0,02A
ACS712	0,040A (±0,011)	0,048A (±0,017)	0,046A (±0,006)	0,044A (±0,017)

* Desviación estándar ±0,00A.

El resultado de t-test para los treinta puntos medidos con carga fija en Tabla II muestra una diferencia significativa entre los instrumentos de medida y la referencia: PowerMonitorV3 p-value 0.036 < 0.05. ACS712 p-value $1 \times 10^{-13} < 0.05$.

Es posible distinguir consumos entre algunos de los algoritmos de cifrado con las medidas efectuadas con el Fluke 115, a diferencia de las realizadas con Power Monitor 3 que no dispone de suficiente precisión, o el ACS que no ofrece suficiente seguridad en las medidas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Se muestra un montaje en la Figura 1, a la izquierda el medidor Fluke115 y en la esquina superior derecha el Power Monitor V3.

Fig. 1. Imagen del montaje

V. Conclusiones y líneas futuras

Existen dispositivos de medida en una amplia gama de precios que permiten obtener el consumo instantáneo de sistemas basados en microcontroladores. Una tarea inicial, no menos importante, es la selección de los dispositivos de medición que se van a analizar, atendiendo a su precio, disponibilidad y especificaciones técnicas, proponiendo la selección indicada en la Tabla I.

Se ha realizado un análisis preliminar con dos sensores: el Power Monitor V3, por ser compacto y fácil de usar, y el ACS712, que requiere un montaje electrónico dedicado.

Respecto al Power Monitor V3, atendiendo a los resultados del t-test, en concreto a la métrica p-value, se determina que hay diferencias significativas entre las medidas de este instrumento frente al de referencia (Fluke 115). Además, también se concluye que no es suficientemente preciso para poder discernir entre algoritmos de cifrado.

Los resultados del ACS712 muestran poca exactitud y una variabilidad alta por encima de la de referencia. El p-value evidencia que difiere de la referencia. Al requerir de un muestreador externo es muy susceptible a la calidad del montaje (cableado, distancia al circuito de muestreo, etc). Utilizarlo en la práctica requiere de conocimientos más específicos que otros sensores.

El trabajo continuará con el análisis preliminar de otras unidades (INA219, Power Profiler II). Cuando se lleven a cabo las pruebas descritas en la metodología, se presentarán los resultados completos y recomendaciones sobre cada uno de ellos.

Agradecimientos

Esta investigación se ha realizado en el marco de los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation). Parcialmente financiado por los proyectos: Proyecto Estratégico C077.23, desarrollado bajo el amparo del con venio entre INCIBE y la Universidad de Zaragoza; y NEWLAN (PID2022-136476OB-100) del gobierno de España (financiado por MICIU/AEI/10.13039/ 501100011033/FEDER EU, ERDF/EU). También por el Departamento de Educación, Ciencia y Universidades del Gobierno de Aragón: Grupo EduQTech (Ref T29_23R) y Grupo CeNIT (Ref T31_23R).

Referencias

[1] [M.A. González de la Torre, V. Sarasa Laborda, L. Hernández-Álvarez, I. Morales Sandoval, L. Hernández Encinas; “Ataques por canal lateral contra AES mediante correlación de consumo de potencia”, en IX Jornadas Nacionales de Investigación en Ciberseguridad (JNIC’2024), pp. 420-427, 2024. https://hdl.handle.net/11441/160773

[2] G. L. Álvarez-Rosado, K. A. Martínez-Hernández, M. A. Camarillo-Ramos, V. Quintero-Rosas, A. Díaz-Ramírez, and R. López-Avitia, “Energy consumption of an internet of things development board”, Rev. Ci. Tec., vol. 5, no. 4, p. e234, pp. 376-386. 2022. https://doi.org/10.37636/recit.v54363375.

[3] Chip Whisperer https://eu.mouser.com/manufacturer/newae-technology/ Última visita en marzo de 2025.

[4] Haberman M.A., Spinelli E.M.: “A Noncontact Voltage Measurement System for Power-Line Voltage Waveforms”, en IEEE Transactions on Instrumentation and Measurement, 69 (6), art. no. 8755857, pp. 2790 – 2797, 2020.

[5] Shenil P.S., George B.: “Nonintrusive AC Voltage Measurement Unit Utilizing the Capacitive Coupling to the Power System Ground”, en IEEE Transactions on Instrumentation and Measurement, 70, art. no. 9293365, 2021.

[6] Lutzen H., Müller J., Kaminski N.: “A Review of Current Sensors in Power Electronics: Fundamentals, Measurement Techniques and Components to Measure the Fast Transients of Wide Bandgap Devices”, en 2023 25th European Conference on Power Electronics and Applications, EPE 2023 ECCE Europe, 2023.

[7] Ziegler S., Woodward R.C., Iu H.H.-C., Borle L.J.: “Current Sensing Techniques: A Review”, en IEEE Sensors Journal, 9 (4), art. no. 4797906, pp. 354 – 376, 2009.

[8] Crescentini M., Syeda S.F., Gibiino G.P.: “Hall-Effect Current Sensors: Principles of Operation and Implementation Techniques”, en IEEE Sensors Journal, 22 (11), pp. 10137 – 10151, 2022.

[9] Vala S.S., Basit Mirza A., Luo F.: “A TMR-based Integrated Current Sensing Solution for WBG Power Modules”, en Conference Proceedings - IEEE Applied Power Electronics Conference and Exposition - APEC, pp. 2398 – 2402, 2024.

[10] Mușuroi C., Oproiu M., Volmer M., Firastrau I.: “High sensitivity differential giant magnetoresistance (gmr) based sensor for non-contacting dc/ac current measurement”, en Sensors (Switzerland), 20 (1), art. no. 323, 2020.

[11] Adamo F., Attivissimo F., Di Nisio A., Carducci C.G.C., Spadavecchia M., Guagnano A., Goh M.K.: “Comparison of Current Sensors for Power Consumption Assessment of Wireless Sensors Network Nodes”, en 2017 IEEE International Workshop on Measurement and Networking, M and N 2017 - Proceedings, art. no. 8078390, 2017.

[12] Zhang K., Hu S., Liu W.: “Power Consumption Benchmarking of TinyOS Applications with A Dedicated Measurement Adapter”, en 2023 5th International Conference on Circuits and Systems, ICCS 2023, pp. 283 – 287, 2023.

[13] Berthou G., Marquet K., Risset T., Salagnac G.: “Accurate Power Consumption Evaluation for Peripherals in Ultra Low-Power embedded systems”, en GIoTS 2020 - Global Internet of Things Summit, Proceedings, art. no. 9119593, 2020.

[14] Nordic Semiconductors, Power Profiler Kit II, https://www.nordicsemi.com/Products/Development-hardware/Power-Profiler-Kit-2, última visita en marzo de 2025.

[15] C. Medrano-Sánchez, R. Igual-Catalán, V. H. Rodríguez-Ontiveros and I. Plaza-García, "Circuit Analysis of Matrix-Like Resistor Networks for Eliminating Crosstalk in Pressure Sensitive Mats," in IEEE Sensors Journal, vol. 19, no. 18, pp. 8027-8036, 15 Sept.15, 2019, doi: 10.1109/JSEN.2019.2918592.

[16] Librería https://github.com/Seeed-Studio/Seeed_Arduino_mbedtls, última visita en marzo de 2025.

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Una arquitectura de votaciones electrónicas: Un anillo para democratizarlos a todos

Álvaro Larrey-Gómez[1,3], Guillermo Azuara Guillén[2,3], José Luis Salazar Riaño[2,3], Julián Fernández-Navajas[2,3], Jordi Alexander Rodríguez-Coello.

[1]Departamento de Ingeniería Informática e Ingeniería de Sistemas.
[2]Departamento de Ingeniería Electrónica y Comunicaciones.
[3]Instituto de Investigación en Ingeniería de Aragón (I3A).

Universidad de Zaragoza.
Edificio Ada Byron, Campus Río Ebro, Calle María de Luna 1, 50018, Zaragoza, España.

{alarrey, gazuara, jsalazar, navajas}@unizar.es

Los sistemas de participación y votación a distancia están ganando popularidad y relevancia, debido a su eficacia en contextos más allá de los electorales. La participación se lleva a cabo con herramientas informáticas y se transmite por redes de comunicación, lo que plantea desafíos en términos de seguridad, privacidad y confianza para los usuarios. En este trabajo en desarrollo, se presenta una propuesta de mejora para un sistema de participación basado en firmas en anillo. Se busca reforzar la seguridad y versatilidad del sistema, permitiendo realizar votaciones y encuestas ajustadas a las necesidades de los organizadores, con más garantías de seguridad y manteniendo las premisas básicas de un sistema de votación/participación. Se está desarrollando una implementación, para más adelante llevar a cabo pruebas de usabilidad y aceptación para distintos perfiles de la población.

***Palabras Clave*- e-votación, seguridad de datos, firmas en anillo, e-voting, security of data, ring signatures.**

I. Introducción

Tras varias décadas, los sistemas de participación a distancia han ganado tracción y permeado en la sociedad por su utilidad en la toma de decisiones y recolección de información a través de encuestas. Han adquirido especial relevancia desde la pandemia de COVID-19.

La característica principal de estos sistemas es que la participación se lleva a cabo con herramientas informáticas y se transmite por redes de comunicación, permitiendo incorporar conceptos telemáticos, criptológicos y de teoría de la información. Sin embargo, esto presenta retos de seguridad y confianza [1, 2].

Existen numerosas propuestas [3], cada una con ventajas e inconvenientes debido a distintos enfoques, necesidades y tecnologías adoptadas. Algunas buscan ser alternativas al sistema electoral a nivel de país [4], mientras que otras son portales fáciles de usar desde un navegador web [5].

Frente a la abundancia de propuestas teóricas, existe una notable falta de implementaciones desplegables. Por ello, este trabajo en desarrollo se centra en una arquitectura pragmática y verificable a través de pruebas en la población.

Mientras que tecnologías como blockchain [6, 7] a menudo introducen complejidades de infraestructura, esta propuesta se centra en utilizar firmas en anillo enlazables, una construcción criptográfica simple de implementar, robusta y autónoma [8].

Brevemente, las firmas en anillo permiten que un miembro anónimo de un grupo firme mensajes de manera indistinguible al resto de miembros del grupo, es decir, donde cualquiera de ellos podría haber firmado el mensaje. Si estas firmas son enlazables, incluyen un campo adicional llamado “linking tag” o etiqueta, único para cada integrante, que permite relacionar firmas emitidas por el mismo manteniendo su anonimato.

En este trabajo en desarrollo, se presenta una propuesta de mejora para un sistema de participación basado en firmas en anillo. Se desarrolla una implementación del sistema y se planifica llevar a cabo pruebas de usabilidad y aceptación para distintos perfiles de la población.

II. Punto de partida y motivación

Se parte de la arquitectura descrita en [9, 10], un modelo bien documentado, para proponer mejoras en la seguridad y resiliencia del sistema.

La Fig. 1. ilustra una implementación de la arquitectura base. Es una aplicación de votaciones centralizada en una única máquina, cuya base de datos almacena todo lo relacionado para su funcionamiento (credenciales de los usuarios, votaciones, censos, resultados, etc.).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

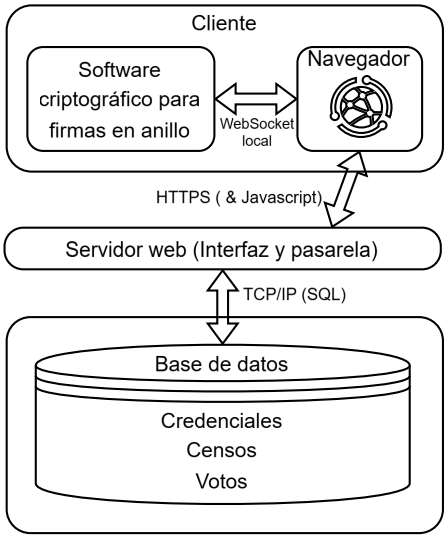


Fig. 1. Esquema de la arquitectura considerada base

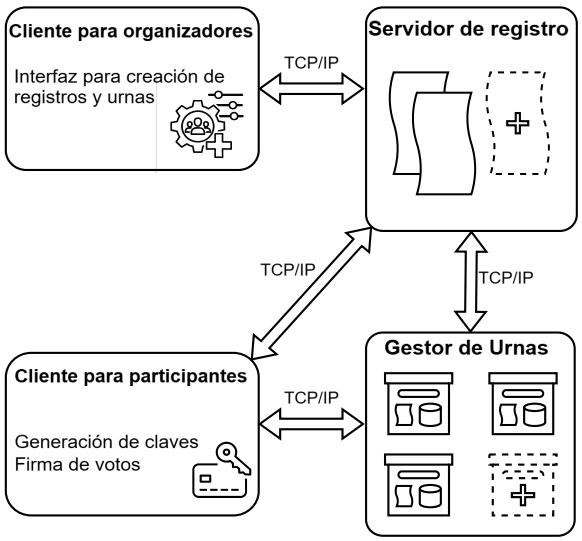


Fig. 2. Esquema de la arquitectura propuesta

Esta arquitectura centralizada presenta varios aspectos ventajosos. Principalmente, que el uso de firmas en anillo enlazables permite que los participantes disfruten de privacidad y anonimato irrevocables. Además, tiene una interfaz de usuario amigable tanto para los organizadores como para los votantes. También, es trivial realizar copias de seguridad y auditorías pues todo se encuentra en una misma base de datos. Es versátil y útil, lo que la hace un punto de partida idóneo para conseguir objetivos más ambiciosos, reforzando la seguridad teniendo en cuenta desde la fase de diseño los aspectos usuales: el principio de mínimo privilegio, el modelado de posibles actores y amenazas, segmentar tanto en red como en servicios y comunicar las partes con interfaces ad-hoc (API), entre otros.

Con la intención de abarcar más escenarios de uso, se desea un sistema todavía más versátil que no sólo permita realizar votaciones, sino que de forma general permita tomar muestras anónimas de un subconjunto de población acotado. Estas muestras pueden ser encuestas de consumo (qué producto tiene mejor acogida), de satisfacción (opinión cualitativa sobre un servicio), etcétera.

En definitiva, se le quiere dar a los organizadores la libertad de crear votaciones y encuestas ajustadas a sus necesidades, de una forma igual de rápida y fácil, con más garantías de seguridad y manteniendo las premisas básicas de un sistema de votación/participación [10].

III. Arquitectura propuesta

A continuación se detalla el funcionamiento, la interacción entre los distintos actores, las ventajas y la implementación de la arquitectura desarrollada.

Atendiendo a la Fig 2, el primer punto a destacar es que consta de cuatro partes diferenciadas: cliente para organizadores, cliente para participantes, servidor de registro y gestor de urnas, con responsabilidades distintas y no solapadas:

Los organizadores pueden crear registros y urnas. Dependiendo de la implementación también pueden asistir en el registro de los participantes.

Los participantes generan pares de claves pública/privada, registran su clave pública y envían votos directamente a la urna deseada.

El servidor de registro genera los parámetros necesarios para conformar un anillo y mantiene las colecciones de claves públicas. Para añadir la clave pública de un participante, se realiza un proceso de identificación de su persona física, que varía según la implementación del sistema.

El gestor de urnas mantiene las distintas votaciones y/o encuestas, que son autocontenidas y prevalecen después del proceso de votación para mantener los resultados públicos. Esto significa que cada urna guarda una copia del censo en el momento de creación, los votos recibidos, los resultados finales y la descripción de la votación, incluyendo el formulario a cumplimentar, fechas de inicio/fin y la configuración adicional elegida por el organizador. Estos datos son públicos en todo momento, excepto los votos emitidos, que se publican al final del proceso junto con los resultados.

Cuando se recibe un voto, se aplica un proceso de validación de la firma y del “linking tag” antes de aceptarlo. No es posible conocer la clave pública/privada que emite la firma, solo comprobar que procede del anillo y detectar si el firmante ha enviado otras firmas anteriormente.

Las urnas pueden tener configuraciones distintas según las necesidades del organizador. Los formularios admiten múltiples campos y opciones a elegir. Los participantes pueden sobrescribir sus votos para mitigar coacción, o se puede guardar la fecha y hora de los votos recibidos. También es posible reutilizar un registro para múltiples votaciones, por rondas o en paralelo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

A. Descripción del proceso de votación

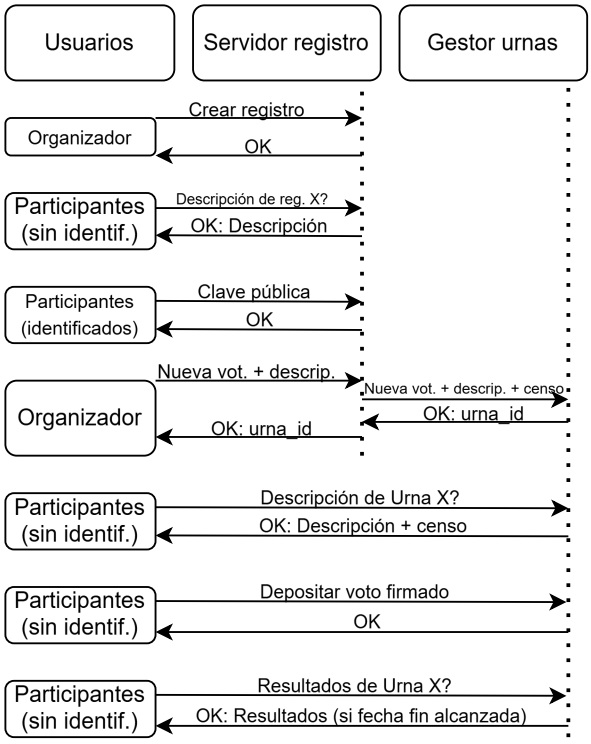


Fig. 3. Flujo de información en un proceso de votación completo

En la Fig. 3 se detallan las comunicaciones necesarias para realizar una votación. Para entender el proceso completo, se deben considerar varios aspectos adicionales:

El organizador introduce la descripción de la votación y las opciones deseadas. Por ejemplo, si un segundo voto de un firmante debe ser rechazado, sobrescrito o añadido a un historial (para detectar tendencias y cambios de opinión). Esta configuración se transfiere a la urna, que debe ser autocontenida y pública.

Los participantes necesitan una descripción inicial del censo para generar localmente su par de claves privada/pública.

No se especifica cómo se comunica a los participantes la dirección IP del registro y la urna, ya que depende de la implementación. El organizador puede compartir esta información de diversas maneras (a través del sistema, mensajería, QR, etc.).

Tampoco se detalla el proceso de identificación del participante, siendo que está condicionado por el sistema de autenticación elegido. En general, el participante debe acreditar su identidad ante una entidad competente, que notificará al servidor de registro.

B. Seguridad y robustez de la arquitectura propuesta

Como se indicó previamente, las distintas partes del sistema tienen responsabilidades distintas y no solapadas, lo que lo hace robusto. En materia de seguridad, una parte comprometida o corrupta puede evitar el correcto funcionamiento del sistema en general, pero no afectar a las otras partes.

A continuación se examinan diversos escenarios en los que un atacante consigue tomar el control de una parte del sistema con el objetivo de manipular resultados o revocar el anonimato:

Un cliente de organizador comprometido puede suplantar a uno legítimo y crear censos y votaciones, pero no puede revocar el anonimato ni manipular resultados, ya que la asociación de la persona física con la clave pública se mantiene privada en el servidor de registro.

Un organizador corrupto no puede influenciar en las votaciones a través del sistema, necesitaría aplicar coerción o coacción por vías externas.

Un cliente de participante comprometido conoce los pares de claves generados en el terminal atacado, permitiéndole suplantar esa identidad en votaciones activas y revocar anonimato en resultados anteriores si se publicaron los “linking tag”. Esto no afecta a otros electores ni impide el proceso de votación.

Un participante corrupto puede intentar forjar sus propias firmas para introducir votos extra. Las urnas descartarían estos votos durante la validación, ya sea porque no corresponden a una clave registrada en el censo, porque el “linking tag” está duplicado o porque este no verifica con la firma.

Un servidor de registro comprometido conoce la asociación de la persona física con su IP y clave pública. Puede crear urnas con censos arbitrarios y emitir más votos. Dado que las urnas son públicas, se pueden auditar los resultados y detectar desajustes. Las personas físicas pueden denunciar que su clave pública no se encuentra en un censo, aunque esto no es fácil de demostrar.

Un gestor de urnas comprometido conoce la dirección IP y el contenido del voto recibido, pero no puede deducir la clave pública asociada a ese voto, manteniendo el anonimato. Puede publicar votaciones con censos y resultados arbitrarios y modificar resultados anteriores. Estos ataques son detectables por cualquier persona que quiera auditar el proceso, cotejando los censos con las listas presentes en el servidor de registro o pidiendo a los electores que comprueben si su voto y “linking tag” coinciden. Para detectar estas manipulaciones en las urnas es conveniente hacer copias de seguridad o utilizar tecnología blockchain como repositorio público.

De manera esencial, si las firmas en anillo requieren parámetros adicionales, un atacante podría diseñar una combinación de parámetros que debilite las firmas, permitiendo conocer el par de claves que las emiten. Este ataque requiere de un servidor de registro o un gestor de urnas comprometido. Como los parámetros son públicos, se pueden auditar e invalidar las votaciones y censos afectados.

C. Contraste con el sistema base

La mejora principal de la arquitectura propuesta frente al sistema base es la separación de responsabilidades. En el modelo base, un administrador comprometido puede

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

vincular votante y voto. En nuestra propuesta el registro conoce la identidad pero no el voto, y la urna únicamente procesa votos anónimos, impidiendo esta correlación maliciosa.

A diferencia del sistema base que es centralizado, el propuesto se encuentra separado en partes conectadas en red, lo que puede dificultar el despliegue y las copias de seguridad a realizar dependiendo de la implementación. Utilizar tecnologías de virtualizado en contenedores ayudaría a mitigar este problema.

Se ha mejorado la versatilidad del sistema para tomar muestras de distintas índoles, como encuestas y opiniones. El organizador puede configurar el nivel de trazabilidad deseado para una o varias votaciones relacionadas, teniendo en cuenta que más trazabilidad implica posibles ataques de revocación del anonimato.

En esta propuesta se aprovechan al máximo las capacidades de las firmas en anillo enlazables, permitiendo el registro con antelación y la emisión del voto de manera espontánea desde diferentes dispositivos y momentos, sin vincular el sufragio con la identidad del votante. Esto mejora la seguridad y confianza del usuario.

Un registro es reutilizable para múltiples urnas, evitando generar nuevas claves para cada votación y permitiendo más opciones de trazabilidad.

Además, los resultados publicados por la urna pueden ser libres de etiqueta ("linking tag"), aumentando la resistencia a la coacción y ataques por correlación, aunque a costa de perder verificabilidad y trazabilidad.

Un mismo sistema puede servir a varios organizadores y grupos de personas. Asimismo, una persona puede estar registrada en varios grupos con claves distintas, evitando ataques de correlación y congruencia.

D. Implementación

La implementación del sistema sigue en desarrollo y está previsto realizar pruebas de usabilidad y aceptación por distintos perfiles de la población.

La estructura segmentada facilita iteraciones y pruebas. Se han integrado conceptos de ciberseguridad desde el inicio, como el saneado de valores de entrada con Regex y límite de peticiones desde una misma IP para mitigar ataques comunes. Las comunicaciones entre las distintas partes del sistema se hacen mediante API REST sobre TLS/HTTPS.

Para mantener la simplicidad del programa hasta obtener los resultados de las pruebas de usabilidad, la identificación de las personas físicas se realiza en persona por el organizador, pulsando un botón en su terminal para permitir el registro. En fases posteriores se integrará con una entidad de identificación como CERES de la FNMT.

IV. Conclusiones y trabajos futuros

Como se ha expuesto, esta arquitectura mejora la resiliencia del sistema de participación, así como la seguridad y privacidad para los usuarios. Proporciona una plataforma para varios entornos y tipos de votación.

Está previsto realizar pruebas de campo en diversos entornos relacionados con población de edad avanzada o sin muchos conocimientos en el manejo de dispositivos electrónicos para constatar la usabilidad y confianza derivada de la interfaz de usuario.

Se estudiará utilizar tecnología blockchain como repositorio estable de anillos y resultados, para facilitar labores de auditoría. También el uso de cifrado homomórfico como capa de privacidad adicional durante el tránsito, almacenaje y recuento de los votos.

Agradecimientos

Los investigadores pertenecen al I3A (Instituto Universitario de Investigación en Ingeniería de Aragón, Universidad de Zaragoza). Esta investigación se ha realizado en el marco de los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation). Parcialmente financiado por los proyectos: Proyecto Estratégico C077.23, desarrollado bajo el amparo del convenio entre INCIBE y la Universidad de Zaragoza; y NEWLAN (PID2022-136476OB-100) del gobierno de España (financiado por MICIU/AEI/ 10.13039/501100011033/FEDER EU, ERDF/EU). También por el Departamento de Educación, Ciencia y Universidades del Gobierno de Aragón, Grupo CeNIT (Ref T31_23R).

Referencias

[1] S. Wolchok, E. Wustrow, J. A. Halderman, H. K. Prasad, A. Kankipati, S. K. Sakhamuri, V. Yagati, and R. Gonggrijp, ''Security analysis of india's electronic voting machines,'' *Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS 2010,* ACM, pp. 1–14, 2010.. DOI: 10.1145/1866307.1866309

[2] T. Kohno, A. Stubblefield, A. D. Rubin, and D. S. Wallach, ''Analysis of an electronic voting system'', *2004 IEEE Symposium on Security and Privacy (S&P 2004)*, IEEE Computer Society, p. 27, 2004. Doi: https://doi.org/10.1109/SECPRI.2004.1301313

[3] Barelli, Riccardo & D'Onghia, Mario & Longari, Stefano, "Towards Secure Electronic Voting: a Survey on E-Voting Systems and Attacks.", *IEEE Access*, Vol 11, pp. 1-31, 2023. DOI: 10.1109/ACCESS.2025.3569334.

[4] Agate V., De Paola A., Ferraro P., Lo Re G., Morana M, "SecureBallot: A secure open source e-Voting system", Journal of Network and Computer Applications, 191, 2021. DOI: 10.1016/j.jnca.2021.103165

[5] Adida, Ben, "Helios: Web-based open-audit voting", Proceedings of the 17th USENIX Security Symposium, pp. 335 – 348, 2008.

[6] Russo A., Anta A.F., Vasco M.I.G., Romano S.P, "Chirotonia: A Scalable and Secure e-Voting Framework based on Blockchains and Linkable Ring Signatures", Proceedings - 2021 IEEE International Conference on Blockchain, Blockchain 2021, pp. 417-424. DOI: 10.1109/Blockchain53845.2021.00065

[7] Qu W., Wu L., Wang W., Liu Z., Wang H, "A electronic voting protocol based on blockchain and homomorphic signcryption", Concurrency and Computation: Practice and Experience, 34, 2022.

[8] Tsang, P.P., Wei, V.K., "Short Linkable Ring Signatures for E-Voting, E-Cash and Attestation", *Information Security Practice and Experience. ISPEC 2005. Lecture Notes in Computer Science*, vol 3439, 2005. https://doi.org/10.1007/978-3-540-31979-5_5

[9] José Luis Salazar, Joan Josep Piles, José Ruíz, José María Moreno-Jiménez, "E-cognocracy and its voting process", *Computer Standards & Interfaces*, Volume 30, Issue 3, Pages 124-131, ISSN 0920-5489, 2008. DOI: 10.1016/j.csi.2007.08.017.

[10] José Luis Salazar, Joan Josep Piles, José Ruiz-Mas, José María Moreno-Jiménez, "Security approaches in e-cognocracy", *Computer Standards & Interfaces*, Volume 32, Issues 5–6, Pages 256-265, ISSN 0920-5489, 2010. DOI: 10.1016/j.csi.2010.01.004.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Deployment, Analysis and Optimization Simulator for 5G Wireless Networks

Juan E. Galeote-Cazorla[(1)], Alejandro Ramírez-Arroyo[(2)], Jesús Galeano-Brajones[(3)], Pablo Padilla[(1)], Juan F. Valenzuela-Valdés[(1)]
juane@ugr.es, araar@es.aau.dk, jgaleanobra@unex.es, {pablopadilla, juanvalenzuela}@ugr.es
[(1)]Department of Signal Theory, Telematics and Communications, Research Centre for Information and Communication Technologies (CITIC-UGR), University of Granada, 18071, Granada, Spain.
[(2)]Department of Electronic Systems, Aalborg University (AAU), 9220 Aalborg, Denmark.
[(3)]Department of Computing and Telematics Engineering, Universidad de Extremadura, 06800, Mérida, Spain.

The advent of fifth-generation (5G) wireless network deployments presents significant challenges for service providers. The high losses and limitations of propagation mechanisms inherent to the millimeter wave (mmWave) band, combined with the complexity of real-world environments, necessitate simulation tools that account for all relevant factors. In this context, we have developed an innovative simulator designed to support the deployment, analysis and optimization of new 5G wireless networks. Following a non-standalone (NSA) deployment, the simulator can evaluate network performance by analyzing the behavior of mobile users within the simulated region. This paper summarizes the main functionalities and key features of the simulator. Additionally, a real-world deployment and mobility analysis use case is presented to demonstrate its capabilities.

***Palabras Clave*—5G systems, coverage, deployment, networks, simulation**

I. Introduction

Fifth-generation (5G) mobile wireless networks are already a reality. By the end of 2023, the number of subscribers had reached 1.4 billion and it is expected to exceed 5.3 billion by 2029 [1]. With their arrival, new transformative applications are emerging such as device-to-device (D2D) communication, machine-to-machine (M2M) communication, the Internet of things (IoT) and the Internet of vehicles (IoV) [2]. According to the International Telecommunication Union (ITU), the services offered by this generation are classified as: enhanced mobile broadband (eMBB), ultra-reliable and low-latency communication (uRLLC) and massive machine type communication (mMTC) [3], [4].

Some of the main characteristics of 5G systems include data rates in the order of gigabits per second and latencies below one millisecond. The key technology enabling these capabilities is the use of the millimeter wave (mmWave) band [2], which offers a wide range of frequencies with largely unallocated spectrum, thereby allowing higher transmission bandwidths to support a greater number of users [5]. However, mmWaves introduce new challenges in terms of propagation. In line-of-sight (LoS) scenarios, coverage areas are limited due to increased propagation losses. Additionally, in non line-of-sight (nLoS) scenarios, the main propagation mechanisms (reflection, diffraction and scattering) are significantly impaired, resulting in a notable reduction in received power [6]. Moreover, especially in outdoor environments, atmospheric absorption becomes a relevant factor in the mmWave band, leading to increased losses under conditions such as fog or rain [7].

The 3rd Generation Partnership Project (3GPP) defines two deployment options for 5G: non-stand alone (NSA) and stand alone (SA) architectures [8]. The NSA option leverages the existing long term evolution (LTE) infrastructure alongside the 5G radio access network (RAN) and its new radio (NR) interface. Alternatively, the SA option proposes the use of the NR interface connected to the new 5G core network (CN). Currently, most deployments are NSA; however, the trend is shifting toward SA [1]. The former focuses on eMBB and offers lower costs for service providers. The latter is clearly more expensive but allows providers to deliver the full range of services expected from 5G networks [9].

Deploying a new 5G network presents a significant challenge for service providers, who must meet minimum key performance indicators (KPIs) to ensure compliance with quality of service (QoS) requirements. One of the most limiting factors is the propagation channel in which the system operates [10], which is entirely dependent on the geographic region where the deployment is planned. The terrain topography, whether the area is urban or rural, and the atmospheric conditions all have a substantial impact on system performance. Due to the high complexity

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

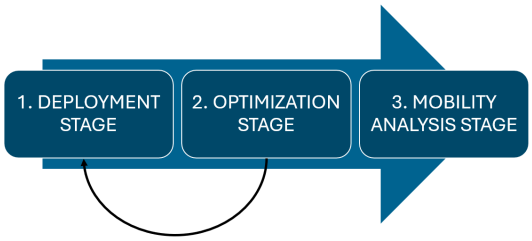


Fig. 1. Global workflow of the 5G simulation tool.

of accounting for the possible scenarios, providers must rely on simulation tools. Solutions such as NYUSIM [11] and QuaDRiGA [12] enable users to simulate a wide variety of propagation channels using different standardized models (e.g. 3GPP models [13]) depending on the general characteristics of the scenario. Alternatively, approaches such as Longley-Rice [14] or ray tracing [15] perform simulations tailored to specific environments, taking into account terrain topography or buildings geometry.

In this work we present an innovative simulator for studying 5G communication systems (see Fig. 1). It is capable of performing new 5G-NSA deployments in any real-world region specified by its geographical coordinates avoiding reliance on abstract models. To achieve this, a radio frequency (RF) coverage analysis is conducted using propagation models that incorporate terrain topography, buildings geometry and atmospheric effects. Optimization algorithms are available for various deployment variables, including nodes locations, operating frequencies and cell orientations in order to improve the overall coverage in the region. Once the deployment is completed, the simulator enables performance evaluation by simulating mobile users following realistic routes generated by a traffic simulator. Parameters such as received power, signal-to-interference-plus-noise ratio (SINR) and link capacity are computed for each user at every simulation time step.

The remainder of this paper is organized as follows. Section II presents the key capabilities of the 5G system simulation tool, with a focus on the deployment and mobility aspects. Section III introduces a specific use case involving deployment and performance analysis. Finally, Section IV summarizes the conclusions and outlines future work.

II. Simulator Key Capabilities

The developed simulator is primarily written in the MATLAB programming language. Additionally, it integrates various APIs to support specific functionalities. For example, it operates with real-world geographical scenarios —users can specify a simulation region using latitude and longitude boundaries. Several open data sources are consulted to extract the required geographical information: GMTED2010 for terrain elevation, OpenStreetMap for buildings geometry and metadata, and OpenCellID for actual LTE node locations (classified by service providers) [16], [17], [18].

A. Deployment and Optimization

The deployment stage begins by placing 5G macro cells at the actual LTE node locations. Each macro cell is divided into three sectors following 3GPP recommendations [13]. These sectors are served by 20 dBi gain directive antennas[1] with a half power beam width (HPBW) of 16º in the H-plane. Transmission power is specified by the user, along with the available operating frequencies. The latter are automatically assigned to the cells according to predefined rules aimed at reducing interference between the base stations (BSs). Next, an initial SINR coverage map is computed for all macro cells using either ray tracing or Longley-Rice propagation models [15], [14]. These models can be combined with supplementary models for rain, fog and atmospheric gases [19].

Once the macro cells are deployed, the next step involves placing micro cells. To this end, three different strategies are proposed: locating the new nodes (i) at the vertices of a hexagonal grid, (ii) at the hotspot areas[2] or (iii) in blind spots identified in the initial coverage map. In this case, the BSs are modeled as isotropic antennas without cell sectorization. As before, the user defines the transmission power and the available operating frequencies. A new coverage map is then computed, incorporating the contributions of both macro cells and micro cells. Finally, backhaul links are established by identifying the nearest macro cell to each micro cell.

An additional capability of the simulator is the optimization option to maximize the coverage. In order to do so, the optimized parameters are: (i) the angular orientation of the macro cells sectors, (ii) the assigned operation frequency of each BS and (iii) micro cells location variations with respect to their initial position. The objective function is the ratio of covered surface with a SINR greater than 0 dB. This functionality is based on genetic algorithms (GAs), specifically the so-called non-dominated sorting genetic algorithm II (NSGA-II) [20]. This is an elitist GA that starts by creating an auxiliary population Q_t from the initial population P_t through the employment of genetic operators, that is, selection, crossover, and mutation. Subsequently, the combined set $P_t \cup Q_t$ is ranked based on each solution's level of Pareto non-domination, organizing them in ascending order. The top-ranked solutions are then advanced to the next generation P_{t+1}. For solutions with identical ranks, NSGA-II employs the crowding distance as a density estimator to determine priority among solutions. This distance is calculated based on the average distance to the immediate neighbors of a solution to the left and right within the same front. By doing so, NSGA-II aims to give precedence to solutions located in less densely populated areas of the approximated Pareto front, promoting a more diverse and well-distributed set of solutions that cover the objective space more comprehensively. This method ensures that the selection process during the evolutionary cycle

[1] Implemented as 8×8 uniform planar arrays (UPAs).

[2] In this context, a hotspot refers to a location with a high density of users, such as a campus or a shopping center.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

favors those solutions that contribute to the exploration and preservation of the front's extent, facilitating a more thorough examination of potential solutions.

B. Users and Mobility

After the deployment stage is completed, the simulator can analyze network performance for users moving within the selected region. A time window is defined in the simulator inputs, along with a specific sample step. User routes are then generated using one of the following alternatives:

- *Random start-end points:* For a given number N of users (specified as input), N start and end random points are generated. After that, using the HERE Routing API [21], we generate real possible pedestrian routes between them.
- *Manual:* For individual users, routes can be manually specified or extracted from a measurement application. This enables comparisons with real-world performance measurements.
- *Mobility simulator:* The simulation tool integrates the TraCI API [22] to generate routes based on the SUMO traffic simulator [23], allowing separate analysis of pedestrians and vehicles following realistic mobility patterns.

To begin the analysis of mobile users, the simulator computes both the SINR and the received signal power from each deployed BS, using the specified propagation model. Terminals are modeled as isotropic antennas for computational simplicity. Next, the channel capacity is calculated for all possible pairings. Two alternatives are considered depending on the SINR value. For SINR values below 0 dB, capacity is computed assuming a single-input single-output (SISO) link [24]:

$$C_{\text{SISO}} = \text{BW} \log_2(1 + \text{SINR}), \tag{1}$$

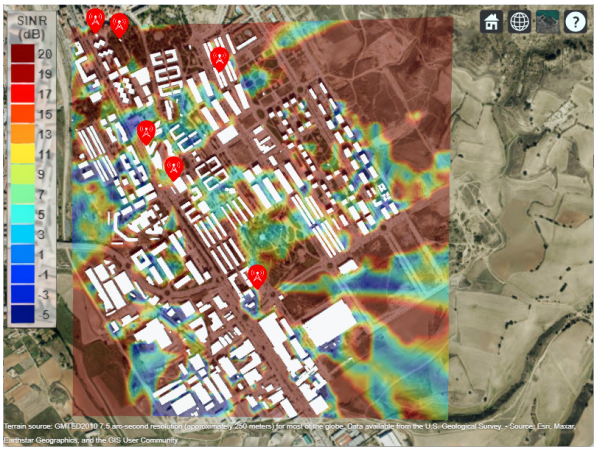
where BW is the bandwidth assigned to the user. In contrast, for SINR values greater than 0 dB, capacity is computed assuming a multiple-inputs multiple-outputs (MIMO) link [25]:

$$C_{\text{MIMO}} = \text{BW} \log_2 \left| \mathbf{I}_{N_{Rx}} + \frac{\text{SINR}}{N_{Tx}} \mathbf{H}\mathbf{H}^H \right|, \tag{2}$$

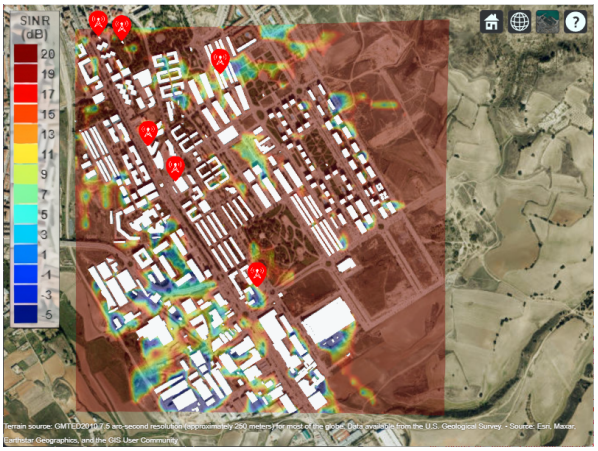
where N_{Rx} and N_{Tx} are the number of antennas in the receiver and transmitter arrays respectively, $\mathbf{I}_{N_{Rx}}$ is the identity matrix of order N_{Rx}, and $\mathbf{H}$ represents the MIMO channel matrix. In this case, the receiver antenna is modeled as a 2×4 UPA [26], resulting in an 8×64 MIMO system. The estimation of the $\mathbf{H}$ matrix is based on a ray tracing analysis, which provides the angle of departure (AoD) and angle of arrival (AoA) for each individual ray contribution [27]. Finally, each user is paired with the BS that offers the highest capacity at each simulation time step.

III. Use Case Example

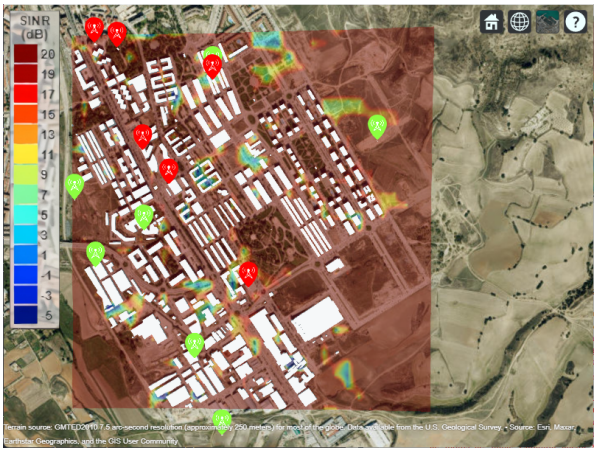
As a representative use case example, we performed a 5G deployment in the city of Cuenca (Spain), followed



(a) Initial deployment —no optimized (91.3% coverage).



(b) Deployment with optimized sectors angles (95.1% coverage).



(c) Optimized deployment with micro cell at blind spots (98.3% coverage).

Fig. 2. Stages of 5G network deployment in the selected urban area of Cuenca. Red markers indicate macro cell locations, while green markers represent micro cells.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

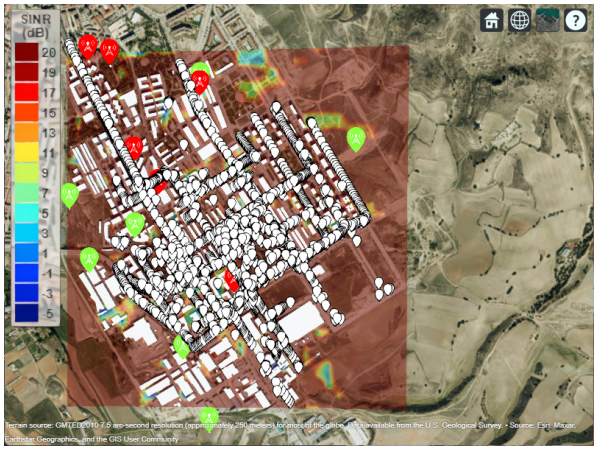


Fig. 3. Distribution of users (white markers) across the simulation region.

by a mobile users analysis. Specifically, an urban area of approximately 2.7 km^2 was selected. The southeast corner of the region corresponds to the geographical coordinates 40º 3.030' N and 2º 8.076' W, while the northwest corner is located at 40º 3.942' N and 2º 6.942' W.

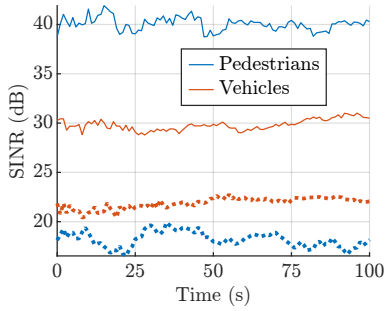
A. New Network Deployment

Firstly, we carried out the deployment simulation in the selected region using the Longley-Rice propagation model. Fig. 2 illustrates the sequence of coverage maps generated throughout the deployment process. In Fig. 2(a), six macro cells are shown deployed at Movistar's LTE node locations. The transmission power was set to 25 W, and four equidistant operating frequencies were available within the band from 3 GHz to 4 GHz. The angular orientations of the sectors were randomly assigned, resulting in 91.3% coverage (defined as the surface area with SINR greater than 0 dB).

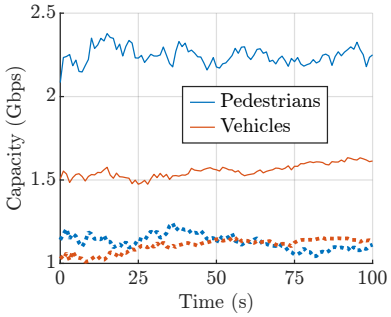
When the sector angles are optimized [see Fig. 2(b)], the coverage increases to 95.1%, due to reduced overlap between sectors. Subsequently, seven micro cells were deployed to cover the remaining blind spots. Their transmission power was set to 3 W and the operating frequencies were selected from 6 GHz and 7 GHz. Finally, after optimizing both the micro cells positions and frequency assignments, the coverage reached 98.3% [see Fig. 2(c)], representing a 7% improvement over the initial deployment.

B. Mobility Analysis

We conducted a brief mobility analysis for both the initial and the optimized deployments. A transmission bandwidth of 100 MHz was assigned to each user. The total simulation time was 100 s, with a sampling interval of 1 s. Using the SUMO simulator, we generated 1,543 users along with their respective routes throughout the region. Fig. 3 shows the distribution of users across the 5G network, where it can be observed that their positions align with the urban layout.



(a) Mean SINR per user.



(b) Mean capacity per user.

Fig. 4. SINR and capacity results from the 100 s mobility analysis with a time step of 1 s. Dashed lines correspond to the initial deployment while solid lines represent the optimized deployment.

Fig. 4 presents the results for SINR and capacity, separated between pedestrians and vehicles. Both parameters remain relatively stable over time. In the initial deployment (dashed lines), the average SINR values are 18.2 dB for pedestrians and 21.9 dB for vehicles, that is, 3.7 dB higher. However, the average capacity is 1.1 Gbps for both user types. In contrast, the optimized deployment (solid lines) yield average SINR values of 40.1 dB for pedestrians and 29.8 dB for vehicles, a difference of 10.3 dB. The corresponding average capacities are 2.2 Gbps and 1.6 Gbps respectively, both higher than in the initial deployment. Notably, capacity for pedestrians increased by 1.1 Gbps, indicating that the optimized deployment had a significant impact on pedestrian areas.

IV. Conclusions

In this work we have presented a novel simulator for analyzing and studying 5G wireless network deployments. The tool supports NSA deployments in any real-world geographical region by leveraging existing LTE nodes as macro cells and placing new micro cells according to various strategies. Using propagation models such as Longley-Rice or ray tracing, it can compute coverage maps that account for terrain topography, buildings geometry and atmospheric effects through supplementary models. Additionally, the simulator incorporates an optimization module to maximize the covered area. Once a deployment is completed, its performance can be evaluated by analyzing mobile user parameters such as received power, SINR and capacity. User mobility can be modeled using different

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

sources, including manual input, routing APIs, and traffic simulators that incorporate both pedestrian and vehicles movement following realistic routes.

The simulator is entirely developed in the MATLAB programming language, following a modular architecture that facilitates the integration of new functionalities. Some examples of future enhancements include: (i) support for SA deployments, (ii) customizable antenna models for receivers, (iii) alternative deployment strategies and (iv) parallel computing capabilities to reduce execution time.

ACKNOWLEDGEMENTS

This work has been supported by grant TED2021-129938B-I00 funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. It has also been supported by grants PID2020-112545RB-C54, PDC2022-133900-I00, PDC2023-145862-I00 and TED2021-131699B-I00, funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR; and in part by the predoctoral grant FPU22/03392.

REFERENCIAS

[1] P. Jonsson et al., "Ericsson Mobility Report," Ericsson, Kista, Sweden, Nov. 2023.

[2] M. Agiwal, A. Roy and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617-1655, 2016.

[3] *Minimum Requirements Related to Technical Performance for IMT–2020 Radio Interface(s),* Rep. ITU-R M.2410-0, International Telecommunication Union, Geneva, Switzerland, Nov. 2017.

[4] *Study on New Radio (NR) Access Technology Physical Layer Aspects,* 3GPP TR 38.802, 3rd Generation Partnership Project, Mar. 2017.

[5] T. S. Rappaport et al., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, pp. 335-349, 2013.

[6] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios and J. Zhang, "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks—With a Focus on Propagation Models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213-6230, 2017.

[7] L. Wei, R. Q. Hu, Y. Qian and G. Wu, "Key elements to enable millimeter wave communications for 5G wireless systems," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 136-143, 2014.

[8] *Summary of Rel-15 Work Items (Release 15),* 3GPP TR 21.915, 3rd Generation Partnership Project, Sep. 2019.

[9] M. Agiwal, H. Kwon, S. Park and H. Jin, "A Survey on 4G-5G Dual Connectivity: Road to 5G Implementation," *IEEE Access*, vol. 9, pp. 16193-16210, 2021.

[10] M. Shafi et al., "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201-1221, 2017.

[11] S. Sun, G. R. MacCartney and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5G wireless communications," in *2017 IEEE International Conference on Communications (ICC)*, Paris, France, pp. 1-7, 2017.

[12] S. Jaeckel, L. Raschkowski, K. Börner and L. Thiele, "QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 3242-3256, 2014.

[13] *Study on channel model for frequencies from 0.5 to 100 GHz (Release 17),* 3GPP TR 38.901, 3rd Generation Partnership Project, Dec. 2023.

[14] A. G. Longley and P. L. Rice, "Prediction of tropospheric radio transmission loss over irregular terrain," U.S. Department of Commeree, ESSA Rep. ERL-79-ITS-67, 1968.

[15] F. Fuschini, E. M. Vitucci, M. Barbiroli, G. Falciasecca and V. Degli-Esposti, "Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques," *Radio Science*, vol. 50, no. 6, pp. 469-485, 2015.

[16] *GMTED2010 Global Multi-resolution Terrain Elevation Data 2010,* U.S. Geological Survey, 2010.

[17] *OpenStreetMap,* OpenStreetMap Foundation. [Online]. Available: https://www.openstreetmap.org/

[18] *OpenCellid,* Unwired Labs. [Online]. Available: https://opencellid.org/

[19] *Attenuation by atmospheric gases and related effects*, Rec. ITU-R P.676-13, International Telecommunication Union, Aug. 2022.

[20] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[21] *HERE Routing API v8 - Devoloper Guide,* Here. [Online]. Available: https://www.here.com/docs/bundle/routing-api-developer-guide-v8/

[22] A. Acosta et al, "TraCI4Matlab: Enabling the Integration of the SUMO Road Traffic Simulator and Matlab® Through a Software Re-Engineering Process," *Modeling Mobility with Open Data*, Springer International Publishing, pp. 155–70, 2015.

[23] P. A. Lopez et al., "Microscopic Traffic Simulation using SUMO," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, pp. 2575-2582, 2018.

[24] K.-L. Du and M. N. S. Swamy, Wireless Communication Systems: From RF Subsystems to 4G Enabling Technologies. Cambridge: Cambridge University Press, 2010.

[25] S. Loyka and G. Levin, "On physically-based normalization of MIMO channel matrices," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1107-1112, 2009.

[26] M. Abdullah et al., "Future Smartphone: MIMO Antenna System for 5G Mobile Terminals," *IEEE Access*, vol. 9, pp. 91593-91603, 2021.

[27] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436-453, 2016.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Propuesta de mejora de Eclipse Zenoh sobre protocolos multi-flujo

Miguel Barón[†,§], Luis Diez[†], Raúl Torrego[§], Ramón Agüero[†]
[†]Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria, Santander, 39005, España
[§]Centro Tecnológico de Investigación Ikerlan, Arrasate/Mondragón, 20500, España
Correo electrónico: {ldiez, ramon}@tlmat.unican.es, {mbaron, rtorrego}@ikerlan.es

La importancia del protocolo Eclipse Zenoh en entornos IIoT está creciendo de manera constante, por su capacidad de unificar comunicación, almacenamiento y consultas de datos con baja latencia y una gran flexibilidad topológica. Por otro lado, el protocolo QUIC se ha consolidado como una alternativa al tradicional TCP, destacando por su reducción de latencias gracias al uso de *multistreaming* (multi-flujo). Este trabajo en curso propone introducir la capacidad de soportar múltiples flujos en Zenoh, con el fin de aprovechar las ventajas que ofrece QUIC como capa de transporte subyacente. Se plantea la motivación de esta propuesta, el desarrollo previsto y el diseño experimental, basado en diversas topologías y tecnologías de acceso, que permitirá comparar el rendimiento de Zenoh sobre QUIC frente a su implementación actual sobre TLS/TCP.

***Palabras Clave*—QUIC, Zenoh, multi-flujo, DRL, IIoT**

I. Introducción

La transformación digital de los entornos industriales ha consolidado el paradigma de la Internet Industrial de las Cosas (IIoT, *Industrial Internet of Things*), en el que miles de dispositivos, sensores y sistemas interconectados comparten datos en tiempo real para mejorar la trazabilidad, optimizar la toma de decisiones y aumentar la eficiencia operativa. Este nuevo escenario impone requisitos de comunicación mucho más estrictos que los de las redes tradicionales, como baja latencia, alta fiabilidad, tolerancia a fallos, seguridad, determinismo y escalabilidad en entornos distribuidos y heterogéneos.

En entornos IIoT combinados con arquitecturas de computación en la nube y en el borde de la red (*Cloud* y *Edge Computing*, respectivamente), los protocolos tradicionales, como el protocolo de transferencia de hipertexto (HTTP, *Hypertext Transfer Protocol*), diseñados para redes de propósito general, introducen una sobrecarga innecesaria y, además, no están optimizados para dispositivos embebidos. En consecuencia, se han propuesto protocolos más ligeros como *Message Queuing Telemetry Transport* (MQTT) [1] y *Constrained Application Protocol* (CoAP) [2], ya que presentan una mayor eficiencia en estos escenarios. Asimismo, las soluciones middleware, como *Data Distribution Service* (DDS) [3], también han ganado relevancia en arquitecturas distribuidas como *Robot Operating System* (ROS) 2, gracias a su modelo orientado a los datos y sus avanzadas capacidades de calidad de servicio (QoS, *Quality of Service*).

En este contexto, Eclipse Zenoh [4], desarrollado por ZettaScale, ha irrumpido recientemente como una alternativa interesante, gracias a que introduce un protocolo de comunicación eficiente, centrado en los datos, diseñado para unificar los modelos de publicación/suscripción, consulta y almacenamiento bajo una arquitectura ligera y distribuida. Como refleja la Figura 1, esta solución permite despliegues centralizados mediante un *router* Zenoh (*brokered*), distribuidos (*routed*) o *peer-to-peer* (P2P) (*clique* o *mesh*). En Zenoh, los nodos pueden funcionar como clientes, que requieren conectarse a un nodo intermedio (par o *router* Zenoh) para comunicarse; o como pares, que participan activamente en el encaminamiento, el descubrimiento y el almacenamiento distribuido. Además, Zenoh puede operar sobre diversas capas de la pila de protocolos (véase la Figura 2). Estas características lo convierten en una opción versátil para sistemas IIoT.

Zenoh está siendo adoptado en ámbitos como la robótica [6] y las comunicaciones *Vehicle-to-everything* (V2X) [7]. Diversos trabajos lo analizan y comparan con soluciones como MQTT, DDS, Kafka o HTTP [8], [9], [10], destacando su baja latencia, flexibilidad y mejor rendimiento en redes con recursos limitados.

Paralelamente, han surgido propuestas para incorporar protocolos de transporte novedosos, como QUIC [11], a las arquitecturas de comunicación IIoT. QUIC fue diseñado inicialmente como una alternativa a la combinación de *Transmission Control Protocol* (TCP), *Transport Layer Security* (TLS) y HTTP/2 [12] para la web, pero sus características lo hacen especialmente atractivo en entornos industriales exigentes.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

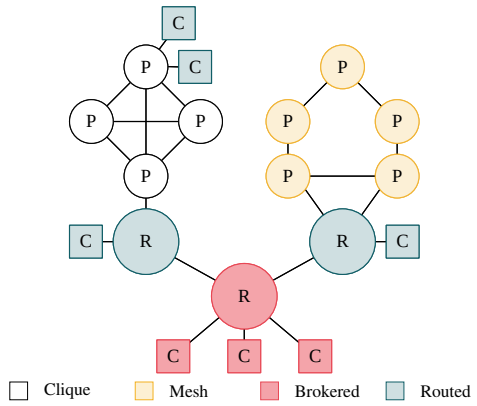


Fig. 1. Topologías con Zenoh. P: Par; C: Cliente; R: *Router* Zenoh [5].

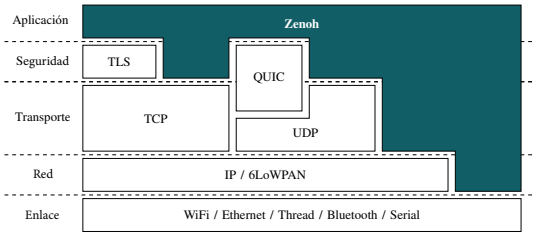


Fig. 2. Arquitectura de Zenoh [5].

Gracias a sus esquemas de seguridad intrínsecos (cifrado integrado), multiplexación de flujos sin *Head-of-Line Blocking* (HOLB), rápido establecimiento de conexión en cero tiempos de ida y vuelta (0-RTT) y mejor control de congestión, QUIC puede reducir significativamente la latencia y mejorar la fiabilidad de las comunicaciones. Recientemente ha habido varios trabajos que han explorado la integración de QUIC con protocolos como MQTT [13], CoAP [14] o *Advanced Message Queuing Protocol* (AMQP) [15], en los que se plantea como una alternativa prometedora a TCP o *User Datagram Protocol* (UDP) en comunicaciones industriales.

En un trabajo previo [5], se evaluó el rendimiento de Zenoh sobre diferentes topologías y condiciones de red, utilizando TCP y UDP. Se obtuvieron mejores resultados que con el protocolo MQTT bajo diversas configuraciones de QoS. A partir de esa experiencia, se propone en este trabajo estudiar la viabilidad y el impacto de utilizar QUIC como capa de transporte subyacente en Zenoh.

Aunque Zenoh ya contempla QUIC como protocolo de transporte, su implementación actual solo permite el uso de un único flujo bidireccional. Esta limitación impide aprovechar las capacidades de multiplexación de QUIC, lo que conlleva posibles efectos de HOLB y mayores retardos en la comunicación. El principal objetivo de esta línea de actuación es habilitar la posibilidad de utilizar múltiples flujos bidireccionales en la implementación en Rust de Zenoh[1] sobre QUIC. Esta mejora sentará las bases para futuras optimizaciones del *scheduling* de flujos mediante técnicas de aprendizaje máquina (ML, *Machine Learning*)

[1]https://github.com/eclipse-zenoh/zenoh

Tabla I
CLASES DE PRIORIDAD Y SUS VALORES ASOCIADOS.

Clase de Prioridad	Valor	Clase de Prioridad	Valor
Control	0	DataHigh	4
RealTime	1	Data[1]	5
InteractiveHigh	2	DataLow	6
InteractiveLow	3	Background	7

[1] Prioridad por defecto.

como el aprendizaje de refuerzo profundo (DRL, *Deep Reinforcement Learning*), con el propósito de reducir la latencia y aumentar la eficiencia de las comunicaciones en entornos IIoT.

II. IMPLEMENTACIÓN ACTUAL Y MEJORA PROPUESTA

Zenoh permite asignar diferentes prioridades a las colas internas que implementa (véase la Tabla I), lo que resulta especialmente útil para reducir la latencia en comunicaciones con requisitos temporales estrictos, como comandos de control o eventos en tiempo real. No obstante, esta priorización se aplica únicamente a nivel de aplicación. En escenarios con alta congestión o pérdidas frecuentes, no es posible garantizar una entrega preferente de los mensajes más importantes, ya que sus niveles de prioridad no son visibles para el protocolo de transporte subyacente.

La implementación de Zenoh presenta una arquitectura modular basada en abstracciones para gestionar sus comunicaciones. Como base se encuentra la librería (*crate* en Rust) `zenoh-transport`, que abstrae los detalles del transporte subyacente, gestionando sesiones, descubrimiento, fragmentación de mensajes, fiabilidad y colas de prioridad. Esta capa interactúa con enlaces concretos mediante interfaces genéricas definidas en `zenoh-link`, como `LocatorInspector` o `LinkManagerBuilderUnicast`. A su vez, estas interfaces se implementan como `zenoh-links`, proporcionando adaptadores para protocolos concretos como QUIC (`zenoh-link-quic`), TCP o UDP, apoyándose en utilidades compartidas (*traits*) del *crate* `zenoh-link-commons`. Esto permite que Zenoh seleccione y utilice automáticamente el transporte adecuado en función del *locator* especificado (por ejemplo, `quic/host:port`) o de la configuración, sin que el usuario deba gestionar los detalles de bajo nivel.

Actualmente, `zenoh-link-quic` no explota todas las posibilidades que ofrece QUIC, tal como lo implementa Quinn, el *crate* de QUIC desarrollado en Rust que utiliza Zenoh, especialmente en lo que respecta a la capacidad de gestionar múltiples flujos. En su estado actual, Zenoh se limita a usar un único flujo bidireccional por conexión, y rechaza los unidireccionales. Esto impone un cuello de botella e impide la gestión de flujos concurrentes con diferentes prioridades o características. Por ello, se propone una extensión de esta capa, implementando un decisor que permita instanciar múltiples flujos QUIC dinámicamente en función de eventos, como las expresiones de clave (KEs, *Key Expressions*), las prioridades definidas en la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

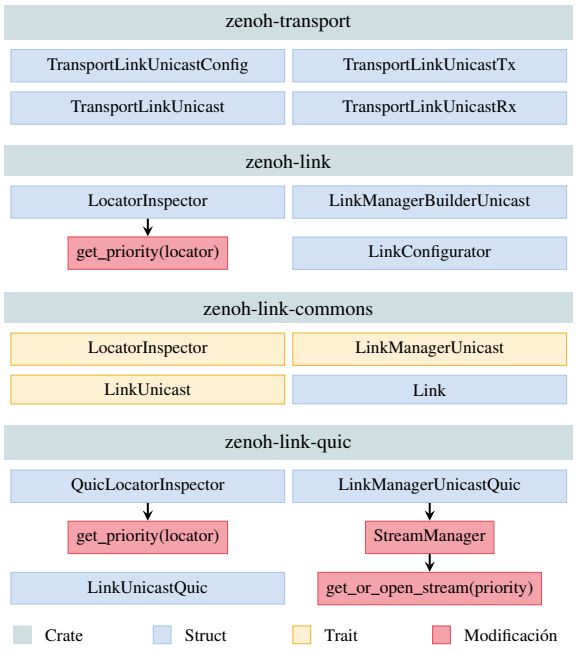


Fig. 3. Diagrama simplificado de la arquitectura modular de Zenoh para la gestión de la capa de enlace, junto con las modificaciones propuestas para soportar el multi-flujo con Zenoh sobre QUIC.

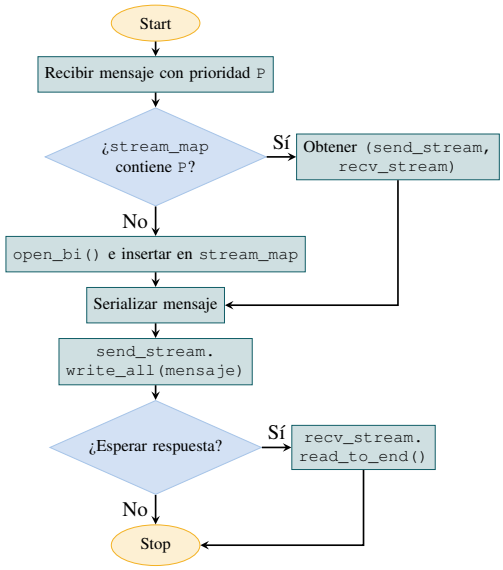


Fig. 4. Lógica del *StreamManager* propuesto para la creación dinámica de flujos QUIC en función de las colas de prioridades.

capa de aplicación o criterios de origen/destino, sin afectar a otras implementaciones de enlaces en Zenoh.

La Figura 3 muestra una visión simplificada de las principales estructuras implicadas en la arquitectura de Zenoh sobre QUIC, incluyendo las interfaces genéricas y sus implementaciones específicas. Inicialmente, se propone aprovechar las prioridades de la propia capa de aplicación (propagadas desde `zenoh-transport` al *locator* gestionado por `zenoh-link-quic`) para abrir dinámicamente flujos QUIC independientes, según el nivel de prioridad, permitiendo así una gestión más eficiente del tráfico. Para ello, será necesario implementar una función `get_priority()` en `QuicLocatorInspector` y exponerla a través de `LocatorInspector`, de modo que la información de prioridad pueda consultarse durante el establecimiento del enlace. Además, se diseñará un nuevo componente `StreamManager`, encargado de gestionar las colas de prioridad y crear los flujos. Se integrará en la estructura `LinkManagerUnicastQuic` y la Figura 4 ilustra la lógica de su funcionamiento.

Tras acometer las modificaciones correspondientes, será fundamental que el sistema sea configurable, permitiendo al usuario definir el número máximo de flujos simultáneos y el modo de asignación: ya sea un flujo por nivel de prioridad, múltiples flujos para la misma prioridad (para evitar congestión), o bien la multiplexación de distintos niveles de prioridad en un único flujo, para optimizar así el uso de recursos.

En fases futuras de este desarrollo, esta lógica podrá extenderse para considerar otros criterios: KEs, el destinatario u otros atributos relevantes. Además, dado que Quinn ya incorpora un *scheduler* basado en prioridades, combinado opcionalmente con algoritmos como *Round Robin* (RR) o *First In First Out* (FIFO) para flujos con la misma prioridad, sería muy interesante que pudiera reconocer y mapear las prioridades de Zenoh con las suyas propias, garantizando una gestión de tráfico coherente y optimizada a nivel de transporte. De hecho, se contempla como línea futura la mejora de este *scheduler* mediante un modelo de DRL, que permita adaptar dinámicamente la programación de flujos en función del estado de la red y del sistema, maximizando así el rendimiento en escenarios industriales exigentes.

La propuesta completa se ilustra en la Figura 5, donde se muestra la arquitectura extendida de Zenoh sobre QUIC. El módulo *Decider*, en la capa de enlace de Zenoh y cuya implementación se describe en este artículo, selecciona dinámicamente el flujo QUIC adecuado, en función de los parámetros correspondientes. Posteriormente, los flujos se gestionan por un *Scheduler* en Quinn, cuya política de planificación se optimiza mediante un agente DRL entrenado con retroalimentación del entorno de red.

III. Metodología y Evaluación Experimental

En entornos industriales (IIoT) con una alta densidad de dispositivos, enlaces inalámbricos propensos a pérdidas y aplicaciones con estrictos requisitos de latencia y fiabilidad, es fundamental realizar un análisis previo del comportamiento de los protocolos de comunicación bajo condiciones realistas.

La línea de trabajo que se introduce en este artículo se centrará en estudiar el impacto que tiene la gestión multi-flujo en el rendimiento global del sistema, evaluando la estrategia de mapeo de colas de prioridad sobre QUIC implementada en el decisor de Zenoh. Para ello, se comparará con su implementación estándar sobre TCP con TLS, en escenarios representativos del ámbito industrial.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

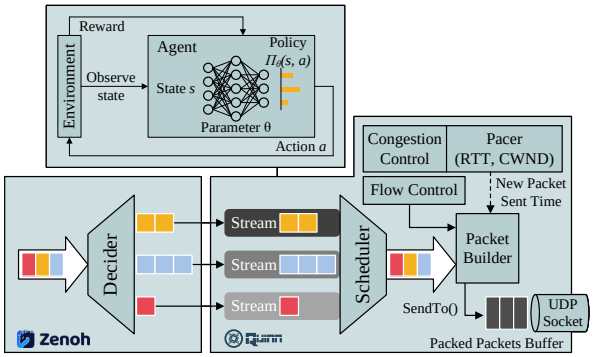


Fig. 5. Arquitectura propuesta para la integración de Zenoh sobre QUIC con planificación inteligente basada en DRL. Basado en [16].

Se usarán dispositivos Raspberry Pi 4 conectados mediante redes Wi-Fi y 5G, esta última proporcionada por una estación base Amarisoft. Además, se reproducirán distintas topologías de red, como la comunicación P2P, arquitecturas con *routers* Zenoh intermedios (*brokered*), y configuraciones más complejas con múltiples *routers* en cadena (*routed*) (véase la Figura 1).

El análisis de rendimiento incluirá métricas como la latencia de transmisión o el RTT, la variación temporal entre entregas consecutivas (*jitter*), la tasa efectiva de transmisión (*throughput*), el ratio de finalización (*completion ratio*) entre QUIC y TCP o la tasa de pérdida de mensajes. Estas métricas se medirán tanto sobre canales ideales como en condiciones más realistas, con pérdidas generadas por la propagación, congestión o variabilidad en la conectividad.

IV. Conclusiones y trabajo futuro

En este trabajo se ha presentado una propuesta para extender la implementación de Zenoh sobre QUIC, habilitando la apertura dinámica de múltiples flujos en función de la prioridad de los mensajes. Esto no solo permitirá reducir la latencia y mejorar la eficiencia bajo diferentes condiciones de carga, sino que también facilitará integrar esquemas avanzados de planificación adaptativa, basados en DRL, en la capa de transporte.

Los próximos pasos se centrarán en la implementación y evaluación de esta propuesta en entornos realistas, así como en la integración de técnicas de optimización que aprovechen al máximo las capacidades de Zenoh y QUIC como protocolos de comunicación en sistemas industriales.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el proyecto XWAVE del programa de investigación ELKARTEK del Gobierno Vasco (Ref. KK-2025/00056), por el Ministerio de Asuntos Económicos y Transformación Digital de España mediante el proyecto SITED (PID2021-125725OB-I00), y por el Gobierno de Cantabria, Fondo Europeo de Desarrollo Regional (FEDER), a través del programa "Ayudas a proyectos de investigación con alto potencial industrial de agentes tecnológicos de excelencia para la competitividad industrial (TCNIC)" (2023/TCN/002).

Referencias

[1] A. Banks, E. Briggs, K. Borgendale, and R. Gupta, "MQTT Version 5.0," OASIS Standard, Tech. Rep., Mar. 2019. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html

[2] Z. Shelby, K. Hartke, and C. Bormann, "The Constrained Application Protocol (CoAP)," RFC 7252, Jun. 2014. [Online]. Available: https://www.rfc-editor.org/info/rfc7252

[3] Object Management Group (OMG), "Data Distribution Service (DDS)," https://www.omg.org/spec/DDS/, 2015, version 1.4.

[4] A. Corsaro, L. Cominardi, O. Hecart, G. Baldoni, J. Avital, J. Loudet, C. Guimares, M. Ilyin, and D. Bannov, "Zenoh: Unifying Communication, Storage and Computation from the Cloud to the Microcontroller," in *2023 26th Euromicro Conference on Digital System Design (DSD)*. Los Alamitos, CA, USA: IEEE Computer Society, Sep. 2023, pp. 422–428.

[5] M. Barón, L. Diez, M. Zverev, J. R. Juárez, and R. Agüero, "On the performance of Zenoh in Industrial IoT Scenarios," *Ad Hoc Networks*, p. 103784, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870525000320

[6] D. Cacciabue, J. Marino, F. Aglieco, M. Levorato, D. Perroni, and F. Risso, "Benchmarking Different Strategies for Offloading ROS2 Computation to the Edge," in *2024 IEEE 10th International Conference on Network Softwarization (NetSoft)*, 2024, pp. 49–54.

[7] L. Lusvarghi, B. Coll-Perales, J. Gozalvez, K. Aghababaiyan, M. Almela, and M. Sepulcre, "Characterization of In-Vehicle Network Sensor Data Traffic in Autonomous Vehicles," in *2024 IEEE Vehicular Networking Conference (VNC)*, 2024, pp. 211–214.

[8] J. J. López Escobar, R. P. Díaz-Redondo, and F. Gil-Castiñeira, "Unleashing the power of decentralized serverless IoT dataflow architecture for the Cloud-to-Edge Continuum: a performance comparison," *Annals of Telecommunications*, vol. 79, no. 3, pp. 135–148, Apr. 2024. [Online]. Available: https://doi.org/10.1007/s12243-023-01009-x

[9] J. Zhang, X. Yu, S. Ha, J. Peña Queralta, and T. Westerlund, "Comparison of Middlewares in Edge-to-Edge and Edge-to-Cloud Communication for Distributed ROS 2 Systems," *Journal of Intelligent & Robotic Systems*, vol. 110, no. 4, p. 162, 2024. [Online]. Available: https://doi.org/10.1007/s10846-024-02187-z

[10] G. Baldoni, J. Quevedo, C. Guimarães, A. de la Oliva, and A. Corsaro, "Data-centric service-based architecture for edge-native 6g network," *IEEE Communications Magazine*, vol. 62, no. 4, pp. 32–38, 2024.

[11] J. Iyengar and M. Thomson, "QUIC: A UDP-Based Multiplexed and Secure Transport," RFC 9000, May 2021. [Online]. Available: https://www.rfc-editor.org/info/rfc9000

[12] A. Langley, A. Riddoch, A. Wilk, A. Vicente, C. Krasic, D. Zhang, F. Yang, F. Kouranov, I. Swett, J. Iyengar, J. Bailey, J. Dorfman, J. Roskind, J. Kulik, P. Westin, R. Tenneti, R. Shade, R. Hamilton, V. Vasiliev, W.-T. Chang, and Z. Shi, "The QUIC Transport Protocol: Design and Internet-Scale Deployment," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, ser. SIGCOMM '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 183–196. [Online]. Available: https://doi.org/10.1145/3098822.3098842

[13] F. Fernández, M. Zverev, P. Garrido, J. R. Juárez, J. Bilbao, and R. Agüero, "Even Lower Latency in IIoT: Evaluation of QUIC in Industrial IoT Scenarios," *Sensors*, vol. 21, no. 17, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/17/5737

[14] J.-H. Jung, H.-B. Nam, D.-K. Choi, and S.-J. Koh, "Use of QUIC for CoAP transport in IoT networks," *Internet of Things*, vol. 24, p. 100905, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660523002287

[15] F. Iqbal, M. Gohar, H. Karamti, W. Karamti, S.-J. Koh, and J.-G. Choi, "Use of QUIC for AMQP in IoT networks," *Computer Networks*, vol. 225, p. 109640, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128623000853

[16] F. Fernández, F. Khan, M. Zverev, L. Diez, J. R. Juárez, A. Brunstrom, and R. Agüero, "Exploiting stream scheduling in quic: Performance assessment over wireless connectivity scenarios," *Ad Hoc Networks*, vol. 164, p. 103599, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870524002105

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Aprendizaje *learning by doing* y evaluación sin estrés para la enseñanza de protocolos de enrutamiento

Antonio Ruiz-Martínez, Sara N. Matheu-García, Jorge Gallego-Madrid, Noelia Pérez-Palma,
Juan Antonio Martínez-Navarro, Ana Hermosilla-García, Ramon Sanchez-Iborra
Departamento de Ingeniería de la Información y las Comunicaciones,
Universidad de Murcia
Facultad de Informática. Campus de Espinardo. 30100. Murcia.
arm@um.es, saranieves.matheu@um.es, jorgegm@um.es, noelia.perez3@um.es, juanantonio@um.es,
ana.hermosilla@um.es, ramonsanchez@um.es

Resumen

La enseñanza de protocolos de enrutamiento en Ingeniería Informática sigue presentando dificultades para conectar la teoría con la práctica. Esta dificultad suele abordarse mediante la resolución de problemas teóricos o por medio de pequeñas prácticas de laboratorio con una entrevista final, lo que genera cierto estrés al alumnado. Para abordar esto, presentamos el enfoque seguido en la asignatura de Arquitectura de Redes del Grado en Ingeniería Informática mediante una metodología de aprendizaje *learning by doing* apoyada en un proyecto integrador guiado inspirado en un escenario real. Esta metodología combina explicaciones breves con trabajo autónomo grupal y se realiza un seguimiento semanal de cada grupo. Aquellos que asisten y participan activamente durante las sesiones de prácticas y redactan un buen informe final no necesitan hacer la entrevista. Nuestros resultados muestran que los estudiantes prefieren un seguimiento semanal, y sus buenos resultados académicos confirman que la metodología propuesta aumenta su motivación y satisfacción.

***Palabras Clave*—*learning by doing*, innovación docente, evaluación, redes, metodologías activas, enrutamiento**

I. INTRODUCCIÓN

Las asignaturas que abordan los conceptos básicos de redes de ordenadores están ampliamente adoptadas en los títulos relacionados con las Tecnologías de la Información y las Comunicaciones (TIC), como todos aquellos relacionados con la Ingeniería Informática, la Ciencia e Ingeniería de Datos y, por supuesto, las Telecomunicaciones. En estas asignaturas uno de los principales retos es cómo enseñar de forma práctica los conceptos teóricos, especialmente cuando es necesario abordar grandes escenarios como es el caso de las redes de área amplia (WAN) donde pueden estar conviviendo varios protocolos de enrutamiento.

Para afrontar este reto y disponer de una infraestructura donde enseñar estos conceptos de forma práctica, se ha hecho habitualmente uso de simuladores por varios motivos, entre los que destacan [1], [2]: que permiten experimentar con escenarios y configuraciones complejas a un reducido costo, los estudiantes pueden practicar de forma segura y en cualquier lugar sin temor a causar daño en una red real y, en general, son fáciles de usar.

Sin embargo, en trabajos previos [2], [3], [4], [5], con el objetivo de simplificar el despliegue o presentar escenarios sencillos, el uso de simuladores en el diseño de prácticas suele limitarse a ejercicios aislados centrados en aspectos específicos. Esta aproximación impide que el alumnado adquiera una visión completa de la complejidad inherente al despliegue y configuración de redes formadas por múltiples sistemas autónomos. Como consecuencia, se dificulta abordar la tarea de forma holística, ya que los distintos elementos que componen una topología realista y sus etapas se presentan de forma fragmentada y sin una conexión clara. Asimismo, la evaluación de esta parte práctica suele estar basada en la entrega semanal de boletines de laboratorio y una prueba final, ya sea oral o escrita, lo cual supone una alta carga de trabajo a lo largo del desarrollo de la asignatura, además de un foco de generación de estrés para el alumnado, respectivamente.

En este trabajo, se presenta la metodología docente de prácticas desarrollada en la asignatura Arquitectura de Redes del Grado en Ingeniería Informática de la Universidad de Murcia. La metodología está basada en el despliegue y puesta en marcha de un escenario realista que se realiza progresivamente siguiendo el enfoque de *learning by doing* [6], [7], [3]. Para el despliegue se

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

hace uso de un sencillo, pero realista, simulador de redes para hacer el despliegue y configuración de una topología compleja de red compuesta por varios dominios. Este proceso, junto con cuestiones relacionadas que permiten que el alumnado razone y comprenda la teoría asociada al desarrollo práctico, será explicado en una memoria final que entregarán al final de las sesiones. Todo este proceso se sigue de forma minuciosa por el equipo docente para realizar un seguimiento semanal del progreso de los estudiantes. En base a dicho seguimiento, los estudiantes que hayan demostrado una evolución positiva y adquisición de competencias adecuada a lo largo del curso, evitan la realización de una entrevista final, que sí deberán realizar aquellos que no han seguido o no se les ha podido realizar dicho seguimiento. El objetivo de esta metodología es múltiple: descargar al alumnado de las entregas continuas de boletines de prácticas, fomentar el trabajo continuo de los estudiantes en el laboratorio y eliminar el foco de estrés que puede suponer enfrentarse a una entrevista final. Hasta donde se ha podido analizar, no se han encontrado trabajos similares en la literatura que estén orientados a la realización de un proyecto de redes para el aprendizaje de conceptos de enrutamiento y los trabajos encontrados están limitados a la realización de prácticas con escenarios acotados o dedicados a la evaluación del uso de alguna de las herramientas de simulación.

El resto del documento se organiza de la siguiente forma. La Sección II discute los trabajos previos relacionados. El contexto de la asignatura foco de este estudio se presenta en la Sección III. La metodología de prácticas propuesta se detalla en la Sección IV. Los resultados de la evaluación de dicha metodología se presentan en la Sección V. La Sección VI discute las lecciones aprendidas. Finalmente, las conclusiones de este trabajo se plasman en la Sección VII.

II. TRABAJOS RELACIONADOS

La enseñanza de redes y, en particular, de los protocolos de enrutamiento, requiere que los estudiantes adquieran tanto conocimientos teóricos como habilidades prácticas para la configuración y gestión de redes reales. Tradicionalmente, estas competencias se desarrollan en laboratorios prácticos, donde los estudiantes, de manera individual o en grupo, resuelven problemas específicos de configuración y diagnóstico de redes [8].

Sin embargo, la implementación de laboratorios físicos con topologías complejas supone un alto coste en términos de recursos y tiempo, lo que ha llevado a la popularización de herramientas de virtualización y simulación de redes [1], [9], [10]. Estas herramientas permiten recrear infraestructuras de red de forma realista y accesible, facilitando la centralización de la gestión y la recuperación ante fallos. Estudios recientes [11] indican que el uso de simuladores es tan efectivo como la manipulación de *hardware* real, con la ventaja adicional de ofrecer un entorno controlado para la experimentación sin riesgos.

Dentro de las experiencias que se han realizado con estas herramientas, destacan aquellas realizadas con VNUML-UM [12], [13], GNS3 [14], Omnet++ [15] o Cisco Packet Tracer (CPT), ampliamente utilizadas en entornos académicos para la enseñanza de redes [2]. Actualmente, CPT[1] se ha consolidado como una de las herramientas más utilizadas en la enseñanza de redes debido a su versatilidad, facilidad de uso y la percepción positiva de los estudiantes acerca del simulador para aprender conceptos de redes [16], [17], [18]. De hecho, una reciente revisión sistemática de la literatura la confirma como la herramienta más utilizada para la enseñanza de redes [19]. Su interfaz intuitiva y sus amplias funcionalidades permiten a los estudiantes adquirir confianza en la configuración de redes sin la necesidad de equipamiento físico. En la Facultad de Informática de la Universidad de Murcia, se ha empleado la virtualización como recurso docente durante muchos años [13], [14], [12], [20], mejorando la formación práctica del alumnado. Actualmente, CPT es la herramienta empleada en los laboratorios prácticos de la asignatura foco de este estudio debido a las ventajas que acabamos de comentar.

Los programas de prácticas suelen estar restringidos a sesiones de laboratorio aisladas, donde cada concepto se aborda de manera independiente, sin una conexión clara con el resto de contenidos tratados en la asignatura en cuestión o son tratados con escenarios simplificados que no representan entornos reales [21], [2], [22]. Esta segmentación puede favorecer la memorización de comandos en lugar del desarrollo de una comprensión global de la red, lo que limita la capacidad de los estudiantes para enfrentarse a escenarios reales. De hecho, en el estudio sobre metodologías de enseñanza en redes realizado en [23], se destaca la importancia de estructurar las prácticas dentro de una secuencia didáctica coherente, con un hilo conductor que favorezca la integración progresiva de los conocimientos. Además, metodologías activas como el *learning by doing* [3] o el aprendizaje basado en proyectos permiten centrar el aprendizaje en el estudiante, en contextos reales, aumentando su motivación, interés y compromiso, y mejorando los resultados académicos [24], especialmente en aspectos prácticos [25].

Respecto a la evaluación de las prácticas, los exámenes orales o entrevistas han sido una de las técnicas más utilizadas para evaluar y controlar el rendimiento académico de los estudiantes según [26]. No obstante, en los últimos años, su uso ha disminuido debido a la creciente masificación en las universidades. Las entrevistas requieren un tiempo considerable para aplicarse de manera efectiva y con la profundidad necesaria, lo que resulta poco viable cuando el número de estudiantes es elevado. Su mayor ventaja es que permiten evaluar con mayor precisión el nivel de conocimientos del estudiante, al mismo tiempo que fomentan habilidades como la comunicación, la expresión oral y la capacidad de

[1] https://learningnetwork.cisco.com/s/article/el-software-de-simulacion-cisco-packet-trace

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

improvisación. Sin embargo, a pesar de sus beneficios, los exámenes y entrevistas orales pueden generar altos niveles de estrés en los estudiantes, afectando su rendimiento académico [27]. Otra posible estrategia de evaluación es llevarla a cabo por medio de evaluación continua. Así, al alumnado se le pide que realice varias entregas evaluables a lo largo del período. Este enfoque permite realizar un seguimiento constante del progreso del alumnado, abordar potenciales deficiencias formativas [28], [29] así como realizar cambios o ajustes en el desarrollo de la asignatura. Sin embargo, este tipo de evaluación, si no está adecuadamente implementada puede suponer una sobrecarga de trabajo para el estudiante y derivar en un incremento del estrés asociado al esfuerzo cognitivo continuado [30] o incluso limitar la efectividad del feedback como herramienta de mejora [31].

Para mitigar estos inconvenientes y promover una experiencia formativa más efectiva, nuestra propuesta se enmarca en un enfoque *learning by doing*, que permite al estudiante aprender a través de la acción, integrando progresivamente los conocimientos en contextos prácticos y realistas. Además, se realiza un seguimiento continuo de los estudiantes, registrando sus avances y resolviendo sus dudas, lo que permite prescindir de las entrevistas orales para aquellos que participan activamente y progresan adecuadamente en los laboratorios.

III. CONTEXTO: ARQUITECTURA DE REDES

El Grado en Ingeniería Informática de la UMU presenta dos asignaturas obligatorias (en los cursos 2º y 3º) que abordan aspectos fundamentales de las redes de ordenadores. En la asignatura Redes de Comunicaciones de 2º curso, se presenta una visión general de los servicios ofrecidos por las distintas capas de la arquitectura TCP/IP, sin ahondar en la funcionalidad específica de cada nivel, y considerando redes tipo LAN. Por su parte, la asignatura objeto de este trabajo, Arquitectura de Redes, impartida en el primer cuatrimestre del tercer curso, parte de los conocimientos y competencias adquiridos en la asignatura anterior, para profundizar en aspectos clave para la comprensión del funcionamiento de las redes WAN e Internet: enrutamiento intra e interdominio, comunicaciones multicast, aspectos avanzados de TCP y, finalmente, calidad de servicio e ingeniería de tráfico. Así, los contenidos teóricos de la asignatura se dividen en seis temas, que son los siguientes:

1) Introducción a la Arquitectura de Redes.
2) Algoritmos y protocolos de enrutamiento.
3) IP Multicast.
4) Aspectos Avanzados del Protocolo TCP.
5) Control de Tráfico y Control de Congestión.
6) Calidad de Servicio e Ingeniería de Tráfico.

Por su parte, el programa de prácticas se centra fundamentalmente en la resolución de un proyecto final que se articula en torno a una secuencia de laboratorios progresivos orientados a la configuración práctica de topologías de red multidominio. La Fig. 1 presenta el escenario de trabajo propuesto durante el curso 2024-25. Como se puede observar, está compuesto por dos organizaciones (sistemas autónomos), cada una haciendo uso de un protocolo en enrutamiento distinto (RIP y OSPF), y con rangos de direcciones IP previamente asignados que deberán ser segmentados mediante *subnetting*. Sobre esta topología, los estudiantes deben configurar cada dominio de forma independiente, para posteriormente interconectarlos y garantizar la conectividad entre todos los nodos de la red. El programa de prácticas es el siguiente:

1) Introducción a Cisco Packet Tracer.
2) Introducción a la Administración de Routers Cisco.
3) Direccionamiento IPv4.
4) Enrutamiento con RIP.
5) Enrutamiento con OSPF.
6) Redistribución de rutas en RIP y OSPF.

Los laboratorios propuestos permiten a los estudiantes alcanzar los siguientes resultados de aprendizaje:

1) Utilizar los conceptos fundamentales referentes al enrutamiento y funcionamiento de la pila de protocolos.
2) Adquirir una visión general de los diferentes factores que afectan el diseño de los protocolos y las redes de comunicaciones, que le permitan analizar y evaluar las diferentes alternativas de diseño para cada tipo de red.
3) Diseñar y poner en marcha redes de comunicaciones, analizando los requisitos específicos, y seleccionando las soluciones más interesantes en cada caso.
4) Configurar y administrar equipamiento de red para que realice las funciones necesarias según el diseño de red.
5) Identificar y resolver problemas y fallos en las redes de comunicaciones, y desviaciones del comportamiento esperado de los protocolos de comunicaciones.

La asignatura de Arquitectura de Redes proporciona los conocimientos y destrezas que sientan las bases para continuar la formación en Ingeniería Telemática dentro del Grado en Ingeniería Informática de la UMU. La siguiente asignatura en esta línea formativa es Servicios Telemáticos, impartida en el segundo cuatrimestre del tercer curso, centrada en el nivel de aplicación de la arquitectura TCP/IP e introduciendo conceptos fundamentales de ciberseguridad.

Estas tres asignaturas (Redes de Comunicaciones, Arquitectura de Redes y Servicios Telemáticos) conforman el núcleo formativo en Ingeniería Telemática para todos los estudiantes del Grado. En el 4º curso existe una mención en Tecnologías de la Información que se puede seguir mediante la elección de una serie de asignaturas optativas que abordan aspectos más específicos y profundos a los tratados en las asignaturas mencionadas anteriormente.

IV. METODOLOGÍA DE PRÁCTICAS

La metodología de prácticas de la asignatura se basa fundamentalmente en el trabajo por parejas durante

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

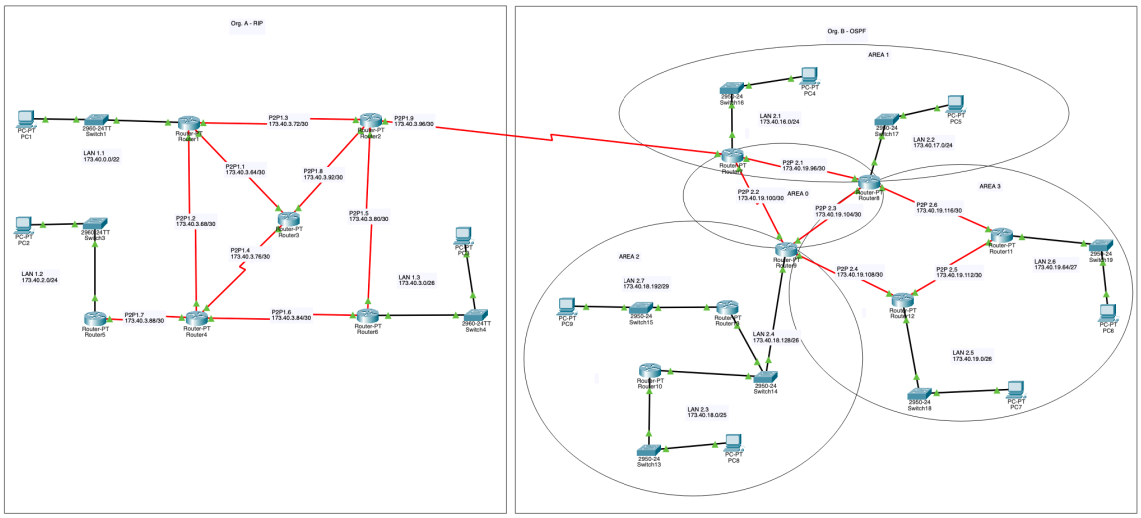


Fig. 1. Escenario final de prácticas a resolver en el curso 2024-25.

sesiones de trabajo tanto dirigido como autónomo. Como se ha mencionado en la sección anterior, el programa de prácticas se divide en seis partes. Durante el desarrollo del cuatrimestre, el profesor introduce cada una de ellas de forma breve, recordando los conceptos clave que ya han sido debidamente explicados en las sesiones de teoría, y dando las explicaciones técnicas necesarias para que los estudiantes puedan trabajar estos conceptos en escenarios realistas (usando el equipamiento de red que simula CPT), de forma que pongan en práctica lo que van aprendiendo de acuerdo al enfoque *learning by doing*. Para cada una de las partes, se dejan entre una y dos sesiones de trabajo autónomo para que los estudiantes asienten los conocimientos adquiridos y los apliquen a la resolución del escenario final. De esta forma, tras la introducción de cada una de las seis partes, los estudiantes tienen tiempo suficiente para asimilar los conceptos y ganar experiencia práctica real. La planificación semanal de las sesiones de prácticas es la siguiente:

- Semana 1: Trabajo autónomo: Instalación de Cisco Packet Tracer.
- Semana 2: Introducción a Cisco Packet Tracer.
- Semana 3: Introducción a la Administración de Routers Cisco.
- Semana 4: Direccionamiento IPv4.
- Semana 5: Enrutamiento con RIP.
- Semana 6: Trabajo autónomo.
- Semana 7: Enrutamiento con OSPF (1a parte).
- Semana 8: Enrutamiento con OSPF (2a parte).
- Semana 9: Trabajo autónomo.
- Semana 10: Redistribución de rutas en RIP y OSPF.
- Semana 11: Trabajo autónomo.
- Semana 12: Trabajo autónomo.
- Semana 13: Trabajo autónomo.
- Semana 14: Entrevista.

Cada una de las prácticas cuenta con un guión explicativo, tanto de los conceptos teóricos como de la parte práctica, junto con un escenario realista de prueba en el que aplicar lo aprendido. De esta forma, se sigue un enfoque de aprendizaje basado en trabajo autónomo con diversos escenarios. En cada uno de éstos se simplifica el escenario para que se centre en los contenidos específicos que se pretenden enseñar en cada sesión. Por ejemplo, en la cuarta parte, que corresponde al enrutamiento con RIP, se proporciona a los estudiantes un guión con un resumen teórico del funcionamiento del protocolo, una recopilación de los comandos necesarios para gestionar RIP desde CPT y una topología de red simplificada en la que aplicar directamente el protocolo. Así, en cada uno de estos escenarios, el estudiante puede focalizarse en los aspectos más importantes a trabajar en cada sesión. De este modo, pueden practicar para, posteriormente, aplicarlo al escenario final.

Durante el curso 2024/24 se ha hecho uso de Cisco Packet Tracer (CPT) v8.2.2. Como se ha comentado anteriormente, CPT permite el diseño y despliegue de escenarios de red complejos de forma sencilla. Su interfaz está diseñada para que la interacción con los elementos de red sea lo más realista posible, de manera que para su configuración se tengan que seguir los mismos procedimientos que se llevarían a cabo si se estuviera haciendo un despliegue físico real. Gracias a este *software* hay una transferencia muy alta de las aptitudes adquiridas en la asignatura cuando los estudiantes se enfrenten a problemas similares en el mundo laboral.

Como objetivo final de las sesiones de prácticas, como se ha mencionado en la Sección III, los estudiantes tienen que resolver un escenario final en parejas. El escenario se presenta en las primeras sesiones para que los estudiantes tengan una visión completa del problema que deben resolver desde el principio. Así, conforme progresan los contenidos y se van presentando los diferentes guiones, los estudiantes pueden ir aplicando estos conocimientos de manera incremental sobre el escenario final. Gracias a las sesiones de trabajo autónomo que hay en cada parte, los estudiantes tienen tiempo suficiente para asimilar los

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

conceptos clave, practicarlos con las topologías de red simplificadas y, posteriormente, aplicarlos en la resolución del escenario final. Además de configurar correctamente la topología final propuesta según las indicaciones dadas, los estudiantes tienen que elaborar una memoria final en la que expresen de forma analítica y detallada el proceso de configuración de todo el escenario y responder a una serie de cuestiones sobre el funcionamiento de la topología y los protocolos empleados, motivando sus argumentos. Las cuestiones les permiten, por medio del despliegue, profundizar en aspectos teóricos de los protocolos como, por ejemplo, los mecanismos de convergencia en RIP o los tipos de áreas en OSPF.

Con respecto a la evaluación de las prácticas, se trata de fomentar el sistema de evaluación continua. Para ello, los profesores de la asignatura hacen un control semanal del avance por parte de los estudiantes agrupados en parejas, anotando sus progresos y el dominio observado de la materia. Al final del cuatrimestre, se analiza si cada pareja ha asistido a la mayoría de las sesiones, si los estudiantes han demostrado haber adquirido las competencias y destrezas esperadas y si han sabido aplicarlas correctamente en el escenario final. En función de este seguimiento continuo, sólo los estudiantes que no han asistido, o bien no han demostrado en las sesiones un dominio suficiente de la materia, son convocados a una entrevista final sobre la resolución del escenario. En ella, para evaluar si han alcanzado los objetivos esperados, se plantean diversas cuestiones relacionadas tanto con el funcionamiento de los protocolos desplegados como con la configuración del escenario realizada por los estudiantes. Esta entrevista tiene un determinado peso en la calificación de prácticas (50%) junto con la valoración que reciba la memoria de prácticas que entreguen. Para aquellos estudiantes que hayan superado el seguimiento continuo, la calificación de prácticas se basará exclusivamente en la calidad de la memoria final presentada. El peso global que tiene la parte de prácticas en la calificación final de la asignatura es de un 40%.

V. EVALUACIÓN DE LA METODOLOGÍA

La evaluación de satisfacción de los estudiantes con la metodología de prácticas expuesta anteriormente fue realizada mediante encuestas anónimas en las que realizamos una serie de preguntas para conocer si las prácticas les habían ayudado a mejorar la comprensión de la teoría, si la asistencia a clase les resultaba beneficiosa y si el diseño de las prácticas les gustaba. A continuación, presentamos tanto la encuesta como los resultados de la misma, así como los datos de asistencia y los resultados académicos obtenidos en la convocatoria ordinaria de la asignatura.

A. Encuesta

La encuesta fue realizada en la última de las sesiones de prácticas comprendiendo 14 preguntas en las que se les pedía que las valoraran de 0 a 10 (siendo 0 la valoración más negativa y 10 la más positiva). Además, la encuesta incluyó una última pregunta de texto libre para que los estudiantes indicaran cualquier aspecto que consideraran destacar y/o mejorar. Las preguntas que respondieron fueron las siguientes:

1) La práctica ha mejorado mi comprensión de los diferentes aspectos teóricos relacionados con el diseño de una red.
2) La topología empleada es adecuada para poder trabajar los distintos aspectos de red de la práctica final.
3) La práctica ha mejorado mi comprensión de RIP.
4) La práctica ha mejorado mi comprensión de OSPF.
5) La práctica ha mejorado mi comprensión de la interconexión entre dos protocolos de enrutamiento interno distintos.
6) La herramienta de Cisco, Packet Tracer, es útil para aprender los distintos aspectos de red cubiertos en la práctica.
7) La metodología de prácticas seguida para aprender los distintos conceptos me parece adecuada.
8) La práctica puede completarse dentro de las horas de laboratorio asignadas.
9) Prefiero asistir a las sesiones de prácticas y realizar un seguimiento semanal, a no asistir y realizar una entrevista final.
10) El profesorado ha proporcionado suficiente orientación y apoyo para completar las prácticas de manera satisfactoria.
11) El profesorado ha fomentado un ambiente participativo y propicio para resolver dudas y profundizar en los conceptos de la asignatura.
12) Las sesiones de trabajo autónomo en el laboratorio, con la posibilidad de consultar dudas al profesor presente, me resultan útiles para avanzar en las prácticas.
13) Prefiero asistir a las sesiones de trabajo autónomo en el laboratorio en vez de realizarlas de forma totalmente autónoma en casa.
14) El curso que viene recomendaría que las prácticas se realizaran de la misma forma.
15) Indica cualquier aspecto que cambiarías o mejorarías de la práctica (aspectos que se deban detallar más, nuevos contenidos, topología, herramienta utilizada, revisión de preguntas, metodología de enseñanza, evaluación, depuración y solución de problemas en la red, etc.).

B. Resultados de la encuesta

Las encuestas fueron respondidas por 111 personas de las 200 personas que asistieron alguna vez a clase, lo que representa el 55,5% de participación en la encuesta. Los resultados obtenidos se muestran en la Tabla I.

A nivel general, podemos observar que las medias de las respuestas oscilan entre 8,31 y 9,81, lo que indica que el estudiantado valoró positivamente la experiencia. Si, además, nos fijamos en que la mediana es de 9 o 10 en la mayoría de las preguntas, podemos señalar que la mayoría del estudiantado dio puntuaciones altas, lo que sugiere un alto nivel de satisfacción.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
DATOS DE LA ENCUESTA POR PREGUNTA

	Pregunta 1	**Pregunta 2**	**Pregunta 3**	**Pregunta 4**	**Pregunta 5**	**Pregunta 6**	**Pregunta 7**	**Pregunta 8**	**Pregunta 9**	**Pregunta 10**	**Pregunta 11**	**Pregunta 12**	**Pregunta 13**	**Pregunta 14**
Media	8,69	8,65	8,67	8,98	8,31	8,72	8,71	8,47	9,81	9,52	9,32	9,56	8,83	9,2
Mediana	9	9	9	9	8	9	9	9	10	10	10	10	10	10
Desviación estándar	1,21	1,59	1,27	1,21	1,64	1,57	1,57	2,25	0,64	0,989	0,926	0,722	1,79	1,28
Mínimo	4	3	5	3	2	1	1	1	6	4	6	7	0	1
Máximo	10	10	10	10	10	10	10	10	10	10	10	10	10	10

Analizando las preguntas relacionadas con el diseño de redes, topología y comprensión de aspectos teóricos de RIP, OSPF e interconexión, tienen medias cercanas o superiores a 8,31, lo que refleja que la práctica es efectiva para abordar estos contenidos. Con respecto al simulador CPT, podemos indicar que fue bien valorado (media de 8,71), por lo que parece que los estudiantes encuentran útil esta herramienta para aprender los conceptos prácticos.

Las preguntas relacionadas con la satisfacción con el profesorado y cómo éste les ha orientado (preguntas 10, 11 y 12) tienen medias superiores a 9,3, lo que muestra que los estudiantes perciben un apoyo significativo por parte del profesorado.

El análisis sobre los aspectos de la metodología de aprendizaje (pregunta 7) y la posibilidad de completar la práctica dentro de las horas asignadas (pregunta 8) también muestran valores altos (medias cercanas a 9), lo que sugiere que el enfoque metodológico y la planificación del tiempo son adecuados. Además, las sesiones autónomas en el laboratorio con apoyo docente también fueron valoradas como muy útiles (media de 9,56).

Por otro lado, si analizamos las preferencias del estudiantado, podemos observar que la mayoría de los estudiantes prefieren asistir a las sesiones de prácticas y realizar un seguimiento semanal, en lugar de no asistir y realizar una entrevista final con una media de 9,81. Esta es la media más alta de todas las preguntas y confirma que la entrevista de prácticas es un aspecto que prefieren evitar. En menor grado, vemos también que prefieren asistir a las sesiones de trabajo autónomo en el laboratorio en vez de realizarlas de forma totalmente autónoma en casa (con una media de 8,83).

Finalmente, una buena forma de saber si la satisfacción con las prácticas es alta es preguntándoles si las prácticas las realizarían de la misma forma para el curso siguiente. La media obtenida en esta pregunta fue de 9,2 por lo que podemos ver la alta satisfacción del estudiantado con el desarrollo de las prácticas.

En base a los resultados obtenidos, podemos concluir que hay un alto grado de satisfacción por cómo se ha desarrollado la metodología de prácticas en la asignatura.

C. Resultados académicos

La buena acogida de la metodología propuesta, según las conclusiones extraídas de la encuesta, es respaldada por los resultados académicos obtenidos. Tal y como se discute a continuación, la adopción de las sesiones semanales de seguimiento y el trabajo autónomo guiado ha permitido que los estudiantes avancen de manera progresiva, lo que ha resultado en la obtención de unas calificaciones, en general, altas.

La Tabla II presenta los datos agregados de la evaluación de las prácticas propuestas durante la convocatoria ordinaria de la asignatura. En primer lugar, destaca el alto porcentaje de presentados (80,78%), lo que demuestra la alta participación e implicación del alumnado. Además, se observa que de los 206 estudiantes presentados, la calificación media fue de 8,20, mientras que la mediana alcanzó un 8,55, con una desviación estándar de 1,72. Estos datos evidencian el buen desempeño general de los estudiantes con una notable concentración de puntuaciones altas. Además, como se aprecia en la Fig. 2, se obtuvo menos de un 6% de suspensos, con un 68,93% del total de los estudiantes presentados obteniendo calificaciones entre 8 y 10. Con estos resultados, la tasa de éxito en la convocatoria ordinaria fue del 94,17%, mientras que la de rendimiento fue del 76,67%. Estos resultados sugieren que el seguimiento continuo aplicado favoreció un entorno de aprendizaje estable y progresivo, donde la preparación constante se traduce en unos buenos resultados.

El seguimiento de la metodología propuesta se evidencia al observar la Fig. 3, que refleja una alta asistencia a las clases de prácticas, con la mayoría de los estudiantes participando en más del 75% de las sesiones. De los 53 estudiantes que no han asistido a ninguna sesión, tan solo 4 de ellos se presentaron en la convocatoria analizada, por lo que se concluye que el número de estudiantes que abordan las prácticas de forma totalmente autónoma es residual. Esta alta asistencia, impulsada por el seguimiento y guía semanal por parte del profesorado, ha contribuido significativamente a que el estudiantado mantenga un nivel adecuado de preparación reflejado en los resultados obtenidos.

Como hemos visto, las tasas de éxito y rendimiento alcanzadas, junto con la alta asistencia a las sesiones de prácticas, validan la efectividad de la metodología implementada, demostrando que un enfoque basado en *learning by doing* y seguimiento continuo puede reforzar significativamente el proceso de aprendizaje en asignaturas técnicamente complejas como Arquitectura de Redes.

VI. LECCIONES APRENDIDAS

A lo largo de este estudio, la implementación de la metodología basada en *learning by doing* y en una evaluación sin estrés ha dejado múltiples enseñanzas, destacando el valor de un enfoque eminentemente práctico y progresivo. El alumnado ha asimilado los conceptos de forma paulatina mediante trabajo autónomo y acompañamiento docente, favoreciendo un aprendizaje profundo y significativo. Los estudiantes refrendan esta percepción a través de sus respuestas a las preguntas enfocadas en este punto dentro de la encuesta realizada; se observa la alta valoración de las sesiones prácticas, donde

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla II
DATOS DE EVALUACIÓN DE PRÁCTICAS EN LA CONVOCATORIA ORDINARIA

Matriculados	Presentados	NP	Media	Mediana	DE	Mínimo	Máximo
253	206	47	8,2	8,55	1,72	1,2	10

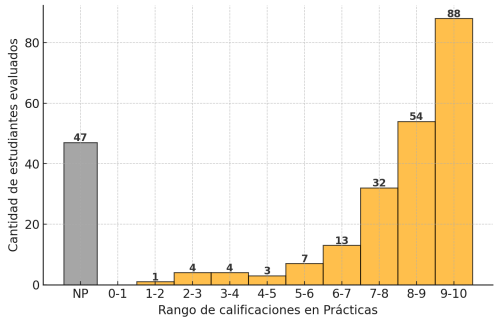


Fig. 2. Histograma de calificaciones de la nota de prácticas.

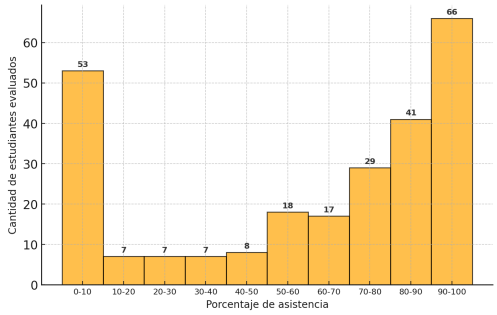


Fig. 3. Histograma de la asistencia a prácticas.

la retroalimentación constante es clave para consolidar el aprendizaje. Además, los estudiantes expresan su interés en ampliar este enfoque práctico a otros contenidos teóricos, considerándolo esencial para afianzar su proceso de aprendizaje.

En relación con el sistema de evaluación, el alumnado valora positivamente la eliminación de las entrevistas finales, destacando que se reduce la presión y permite centrarse en la entrega de la documentación. Sin embargo, se reconoce que este enfoque tiene áreas de mejora, como la menor práctica de habilidades de exposición oral y la dificultad para detectar desequilibrios en el trabajo en equipo. Esta carencia, desde una perspectiva crítica, resalta la importancia de replantear la evaluación para incluir instancias de exposición que fomenten la defensa oral sin añadir una carga de estrés excesiva.

En cuanto a las herramientas empleadas, el simulador CPT es bien aceptado, aunque se ha identificado que algunas funcionalidades, como la agregación de rutas externas, no se replican fielmente con respecto al funcionamiento de equipamiento real. A pesar de esta limitación, se considera que las ventajas aportadas por el uso del simulador son mayores que sus desventajas, ya que suelen resolverse gracias al seguimiento continuo y el apoyo del profesorado.

Finalmente, desde una perspectiva organizativa, se evidencia que el *feedback* del profesorado al alumnado es fundamental para conseguir un proceso de aprendizaje eficaz, facilitando avances progresivos. Asimismo, la distribución temporal de las sesiones y la carga de trabajo son adecuadas, lo que se refleja en la puntualidad de las entregas y en los buenos resultados obtenidos. El material de apoyo y las explicaciones proporcionadas reciben una valoración positiva, evidenciada en el bajo número de consultas registradas al respecto a lo largo del curso.

VII. CONCLUSIONES

En las asignaturas de redes es necesario adquirir no sólo un conocimiento teórico de conceptos como el diseño de topologías y la configuración de protocolos de enrutamiento, sino también unas destrezas prácticas para que los estudiantes puedan desarrollar habilidades que posteriormente utilizarán en entornos reales durante su vida profesional.

Para conseguir esas habilidades prácticas, hemos diseñado una metodología basada en *learning by doing* en la que el alumnado aprende paulatinamente y de forma práctica, mediante el simulador Cisco Packet Tracer, el diseño y la configuración de un escenario WAN que consideramos realista. En la metodología se ha fomentado que el aprendizaje esté centrado en el estudiante mediante un seguimiento semanal continuo por parte del profesorado. Esta supervisión constante permite una evaluación que no necesita de una entrevista final, reduciendo así la presión y el posible estrés que este tipo de pruebas suele generar en el alumnado.

Según los resultados presentados en este trabajo, la metodología de prácticas propuesta ha demostrado ser efectiva para la enseñanza de protocolos de enrutamiento y gestión de topologías de grandes dimensiones. Por un lado, su aplicación se ha traducido en unas altas tasas de éxito y rendimiento, así como de asistencia, lo que refleja su buena adopción por parte del alumnado. Además, se ha obtenido un alto grado de satisfacción debido al seguimiento continuo y a cómo se ha realizado el proceso de evaluación sin necesidad de entrevistas.

Como trabajo futuro, se propone extender el escenario de la práctica final para abordar otros contenidos teóricos, como la calidad de servicio, así como aplicar esta metodología a otras asignaturas del área de Ingeniería Telemática. También se plantea su adaptación a escenarios más complejos, incluyendo contenidos teóricos, con el fin de seguir promoviendo un aprendizaje significativo.

AGRADECIMIENTOS

Este trabajo ha sido apoyado por el proyecto PID2023-148104OB-C43 financiado por

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

MICIU/AEI/10.13039/501100011033 y por FEDER/UE y por el Ministerio de Economía y Transformación Digital, a través del proyecto CERBERUS-MINERVA (TSI-063000-2021-46). Los autores quieren también agradecer el apoyo recibido a través del grupo de *innovación docente para la enseñanza de las TIC y sus fundamentos*.

Referencias

[1] M. Prvan and J. Ozegovic, "Methods in teaching computer networks," *ACM Transactions on Computing Education (TOCE)*, vol. 20, pp. 1 – 35, 2020.

[2] J. Allison, "Simulation-Based Learning via Cisco Packet Tracer to Enhance the Teaching of Computer Networks," in *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1*, ser. ITiCSE '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 68–74.

[3] D. Tokmakov and A. Bekyarova-Tokmakova, "Teaching lab-based computer networking through performance-centered approach," in *ICERI2021 Proceedings*, ser. 14th annual International Conference of Education, Research and Innovation. IATED, November 2021, pp. 2200–2207.

[4] N. M. M. Noor, N. Yayao, and S. Sulaiman, "Effectiveness of using cisco packet tracer as a learning tool: A case study of routing protocol," *International Journal of Information and Education Technology*, vol. 8, no. 1, pp. 11–16, 2018.

[5] N. b. Abdul Rashid, M. Z. bin Othman, R. bin Johan, and S. F. bin Hj. Sidek, "Cisco packet tracer simulation as effective pedagogy in computer networking course," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 13, no. 10, p. pp. 4–18, Sep. 2019.

[6] J. Dewey, "Experience and education," in *The educational forum*, vol. 50, no. 3. Taylor & Francis, 1986, pp. 241–252.

[7] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*. FT press, 2014.

[8] J. Parker, R. Cupper, C. Kelemen, D. Molnar, and G. Scragg, "Laboratories in the computer science curriculum," *Computer Science Education*, vol. 1, no. 3, pp. 205–221, 1990.

[9] J. Gomez, E. F. Kfoury, J. Crichigno, and G. Srivastava, "A survey on network simulators, emulators, and testbeds used for research and education," *Computer Networks*, vol. 237, p. 110054, 2023.

[10] P. H. M. Gurgel, E. F. Barbosa, and K. R. L. J. C. Branco, "Teaching computer networks: A practical approach using virtualization tools," in *2013 IEEE Frontiers in Education Conference (FIE)*, 2013, pp. 1021–1026.

[11] R. C. Kizilirmak, A. Kassayev, and I. A. Ukaegbu, "A comparative study of simulated and hands-on experiments in teaching computer networks laboratory course," in *2023 IEEE Global Engineering Education Conference (EDUCON)*, 2023, pp. 1–4.

[12] A. Ruiz-Martínez, F. Pereñíguez-García, R. Marín-López, P. M. Ruiz-Martínez, and A. F. Skarmeta-Gómez, " Teaching Advanced Concepts in Computer Networks: VNUML-UM Virtualization Tool ," *IEEE Transactions on Learning Technologies*, vol. 6, no. 01, pp. 85–96, Jan. 2013.

[13] A. Ruiz Martínez, R. Marín López, F. Pereñíguez García, P. M. Ruiz Martinez, and A. F. Skarmeta, "Experiencia con la herramienta de virtualización vnuml para la enseñanza de redes de computadores," in *XVI Jornadas de Enseñanza Universitaria de la Informática*, 2010.

[14] F. Pereñíguez-García, A. Ruiz-Martínez, F. R. Muñoz, R. Marín-López, and P. Ruiz-Martínez, "Experimenting with different virtualization tools for the practical learning of computer networks," in *EDULEARN12 Proceedings*, 2012, pp. 714–721.

[15] A. Varga and R. Hornig, "An overview of the omnet++ simulation environment," in *1st International ICST Conference on Simulation Tools and Techniques for Communications, Networks and Systems*, 2010.

[16] G. Mwansa, M. R. Ngandu, and Z. S. Dasi, "Enhancing practical skills in computer networking: Evaluating the unique impact of simulation tools, particularly cisco packet tracer, in resource-constrained higher education settings," *Education Sciences*, vol. 14, no. 10, 2024.

[17] N. H. B. A. Long and S. E. B. Mokhtar, "Enhancing student understanding in data communication and networking course: Hands-on learning with cisco packet tracer." ICBE Publication.

[18] N. M. M. Noor, N. Yayao, and S. Sulaiman, "Effectiveness of using cisco packet tracer as a learning tool: A case study of routing protocol," *International Journal of Information and Education Technology*, vol. 8, pp. 11–16, 2018.

[19] S. Asadi, J. Allison, M. Khurana, and M. Nilashi, "Simulation-based learning for computer and networking teaching: A systematic literature review and bibliometric analysis," *Education and Information Technologies*, vol. 29, no. 12, pp. 15 655–15 690, Aug. 2024.

[20] A. Ruiz Martínez and P. M. Ruiz Martínez, "Enseñanza práctica de arquitectura de redes mediante la virtualización de escenarios de red," *Experiencias Docentes en Redes de Computadores (EDONET)*, vol. 1, no. 1, p. 67–81, 2015.

[21] M. Vijayalakshmi, P. Desai, and M. M. Raikar, "Packet tracer simulation tool as pedagogy to enhance learning of computer network concepts," in *2016 IEEE 4th International Conference on MOOCs, Innovation and Technology in Education (MITE)*, 2016, pp. 71–76.

[22] N. Sarkar, "Teaching computer networking fundamentals using practical laboratory exercises," *IEEE Transactions on Education*, vol. 49, no. 2, pp. 285–291, 2006.

[23] E. M. Castañeda-González, "Estrategias didácticas aplicadas a la enseñanza de las redes de computadoras en la carrera de ingeniería en telecomunicaciones y electrónica," *Maestro y Sociedad*, vol. 13, no. 3, p. 368–378, jul. 2016.

[24] A. Grotta and E. P. V. Prado, "Benefits of the project-based learning to cope with computer programming education: A systematic literature review," in *Conference Proceedings of Pbl2019 Immersive Virtual International Conference Hello Future*, 2019, pp. 12–13.

[25] L. Zhang and Y. Ma, "A study of the impact of project-based learning on student learning effects: a meta-analysis study," *Frontiers in Psychology*, vol. 14, 2023.

[26] C. Calderón Patier and G. Escalera Izquierdo, "La evaluación de la docencia ante el reto del espacio europeo de educación superior (eees)," *Educación XX1*, vol. 11, pp. 237–256, 2008.

[27] T. Ringeisen, S. Lichtenfeld, S. Becker, and N. Minkley, "Stress experience and performance during an oral exam: the role of self-efficacy, threat appraisals, anxiety, and cortisol," *Anxiety, Stress, & Coping*, vol. 32, no. 1, pp. 50–66, sep 2018.

[28] S. M. Ismail, D. R. Rahul, I. Patra, and E. Rezvani, "Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill," *Language Testing in Asia*, vol. 12, no. 1, p. 40, Sep. 2022.

[29] D. Boud, *Enhancing Learning Through Self-assessment*. London: Routledge, Oct. 2013.

[30] N. v. d. Embse, D. Jester, D. Roy, and J. Post, "Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review," *Journal of Affective Disorders*, vol. 227, p. 483–493, 2018.

[31] Y. Rodríguez Rincón, A. Munárriz, and A. Magreñán Ruiz, "A new approach to continuous assessment: Moving from a stressful sum of grades to meaningful learning through self-reflection," *The International Journal of Management Education*, vol. 22, no. 3, p. 101072, Nov. 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Rendimiento académico y asistencia: Un estudio con métodos estadísticos y LLM

Luis Bote-Curiel, Mihaela I. Chidean
Dept. Teoría de la Señal y Comunicaciones, Universidad Rey Juan Carlos, Fuenlabrada, España
luis.bote@urjc.es, mihaela.chidean@urjc.es

En este trabajo se analiza el impacto de la implementación de una política de asistencia obligatoria en el rendimiento académico en la asignatura de Transmisión Digital en el Grado en Ing. Telemática de la Univ. Rey Juan Carlos. Se emplean métodos estadísticos y de análisis de opinión para evaluar el desarrollo de la asignatura entre los cursos académicos 2018-19 y 2024-25. El conjunto de datos analizado incluye las calificaciones finales obtenidas y las opiniones anónimas recabadas al finalizar la evaluación. Los resultados confirman la mejoría en el rendimiento académico del estudiantado al implementar esta política de asistencia, a la vez que revelan opiniones divididas sobre la misma. Se destaca la tensión entre el fomento de la participación y entre la autonomía y motivación intrínseca del estudiantado. Estos resultados plantean la necesidad de adaptar la implantación de esta política para equilibrar la asistencia regular con el desarrollo de la autonomía en la educación superior.

***Palabras Clave*—asistencia a clase, rendimiento académico, métodos estadísticos, modelo de lenguaje de gran tamaño, LLM**

I. INTRODUCCIÓN

La sociedad actual está marcada por la hiperconectividad, en la que el acceso instantáneo a la información, los servicios y la comunicación global ha transformado profundamente nuestras formas de vivir, trabajar y aprender [1], [2]. Esta era digital, impulsada por el desarrollo acelerado de las tecnologías de la información y la comunicación, ha dado lugar a lo que algunos autores denominan la "economía de la prisa", caracterizada por la inmediatez, la multitarea y una constante presión por optimizar el uso del tiempo [3]. En este contexto, el valor del tiempo se ha intensificado como recurso escaso, lo que influye tanto en los hábitos de consumo como en las dinámicas educativas. La educación superior no escapa a esta lógica: el compromiso del estudiantado con su formación se ve condicionado por múltiples estímulos externos, demandas laborales, sociales y una creciente expectativa de eficiencia en la adquisición de conocimientos [4]. Esta realidad plantea desafíos en el diseño de estrategias docentes y la reflexión sobre el papel que juegan aspectos como la asistencia, el ritmo de aprendizaje o la autonomía en los procesos formativos [5], [6].

En el ámbito de la innovación docente, se han explorado numerosas estrategias para aumentar la motivación y el compromiso de los estudiantes universitarios. Entre estas, destacan el uso de tecnologías interactivas como aplicaciones móviles para la participación en clase [5], la gamificación para fomentar el aprendizaje activo y el aprendizaje basado en proyectos para conectar los contenidos teóricos con situaciones reales. Además, se han implementado metodologías como las clases invertidas (*flipped classroom*) para promover una mayor implicación fuera del aula y un aprovechamiento más efectivo del tiempo en las sesiones presenciales. Sin embargo, a pesar de estos avances, algunos estudios alertan sobre la creciente desconexión y distracción en el entorno educativo moderno [6]. En respuesta a esta tendencia, nuestra experiencia propone una medida directa y deliberada: la implementación de un política de asistencia obligatoria a las clases.

Este estudio explora si la asistencia regular y la participación activa en el aula pueden ser factores relevantes en la concentración, el seguimiento de los contenidos y el rendimiento académico de los estudiantes. Para validarlo, se analizarán los datos recopilados en un caso de estudio específico empleando métodos estadísticos y un modelo de lenguaje de gran tamaño (LLM, *Large Language Model*). En concreto, se analizarán los resultados académicos obtenidos en la asignatura de Transmisión Digital, impartida en el tercer curso del Grado en Ingeniería Telemática (GIT) de la Universidad Rey Juan Carlos (URJC) en los últimos siete cursos académicos. En los dos últimos cursos bajo estudio se ha implementado una estrategia de asistencia obligatoria. La motivación de implementar esta estrategia se debe a que tras la etapa post-COVID, el equipo docente ha observado una disminución significativa en la asistencia presencial a clase. Con esta iniciativa, además de intentar revertir esta tendencia, se pretende

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

obtener evidencias que permitan argumentar ante el estudiantado los beneficios de una participación activa y constante en su propio proceso de aprendizaje.

Aunque en los grados adaptados al Espacio Europeo de Educación Superior (EEES) se asume que la asistencia regular, el trabajo diario y la evaluación continua forman parte del modelo educativo, la práctica real puede diferir considerablemente. Al menos en el caso del GIT de la URJC, la situación es otra. Tras revisar las guías docentes del curso 2024-25, se observa que, de las 40 asignaturas que componen el grado, la mayoría aplica algún tipo de evaluación continua basada en entregas periódicas y pruebas distribuidas a lo largo del cuatrimestre. No obstante, salvo TxDig y cuatro asignaturas con prácticas obligatorias en laboratorios docentes con equipamientos específicos, el resto no incluye de forma explícita la asistencia obligatoria a todas las sesiones de la asignatura con independencia de su tipología como parte del sistema de evaluación.

El análisis de los datos se fundamenta en métodos estadísticos ampliamente aceptados para el tipo de variables estudiadas. La fase inicial consiste en un análisis descriptivo para la caracterización de la muestra, detallando la distribución de los resultados académicos (aprobados y suspensos) en los cursos bajo estudio. Esta fase de exploración de datos es una práctica estándar y proporciona el contexto necesario [7]. A continuación, para comparar los resultados entre los grupos de estudiantes que asistieron y no asistieron a clase, se emplea la prueba de independencia chi-cuadrado (χ^2) [8]. La elección de este test estadístico se justifica por ser la herramienta de referencia para analizar la relación entre variables de naturaleza categórica, permitiendo, en este caso concreto, determinar si la frecuencia observada de éxito académico en cada grupo difiere significativamente de la frecuencia esperada bajo la hipótesis nula de independencia.

Por último, el documento está organizado de la siguiente manera. La Sección II se describe la población bajo estudio, así como las herramientas de análisis empleadas. A continuación, la Sección III describen e interpretan los resultados obtenidos. El trabajo finaliza con la Sección IV, donde se resumen las conclusiones alcanzadas y los posibles trabajos futuros.

II. METODOLOGÍA

En esta sección se describe en detalle la metodología empleada. Se contextualiza el trabajo, presentando el entorno en el que se desarrolló y las características de los participantes. Se definen, también, los conjuntos de datos que constituyen la base del trabajo: las calificaciones finales del alumnado y las encuestas de opinión. Finalmente, se exponen los métodos de análisis, tanto cuantitativos como cualitativos, que se aplican a dichos conjuntos de datos para la obtención de los resultados.

A. La asignatura de Transmisión Digital

TxDig es una asignatura obligatoria de 6 ECTS del primer cuatrimestre del tercer curso del GIT. TxDig, junto con Teoría de la Comunicación y Tratamiento de Información Multimedia son la tres únicas asignaturas del

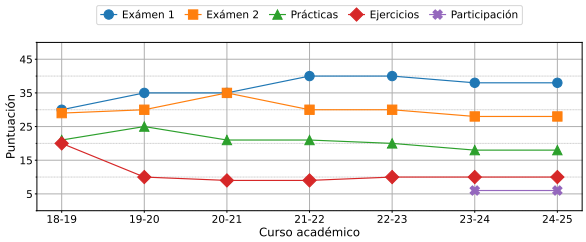


Fig. 1. Evolución del porcentaje asignado en la nota a los ítems de calificación considerados por curso académico.

módulo de "Sistemas de comunicación y audiovisuales" del plan de estudios [9]. Los docentes que imparten la asignatura son del ámbito de la Teoría de la Señal y Comunicaciones.

La asignatura cubre fundamentos de teoría de la información, codificación de fuentes y canales, y técnicas de corrección de errores. También se abordan receptores óptimos en entornos con ruido e interferencia, así como sistemas de acceso múltiple, multiportadora y de espectro ensanchado. A la hora de impartir estos contenidos, el equipo docente se centra en personalizar y contextualizar las explicaciones dentro de la especialidad del grado, utilizando ejemplos y aplicaciones basados en estándares y tecnologías comúnmente empleados en el ámbito de la ingeniería telemática.

El sistema de evaluación de la asignatura se ha mantenido mayoritariamente estable a lo largo de los distintos cursos académicos y se estructura en tres componentes principales [10]: (i) dos exámenes de teoría y problemas, que cubren aproximadamente la mitad del temario cada uno, son re-evaluables en la convocatoria extraordinaria y tienen un peso conjunto en la nota final que oscila entre el 59% y el 70%; (ii) una evaluación de las prácticas realizadas con MATLAB, también re-evaluable, con un peso comprendido entre el 18% y el 25%; y (iii) la entrega de ejercicios seleccionados a lo largo del cuatrimestre, no re-evaluable y con un peso que varía entre el 9% y el 20%. La Fig. 1 muestra la evolución de estos pesos en los cursos académicos bajo estudio.

Durante el curso 2020-21, debido a las medidas excepcionales adoptadas por la COVID-19, el sistema de evaluación se adaptó ligeramente mediante la incorporación de pruebas adicionales asociadas a cada tema impartido. La calificación final se determinaba como el valor máximo entre dos opciones: una que incluía estas pruebas adicionales y otra que no, con el objetivo de favorecer al estudiantado en un contexto especialmente complejo y garantizar la equidad en la evaluación. No obstante, dado que las ponderaciones entre los apartados de teoría, prácticas y ejercicios se mantuvieron constantes, en este estudio no se realizará una distinción específica para dicho curso académico.

A partir del curso académico 2023-24 se implementa en TxDig la política de asistencia obligatoria. Dicha política requiere un mínimo del 70% de participación en las sesiones presenciales para poder ser evaluado en la convo-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
NÚMERO DE ESTUDIANTES MATRICULADOS Y FRECUENCIAS ABSOLUTAS Y RELATIVAS DE ESTUDIANTES APROBADOS Y SUSPENSOS POR CURSO ACADÉMICO. EN LOS CURSOS 2023-24 Y 2024-25, EL FONDO AZUL INDICA LOS RESULTADOS PARA LOS ESTUDIANTES QUE HAN CUMPLIDO EL MÍNIMO DE 70% DE ASISTENCIA REQUERIDO.

	Número matriculados	Aprobados (%)	Suspensos (%)
2018-19	14	3 (27.3%)	8 (72.7%)
2019-20	19	12 (63.2%)	7 (36.8%)
2020-21	21	5 (26.3%)	14 (73.7%)
2021-22	23	3 (13.0%)	20 (87.0%)
2022-23	22	5 (22.7%)	17 (77.3%)
2023-24	23	1 (14.3%)	6 (85.7%)
		11 (68.8%)	5 (31.2%)
2024-25	18	0 (0%)	6 (100%)
		6 (50%)	6 (50%)

catoria ordinaria. Como parte de esta intervención, el incumplimiento de dicho criterio resulta en una calificación de "Suspenso" en esa convocatoria, aunque se conservan las notas parciales para la evaluación de la convocatoria extraordinaria. Como incentivo adicional, en este periodo se introduce también un ítem de evaluación adicional: la participación. Este ítem evalúa la participación activa y eficiente en las actividades formativas realizadas y el hecho de que el estudiantado toma iniciativa y es agente activo en el proceso enseñanza-aprendizaje.

B. Conjunto de datos y métodos estadísticos

El primer conjunto de datos incluye las calificaciones finales del alumnado matriculado en la asignatura TxDig en los cursos académicos 2018-19 a 2024-25. En la Tabla I se presentan los datos procesados a partir de un análisis descriptivo, que incluye las frecuencias absolutas y relativas de alumnado matriculado en cada curso académico. Se distingue entre quienes cumplen o no el criterio de asistencia, así como la distribución de los resultados académicos, diferenciando entre aprobados y suspensos. Este enfoque exploratorio y descriptivo permite caracterizar la muestra y observar tendencias generales en el rendimiento académico a lo largo de los diferentes cursos.

Posteriormente, para los cursos 2023-24 y 2024-25, se construye la Tabla II, que representa una tabla de contingencia que agrupa los resultados académicos (aprobado o suspenso) en función del cumplimiento del criterio de asistencia. Sobre estos datos agregados, se aplica el test estadístico de independencia chi-cuadrado (χ^2), con el objetivo de determinar si existe una asociación estadísticamente significativa entre la asistencia a clase y el éxito académico. Este test permite contrastar la hipótesis nula de independencia entre ambas variables, evaluando si la proporción de aprobados y suspensos difiere significativamente entre los estudiantes que cumplen con el requisito de asistencia y aquellos que no lo hacen.

C. Encuesta de opinión y LLM

El segundo conjunto de datos incluye las respuestas recopiladas en una encuesta de opinión realizada con el objetivo de recopilar la experiencia del estudiantado. Dicha encuesta incluye una variedad de preguntas enfocadas a

Tabla II
TABLA DE CONTINGENCIA DE LOS RESULTADOS ACADÉMICOS (APROBADO/SUSPENSO) EN FUNCIÓN DEL CUMPLIMIENTO DEL CRITERIO DE ASISTENCIA EN LOS CURSOS ACADÉMICOS 2023-24 Y 2024-25.

	Aprobados	Suspensos
No cumple mínimo asistencia	1	12
Cumple minimo asistencia	17	11

evaluar la experiencia en la asignatura y, con más énfasis, la experiencia con respecto a la política de asistencia obligatoria implementada. Esta encuesta se ha realizado únicamente en los cursos donde se ha implementado esta política. El cuestionario es anónimo y de participación voluntaria, abarcando aspectos como la carga académica, el interés en la materia y los logros percibidos, buscando una visión integral de la vivencia del estudiantado.

Se profundizó en la percepción de la política de asistencia obligatoria empleando un total de ocho preguntas. Se incluyen preguntas como "¿Qué opinas sobre la política de asistencia obligatoria en esta asignatura?" y "¿Cómo afectó la política de asistencia obligatoria a tu enfoque y compromiso con la asignatura?". Además, se invitó a la reflexión sobre el futuro de esta política mediante preguntas como "¿Consideras que la política de asistencia obligatoria debería mantenerse, modificarse o eliminarse? ¿Por qué?" y "¿Qué impacto crees que tendría la eliminación de la asistencia obligatoria en el rendimiento académico del estudiantado?". La encuesta también incluye la posibilidad de añadir comentarios y sugerencias adicionales, buscando una retroalimentación completa y constructiva.

La encuesta ha contado con la participación de 29 personas. Si bien todas las preguntas han sido respondidas, el nivel de detalle y profundidad varía entre las respuestas individuales, reflejando la diversidad de experiencias y opiniones dentro del grupo. Se realiza un análisis cualitativo de las respuestas obtenidas, empleando el LLM Gemini 2.0 Flash [11] al que se le proporcionó la siguiente instrucción: "*Agrupa y resume las respuestas a las preguntas 18 a 25, identificando las principales categorías de opinión y destacando los comentarios más representativos, tanto positivos como críticos.*". La Sección III presenta el resultado de este análisis, así como la discusión asociada, ambos basados en la salida del modelo de inteligencia artificial empleado.

III. RESULTADOS Y DISCUSIÓN

En esta sección se presentan e interpreta el impacto de los resultados del trabajo derivados del análisis de los conjuntos de datos de calificaciones finales del alumnado y las encuestas de opinión descritos previamente. Igualmente, se presentan los resultados obtenidos en las Encuestas de Valoración Docente (EVD) de la asignatura a lo largo de los cursos, así como una breve discusión de las limitaciones del presente estudio.

A. Métodos estadísticos

En la Tabla I se observa que, durante los cursos sin requisito de asistencia, la proporción de aprobados y suspen-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

sos se mantiene relativamente constante. El curso 2019-20 constituye una excepción, donde la tasa de aprobados aumenta notablemente debido a las circunstancias excepcionales derivadas de la pandemia de COVID-19 y a la evaluación adaptada que se llevó a cabo en la convocatoria extraordinaria. Sin embargo, en los dos cursos en los que se implementó la obligatoriedad de la asistencia, los resultados académicos muestran una clara diferenciación entre los estudiantes que cumplieron con el criterio de asistencia y aquellos que no lo hicieron. Los estudiantes que no asistieron a clase presentan tasas de aprobados y suspensos similares a las de los cursos anteriores, mientras que entre quienes sí cumplieron con la asistencia obligatoria se observa una mayor proporción de aprobados y una reducción significativa de los suspensos. Estos datos sugieren que la asistencia regular a clase podría estar asociada a un mejor rendimiento académico en el nuevo contexto normativo.

Para analizar si la diferencia observada en los resultados académicos entre los estudiantes que cumplieron con el criterio de asistencia y aquellos que no lo hicieron es estadísticamente significativa, se aplica la prueba de independencia χ^2 sobre la tabla de contingencia representada en la Tabla II. Esta prueba permite evaluar la existencia de una posible asociación entre la asistencia a clase y el rendimiento académico, contrastando las siguientes hipótesis:

- Hipótesis nula: No existe relación entre la asistencia a clase y el resultado académico (aprobado o suspenso); es decir, ambas variables son independientes.
- Hipótesis alternativa: Existe una relación entre la asistencia a clase y el resultado académico, es decir, ambas variables están asociadas.

La aplicación de la prueba arrojó un valor del estadístico χ^2 de 8.10, con 1 grado de libertad y un p-valor de 0.0044. Dado que el p-valor es menor que el umbral de significación habitual (0.05), se rechaza la hipótesis nula de independencia entre la asistencia a clase y el resultado académico y, por tanto, se acepta la hipótesis alternativa de que existe una relación. Es decir, este resultado indica que sí existe una asociación estadísticamente significativa entre el cumplimiento del criterio de asistencia y el rendimiento académico, de modo que la probabilidad de aprobar o suspender está relacionada con la asistencia a clase en el contexto de los cursos analizados.

Estos resultados obtenidos, que demuestran una asociación estadísticamente significativa entre el cumplimiento del criterio de asistencia y el rendimiento académico, están respaldados por la literatura científica. En concreto, estos resultados concuerdan con las conclusiones a las que llega un extenso meta-análisis que establece que la asistencia a clase es el predictor más potente de las calificaciones universitarias, superando incluso a indicadores tradicionales como las notas de acceso a la universidad o el expediente de bachillerato [12]. La relación que encontramos en nuestra asignatura, por tanto, no es un caso aislado, sino que confirma que el simple acto de asistir a clase está intrínsecamente ligado al éxito del estudiantado.

B. Encuesta de opinión

Las respuestas obtenidas en las preguntas 18 a 25 de la encuesta de opinión muestran opiniones dispares con respecto a la política de asistencia obligatoria implantada. La mayoría valora alguno de los aspectos relacionados y beneficios derivados de esta política, como puede ser el mantenerse al día con el material docente. Sin embargo, como es de esperar en un contexto de educación no obligatoria, otras personas consideran esta política innecesaria o excesivamente restrictiva, defendiendo en su justificación, por ejemplo, la autonomía del estudiantado. Como ejemplo de comentarios dispares, destacamos los siguientes comentarios de la primera categoría:

- "*En mi caso sí, al verme obligado a venir, prestaba más atención.*" o
- "*Me ha ayudado a ser más constante y no dejarlo todo para el final.*"

y de la segunda categoría:

- "*Me desanimó por completo, me sentí obligado y con poca o nula intención propia de asistir y aprender en la asignatura.*" o
- "*El que no tiene interés no lo va a tener y menos viéndose obligado a venir.*".

Las dos categorías de opiniones que se han observado en nuestros resultados están alineadas con otros hallazgos del estado del arte. Para el equipo docente es particularmente preocupante la percepción de que la asistencia obligatoria restringe la autonomía del estudiantado, aspecto que podría incluso tener un impacto negativo en la motivación intrínseca [13]. Esta disminución de la motivación podría, en algunos casos, conducir incluso a resultados opuestos a los deseados; por ejemplo, algunas personas han manifestado que, en ausencia de la política de asistencia obligatoria, asistirían a más sesiones al no sentir la presión asociada a la obligatoriedad. Sin embargo, establecida la relación existente entre asistencia y rendimiento académico, tanto en este trabajo como en otros estudios como [12] o [14], consideramos fundamental valorar también estrategias alternativas que fomenten la autonomía, la participación activa y el compromiso personal con el aprendizaje.

Al considerar las preguntas relacionadas con su opinión sobre las ventajas y desventajas, así como la continuidad de esta política de asistencia obligatoria, también se encuentra opiniones interesantes, variadas y con bastante utilidad para el equipo docente para adaptar los requisitos a lo largo de los cursos venideros. Por un lado, sus respuestas muestran que entienden la motivación de esta política e incluso la valoran, con comentarios como pueden ser

- "*Creo que la asignatura tiene conceptos que tienen que quedar muy claros y se consigue yendo a clase.*"

Otras respuestas también muestran el interés en el proceso de mejora continua aportando ideas para modificar esta política, con comentarios como pueden ser:

- "*Ventajas: ayuda a llevar la asignatura al día, mejora comprensión; inconvenientes: para alumnos en segunda o mayor matrícula se hace repetitiva.*"

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Finalmente, otra serie de comentarios muestran el rechazo a la continuidad de la política, aportando argumentos referentes a la etapa educativa y a su propia madurez:

- "*Puedo entender que para aquellos alumnos en primera matrícula, la asistencia obligatoria pueda tener beneficios. Aun así creo que en la universidad no debería aplicarse.*"
- "*Demasiada presión, ya no somos pequeños, deberíamos de poder decidir cuando queremos venir.*"

La variedad de opiniones coincide tanto con lo esperado por parte del equipo docente y como con la literatura de especialidad. La naturaleza no obligatoria de la educación universitaria y el hecho de que el estudiantado son personas adultas son aspectos clave a considerar en la decisión sobre la continuidad de esta política de asistencia obligatoria. No obstante, nuestra responsabilidad como docentes también nos lleva a buscar alternativas que mejoren el aprendizaje y, por consiguiente, el rendimiento académico.

Por un lado, es importante recordar que tanto en este trabajo como en otros se muestra que la asistencia tiene un impacto positivo en el rendimiento académico [12], [14]. Por otro, es fundamental reconocer que el fomento de la autonomía y la responsabilidad como persona adulta es también un aprendizaje esencial en esta etapa educativa [15], [13]. Y también es fundamental recordar el elemento que motivó la implementación de esta política en la asignatura: el cambio de tendencia observado tras la disrupción vital que ha causado la COVID-19. ¿Qué hacer en este escenario y en el futuro próximo? Esta es la pregunta que todos los docentes nos hacemos antes de comenzar cada curso. En este caso concreto, se prevé continuar con esta política en el siguiente curso académico, incorporando modificaciones que fomenten la autonomía y la responsabilidad en la educación superior.

C. Encuesta de valoración docente (EVD)

El Sistema Interno de Garantía de Calidad de la URJC [16] se encarga de realizar las EVD para cada asignatura impartida, ofreciendo una perspectiva externa e independiente de la opinión del estudiantado. Al finalizar cada cuatrimestre, el estudiantado tiene la oportunidad de completar, de manera voluntaria y anónima, una encuesta en la cual se califican los diferentes aspectos de cada asignatura. En la Tabla III se muestran las valoraciones globales obtenidas para la asignatura, junto con el porcentaje de participación en la encuesta, para cada curso académico bajo estudio. Igualmente, se muestran los promedios de las valoraciones globales de todas las asignaturas de tercer curso del grado y del grado en su totalidad. No se dispone aún de valores promediados del tercer curso y del grado para el curso académico 2024-25.

Se observa que la valoración de TxDig ha fluctuado a lo largo de los cursos, con un pico en el curso 2021-22 y 2024-25, y valores más bajos en los cursos 2018-19 y 2022-23. Es conveniente observar también la variabilidad en los porcentajes de participación en las EVD de TxDig, una variabilidad que podría influir en la representatividad de las valoraciones obtenidas. En general, salvo contadas excepciones, la valoración de TxDig es inferior tanto al promedio de las asignaturas de tercer curso del GIT como al promedio de todas las asignaturas de GIT. El equipo docente considera esta tendencia en la valoración de TxDig como esperable, en especial teniendo en cuenta el rendimiento de la asignatura a lo largo de los cursos académicos. Aún así, en este caso, a pesar de que la valoración es superior al umbral de alerta de 3.5, esta tendencia ha animado la implantación de la política de asistencia bajo estudio en este trabajo.

Tabla III
RESULTADOS DE LAS EVD PARA CADA CURSO ACADÉMICO. SE PRESENTAN RESULTADOS DE TXDIG, DEL PROMEDIO DE LAS ASIGNATURAS DEL TERCER CURSO DE GIT Y DEL PROMEDIO DE TODAS LAS ASIGNATURAS DEL GIT (ESCALA DE 1 A 5).

	TxDig (participación)	Promedio asignaturas tercero GIT	Promedio asignaturas GIT
2018-19	3.2 (85.7%)	3.84	3.82
2019-20	3.7 (52.6%)	4.15	4.20
2020-21	3.8 (23.8%)	4.15	4.20
2021-22	4.6 (8.7%)	4.09	4.14
2022-23	3.4 (31.8%)	4.09	4.14
2023-24	3.66 (47.8%)	4.22	4.31
2024-25	4.58 (50%)	-	-

Como hecho a destacar, las EVD obtenidas para los dos cursos en los cuales se ha puesto en práctica la política de asistencia obligatoria se han mantenido estables. Este resultado es alentador y, en combinación con los resultados presentados en las dos secciones anteriores, refuerza la idea de que la dirección tomada en esta asignatura parece adecuada.

D. Limitaciones del estudio

Consideramos que este estudio aporta evidencia valiosa sobre el impacto de la asistencia obligatoria en el rendimiento académico, en especial en la asignatura bajo estudio. Sin embargo, es importante reconocer sus limitaciones. En primer lugar, el tamaño del grupo de la asignatura y en especial de la muestra de la encuesta de opinión es modesto, lo que podría limitar la generalización de los resultados a la totalidad del estudiantado del GIT de la URJC. Además, la participación voluntaria en la encuesta introduce un posible sesgo de auto-selección, ya que es probable que los estudiantes con opiniones más firmes sobre la política de asistencia (ya sean a favor o en contra) hayan estado más motivados para participar.

En segundo lugar, los resultados de este estudio deben interpretarse en el contexto específico de la asignatura TxDig, impartida en un grado de ingeniería en una universidad concreta. La efectividad de la política de asistencia obligatoria podría variar en otras disciplinas, niveles educativos o instituciones, donde las características del estudiantado, los métodos de enseñanza y las políticas académicas podrían ser diferentes. Es importante reconocer que incluso el contexto socioeconómico del estudiantado puede influir en su rendimiento académico y en su capacidad para asistir regularmente a clase.

En tercer lugar, aunque se han considerado y controlado diversas variables relevantes, es posible que otros factores

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

no medidos (como el nivel de motivación intrínseca de los estudiantes, su situación personal o su acceso a recursos de aprendizaje adicionales) hayan influido en los resultados académicos. Finalmente, el análisis cualitativo de las respuestas de la encuesta se realizó con un LLM concreto. Aunque este análisis nos permite identificar patrones y temas relevantes, también es necesario considerar los posibles sesgos algorítmicos o interpretaciones erróneas de los matices del lenguaje.

Por lo tanto, aunque estudio nos aporta información valiosa, reconocemos la importancia de abordar sus limitaciones en futuras investigaciones. En este sentido, las próximas evaluaciones buscarán ampliar la muestra de la encuesta, considerar el contexto socioeconómico del estudiantado y explorar métodos de análisis cualitativo que minimicen los sesgos algorítmicos.

IV. CONCLUSIÓN

En este trabajo hemos analizado el impacto que ha tenido la implantación de una política de asistencia obligatoria a clase en el rendimiento académico en una asignatura de tercer curso, TxDig, del GIT implantado en la URJC. El análisis se ha realizado empleando métodos estadísticos y herramientas de inteligencia artificial para el análisis de los resultados académicos durante los cursos 2018-19 a 2024-25 y de las opiniones del estudiantado, respectivamente. Los resultados obtenidos muestran una asociación estadísticamente significativa entre la asistencia regular y la probabilidad de aprobar la asignatura, evaluada en varios cursos donde se ha implantado una política de asistencia obligatoria. El análisis cualitativo revela una diversidad de percepciones entre el estudiantado. Si bien una parte valora la obligatoriedad como un incentivo para mantener el ritmo y la constancia, otra considera que puede limitar la autonomía y la motivación intrínseca, aspectos fundamentales en la educación superior. Esta dualidad coincide con lo reportado en la literatura y plantea la necesidad de buscar un equilibrio entre el fomento de la participación activa y el respeto a la autonomía del alumnado.

Como líneas de trabajo futuro, estamos trabajando en mejorar el diseño de las encuestas de opinión para obtener una retroalimentación más rica y representativa. Igualmente, estamos valorando el incluir como método de evaluación adicional de la asignatura el uso de incentivos alternativos que promuevan la participación voluntaria y el compromiso personal, reduciendo así la necesidad de medidas obligatorias. Asimismo, se plantea profundizar en el análisis mediante LLM para evaluar de forma más precisa las opiniones y sugerencias del estudiantado, lo que nos permitirá tomar decisiones más informadas sobre la evolución de la asignatura en los cursos venideros.

AGRADECIMIENTOS

Este trabajo ha sido financiado, en parte, por la Comunidad de Madrid en el marco del convenio 2023-2026 con la Univ. Rey Juan Carlos para la concesión de subvenciones directas para el fomento y la promoción de la investigación y la transferencia tecnológica, Línea A: Doctores Emergentes, Proyecto OrdeNGN (Ref. F1177), y también por el Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación (MICIU/AEI/10.13039/501100011033) a través de la ayuda PID2023-151462OB-I00.

REFERENCIAS

[1] International Telecommunication Union (ITU), "Measuring digital development facts and figures 2021," 2021, disponible en: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2021.pdf.
[2] Cisco Systems, Inc., "Cisco Annual Internet Report (2018–2023)," 2020, disponible en: http://www.cisco.com/.
[3] H. Rosa, *Social Acceleration: A New Theory of Modernity*. New York: Columbia University Press, 2013.
[4] S. Hess and F. Tremblay, "Student Engagement and the Role of Technology," *Humans*, vol. 4, no. 4, p. 23, 2024.
[5] C. Guerrero *et al.*, "Use of Mobile Devices in the Classroom to Increase Motivation and Participation of Engineering University Students," *IEEE Latin America Transactions*, vol. 14, no. 1, pp. 411–416, 2016.
[6] M. Pérez-Juárez, D. González-Ortega, and J. Aguiar-Pérez, "Digital Distractions from the Point of View of Higher Education Students," *Sustainability*, vol. 15, no. 7, p. 6044, 2023.
[7] A. B. Downey, *Think Stats: Exploratory Data Analysis*, 3rd ed. O'Reilly Media, Inc., 2025.
[8] A. Agresti, *Categorical data analysis*, 3rd ed. John Wiley & Sons, 2013.
[9] Ministerio de Ciencia, Innovación y Universidades, "Registro de Universidades, Centros y Títulos (RUCT)," disponible en: https://www.educacion.gob.es/ruct/home.
[10] Universidad Rey Juan Carlos, "Guías Docentes URJC," disponible en: https://gestion3.urjc.es/guiasdocentes/.
[11] Google DeepMind, "Gemini 2.0 Flash: Large Language Model," 2024, disponible en: https://deepmind.google/technologies/gemini/.
[12] M. Credé, S. G. Roch, and U. M. Kieszczynka, "Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics," *Review of Educational Research*, vol. 80, no. 2, pp. 272–295, 2010.
[13] R. M. Ryan and E. L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary Educational Psychology*, vol. 25, no. 1, pp. 54–67, 2000.
[14] P. Friedman, F. Rodriguez, and J. McComb, "Why students do and do not attend classes: Myths and realities," *College Teaching*, vol. 49, no. 4, pp. 124–133, 2001.
[15] D. Boud, "Moving towards autonomy," in *Developing student autonomy in learning*, 2nd ed., D. Boud, Ed. London: Kogan Page, 1988, pp. 17–39.
[16] Universidad Rey Juan Carlos. (2024) Calidad. [Online]. Available: https://www.urjc.es/universidad/calidad

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Transmisión de vídeo HD en Canal Físico Restringido usando Tiled Stream Time Division Multiplexing y evaluación de su calidad

David Gómez, Miguel Garcia-Pineda, Mario Montagud.
Departamento de Informática.

ETSE-UV. Universidad de València.

Av. de la Univesitat, s/n. 46701. Burjassot. València.
dagomi@alumni.uv.es, miguel.garcia-pineda@uv.es, mario.montagut@uv.es.

Los sistemas de videoportero instalados masivamente en edificios residenciales emplean todavía un par bifilar sin apantallar que apenas brinda unos pocos cientos de kHz de ancho de banda útil, lo que impide transportar vídeo comprimido con una calidad aceptable. Para superar esta limitación sin sustituir el cableado ni recurrir a módems xDSL/PLC de alto coste, proponemos Tiled-Stream Time Division Multiplexing (TS-TDM), un esquema híbrido que aúna la segmentación espacial del fotograma en tiles con la asignación temporal de cada tile a un slot en Time Division Multiplexing (TDM). Se implementó un prototipo sobre un System on Chip (SoCs) embebidos con codificación H.264/AVC y un búfer cíclico de 25 fps, y se ensayó en un cable bifilar real alimentado por la misma línea. La calidad de la experiencia se midió mediante Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) y Vídeo Multi-Method Assessment Fusion (VMAF). Frente a un flujo H.264 convencional a igual bitrate (380 kbps), TS-TDM logra +2 dB en Peak Signal-to-Noise Ratio - Human Visual System (PSNR-AVG), +33 % en VMAF y reduce la latencia extremo-a-extremo a ≤ 40 ms, sin requerir componentes analógicos adicionales ni filtros paso-alto. Estas mejoras demuestran que la multiplexación temporal inteligente combinada con *tiling* permite llevar vídeo de calidad a infraestructuras heredadas de dos hilos, preservando la interoperabilidad y minimizando el coste de actualización.

***Palabras Clave*- Videoportero, Canal restringido, Tiled Streaming, TDM, Embebidos, Calidad.**

I. Introducción

Los sistemas de videoportero basados en un par bifilar no polarizado sin apantallar constituyen aún hoy la infraestructura dominante en edificios residenciales, pese a su ancho de banda efectivo inferior a unos cientos de kHz y a la atenuación creciente con la distancia. Estas limitaciones convierten la transmisión de vídeo comprimido, por ejemplo H.264/AVC a 25 fps y 720p (HD), en un reto: pequeñas pérdidas de bits o fluctuaciones de retardo rompen la decodificación, y las soluciones de alta capacidad existentes (xDSL, HomePlug/G.hn, IP dedicadas) exigen hardware costoso y elevan la latencia del sistema.

Para sortear estas restricciones proponemos Tiled-Stream Time-Division Multiplexing (TS-TDM) que fusiona la multiplexación por división en el tiempo Time Division Multiplexing (TDM) con la segmentación espacial de vídeo en mosaicos Tiled Streaming (TS): primero se descompone cada fotograma en bloques independientes *tiles* y, después, a cada uno se le adjudica una ranura temporal dentro del ciclo TDM, fija en la modalidad síncrona o asignada dinámicamente en la modalidad asíncrona, de modo que múltiples tiles comparten el mismo canal sin superponerse en el tiempo. Gracias a esta distribución temporal de la carga, TS-TDM permite transportar secuencias de vídeo cuya tasa binaria agregada rebasa el límite teórico del canal sin degradar la calidad percibida, ya que cada *tile* se procesa y transmite de manera autónoma dentro de su *time slot*.

El presente trabajo demuestra experimentalmente que TS-TDM mejora la calidad de vídeo y por lo tanto de forma potencial la calidad de experiencia (QoE) sin incrementar el ancho de banda nominal (380 kbps): se alcanzan ganancias de hasta +50 % en Vídeo Multi-Method Assessment Fusion VMAF y +2 dB en Peak Signal-to-Noise Ratio - Human Visual System (PSNR-AVG) frente al flujo convencional, manteniendo la latencia por debajo de los 40 ms. Para ello se implementa una prueba de concepto en dispositivos embebidos usando GStreamer en C, y se evalúa con un conjunto de 5 métricas objetivas.

Las principales contribuciones de este artículo son:

- Diseño e implementación de TS-TDM en System on Chip (SoCs) comerciales, con buffer cíclico de 25 frames por tile para 25 fps.
- Metodología automatizada de la evaluación de la calidad del flujo multimedia mediante la utilización de diversas métricas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- Comparativa cuantitativa entre tres configuraciones (sin TS-TDM, con TS-TDM y TS-TDM hardwareless ideal) que evidencia la superioridad perceptual de la propuesta incluso bajo codificación H.264/High Profile.

El resto del artículo se organiza como sigue: la sección II revisa el estado del arte; la sección III describe la arquitectura TS-TDM y la metodología experimental; la sección IV presenta y discute los resultados; por último, se expone las conclusiones y futuras líneas de investigación.

II. Estado del Arte

La mayoría de los videoporteros residenciales siguen dependiendo de un par bifilar de 1 mm² sin apantallar no polarizado. La ausencia de blindaje y la necesidad de transportar simultáneamente la alimentación en continua hacen que la atenuación crezca con la frecuencia; el ancho de banda disponible raramente supera unos cientos de kHz, de modo que apenas se admite un flujo analógico de audio y vídeo de muy baja resolución. La propia patente US 9 215 410 B2 [1] destaca que los sistemas clásicos «no permiten múltiples canales de audio/vídeo bidireccionales sobre dos hilos» y carecen de un estándar abierto que lo resuelva.

Cuando se pretende utilizar vídeo H.264/AVC, debido a su alta compatibilidad, la situación se complica: la compresión temporal basada en Group of pictures (GOP) I/P/B exige sincronización precisa y latencia reducida; pérdidas o *jitter* rompen la decodificación y propagan errores de frame a frame[2].

Para sortear este cuello de botella se han investigado tres grandes alternativas:

- xDSL (VDSL2/G.fast) amplía el espectro hasta 30–106 MHz y supera 100 Mbit/s, pero incorpora ecocancelación compleja y añade decenas de milisegundos de retardo [3].
- Power Line Communication (PLC) de banda ancha —HomePlug AV2 y G.hn— alcanza el *gigabit* sobre cable telefónico o eléctrico empleando Orthogonal Frequency Division Multiplexing (OFDM) y Multiple-Input Multiple-Output (MIMO), aunque opera muy por encima del rango electromagnético permisible del par del videoportero y exige filtrado de Electromagnetic Compatibility (EMC) estricto [4], [5].
- Soluciones comerciales "2-Wire HD" (p. ej., el distribuidor Hikvision DS-KAD706-S) combinan señal y alimentación en pasarelas propietarias; reducen obra civil, pero no documentan su gestión de ancho de banda ni publican métricas objetivas de calidad [6].

En cuanto a la multiplexación temporal, matrices profesionales como RTS ADAM-M demuestran la viabilidad de asignar slots TDM a decenas de canales de audio, aunque no afrontan vídeo ni canales ruidosos de banda estrecha [7]

Por otro lado, la comunidad multimedia explota la segmentación espacial (*tiling*) para vídeo 360°. Por ejemplo, C. Ozcinar et al. propone una transmisión adaptativa dependiente del *viewport* [8] y Pang et al. priorizan tiles mediante aprendizaje por refuerzo en VATP360 [9]. Ambas líneas presuponen redes IP de varios Mbit/s y sólo actúan en la capa de aplicación, sin interacción con el medio físico.

Según nuestro conocimiento, ninguno de los trabajos revisados combina la asignación TDM con la división en *tiles* directamente sobre un enlace bifilar < 500 kHz, manteniendo baja complejidad electrónica y evaluando la calidad de experiencia con métricas modernas (PSNR, SSIM, VMAF, etc.).

Nuestra propuesta TS-TDM integra segmentación y multiplexación en el nivel físico, reutilizando el cableado y la modulación existentes, y demuestra mejoras de más del 33 % en métricas objetivas como es VMAF, con latencia extremo-a-extremo ≤ 40 ms, cubriendo así el vacío detectado.

III. Metodología

Este estudio presenta una evaluación comparativa de la calidad visual de flujos de vídeo codificados mediante técnicas de mosaico con multiplexación por división de tiempo Tiled Stream Time Division Multiplexing (TS-TDM) frente a una codificación limitada en el bitrate que soporta el canal.

A. Implementación TS-TDM

La implementación de TS-TDM se ha realizado mediante el uso del framework de GStreamer [10], en el emisor y en el receptor, como se muestra en la Fig. 1, se ha utilizado el lenguaje C para garantizar la implementación en dispositivos embebidos reales.

Para la implementación de TS-TDM, es necesario la elección de un mosaico de corte. Los mosaicos en *Tiled* Stream no estan definidos, sino que se pueden implementar según las necesidades de adaptación del vídeo, en altas resoluciones de 4K, 8K o superiores, el número de *Tiles* será alto con matrices de TS=2x3 o TS=3x3 generando seis y nueve *Tiles* respectivamente, en el caso de estudio la matriz de cortes es de 2 *Tiles*, debido a que la resolución de partida es de 720p y con un corte horizontal de la mitad de la altura total del vídeo es suficiente.

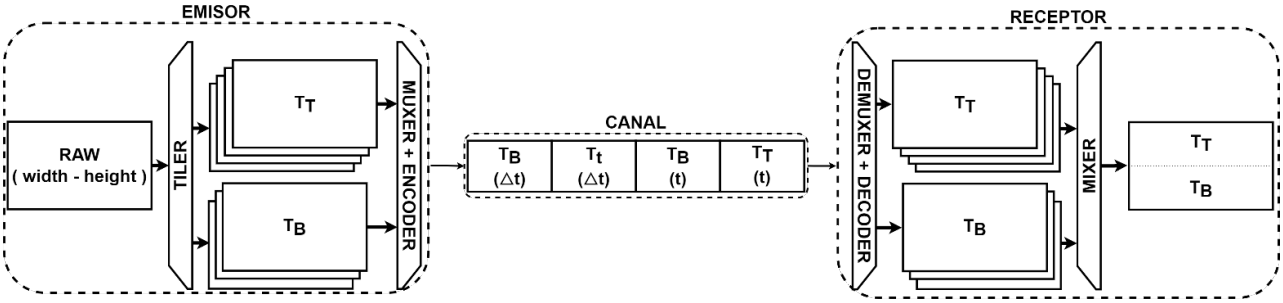


Fig. 1 Esquema implementación TS-TDM

ISBN

GStreamer utiliza diferentes *plugins* que se enlazan para formar una *pipeline*, esta, define el flujo de datos multimedia y realiza una sincronización entre todos los plugins utilizados. Para realizar el desarrollo de TS-TDM, se ha implementado una *pipeline* en el emisor y otra en el receptor. Para el emisor, se accede al contenido del fichero RAW y se define un mosaico de 2x1, Tile-Top (T_T) y Tile-Bottom (T_B). Una vez definido el mosaico, se realiza el corte en dos flujos multimedia independientes sobre el mismo fichero RAW, esto nos garantiza, que el vídeo no sufre ninguna degradación al aplicar TS. Para poder multiplexar los flujos multimedia de T_T y T_B es necesario almacenar 25 *frames* de cada flujo, debido a que el sistema sobre el que realizamos el estudio tiene una tasa de visualización de 25fps, al ser frames completos en formato RAW, es posible modificar el buffer de frames intercalados entre T_T y T_B el estudio del impacto sobre el intercalado de frames, se realizará en un trabajo a futuro.

. Este buffer de vídeo genera un retardo de 0,02 segundos, al tener que almacenar 50 *frames*. Una vez almacenados, se genera el flujo de vídeo, donde los primeros 25 frames van a contener la información de T_T y los siguientes 25 frames la información de T_B , en el siguiente ciclo en $\Delta T+1$, la tercera ventana temporal del TDM contendrá la información de $T_{T\,(\Delta T+1)}$- y la cuarta ventana temporal contendrá la información de $T_{B(\Delta T+1)}$, con esta información, el receptor, puede interpretar en que ventana temporal se encuentra la información necesaria para poder representar sobre un reproductor con matriz de visualización de 2x1 la información de $T_{T-(\Delta T+n)}$ y $T_{B-(\Delta T+n)}$,este proceso es cíclico y se repetirá hasta la finalización de la emisión de vídeo. En la Fig1, podemos ver representado el flujo completo de emisión y recepción mediante un canal de comunicación de un vídeo aplicando TS-TDM.

1. *Data Sets*

El conjunto de data sets utilizados en este trabajo está compuesto por cuatro vídeos de una duración de 6 segundos a 25 fps (150 frames) con diferentes enfoques en su procesado, la duración del vídeo está limitada debido a la capacidad de procesado y memoria del *SoC* embebido del dispositivo físico, el *stream* de audio, es enviado de forma independiente y sincronizado en el reproductor y quedando fuera del estudio presentado.

Los cuatro vídeos intentan cubrir diferentes condiciones de luz encuadres y composición de vídeo identificados como, Terraza Cerca (Luz natural con encuadre cercano), Oficina (Luz artificial con encuadre cercano), Terraza lejos (Luz natural con encuadre lejano) y Noche (Bajas condiciones de luz con encuadre medio), como se muestra en la Fig. 2. Estos cuatro vídeos, son procesados en tres diferentes métodos (Sin TS-TDM, Con TS-TDM y Con TS-TDM Hardwareless). Los videos procesados, parten de una fuente común de generación de vídeo, está fuente es la cámara de vídeo que dispone el producto comercial AMP. VID. DUOX PLUS SKYLINE CONNECT WV[11] para la obtención del vídeo se han realizado modificaciones en el *firmware* (FW) del producto, las cuales permiten la extracción del vídeo directamente del bus de datos de la cámara en formato RAW con los parámetros de origen (720p, NV12 y 25fps) como se muestra en la Fig. 2., con esto se garantiza la fidelidad de las pruebas al realizar las manipulaciones del vídeo y la aplicación de TS-TDM sobre el vídeo RAW sin ningún tipo de procesado previo o posterior, conservando toda la información de origen.

Fig. 2 Muestra de los cuatro vídeos que componen el dataset.

2. *Sin TS-TDM*

El vídeo obtenido en esta etapa será aquel que dispondrá de la calidad de referencia que actualmente tiene el sistema bajo ensayo y con el que compararemos los resultados de las métricas de calidad de vídeo, para valorar el impacto de aplicar TS-TDM. Para la obtención del vídeo procesado por el emisor del vídeo, se ha realizado una modificación en el FW del producto, la cual, antes de realizar la transmisión por el canal, se almacena en memoria volátil del sistema en un fichero en formato h.264, como se muestra en la Fig 3.

El vídeo capturado por la cámara, es procesado por el enconder aplicando los parámetros de compresión Ancho: 720, Alto: 480, Fps: 25, *h264Profile*: *HIGH*, h264Level: *LEVEL*_51, *keyFrameInterval*: 25, Bitrate: 380Kbps. A este vídeo lo identificaremos como Sin TS-TDM.

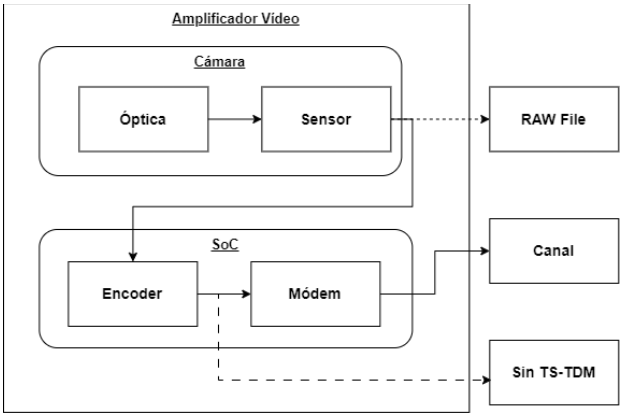


Fig. 3 Obtención de fichero RAW y vídeo Sin TS-TDM

3. *Con TS-TDM*

Utilizando como fuente de vídeo el fichero RAW generado según se muestra en la Fig 4. Aplicaremos sobre el fichero RAW, la estrategia de TS-TDM y obtendremos como

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

resultado un vídeo en formato RAW-TSD-TDM, este fichero se pasa por el *Encoder* que aplica el mismo procesado que en el vídeo Sin TS-TDM, el resultado es un vídeo con las mismas condiciones de contorno, pero aplicando TS-TDM al que identificaremos como Con TS-TDM.

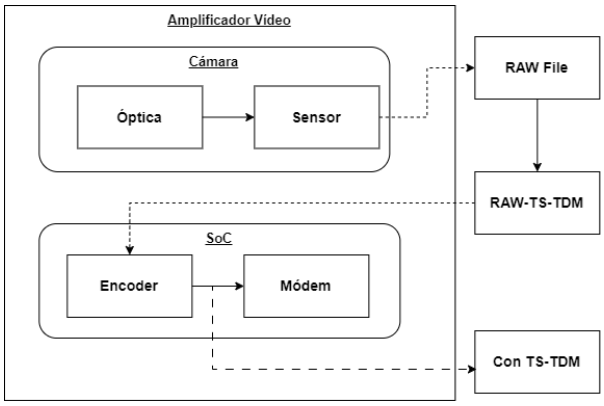


Fig. 4: Obtención vídeo Con TS-TDM

4. *Con TS-TDM Hardwareless*

Debido a las limitaciones y prestaciones de procesado que existen en el System on Chip (SoCs) bajo ensayo, se ha generado un vídeo mediante el uso FFMPEG[12], con la misma configuración del *Encoder* que garantiza un vídeo con las mismas características en términos de ancho de banda, este vídeo ha sido procesado posteriormente con TS-TDM, de esta forma podemos considerar el enconder como ideal, a este vídeo lo identificaremos como Con TS-TDM Hardwareless.

B. *Evaluación*

Se realizará una evaluación objetiva de la calidad del vídeo transmitido a través de un canal físico confiable, sin pérdidas perceptibles durante su uso y con un ancho de banda limitado a 380 kb/s. Esta restricción se debe a condicionantes del sistema, ya que la capacidad disponible para la transmisión de vídeo se encuentra en ese rango. La evaluación se llevará a cabo sobre vídeo en resolución 720p, sin considerar la carga computacional, y utilizando técnicas de TS-TDM. La implementación de TS-TDM ha sido realizada sobre un SoC real en la generación de los *datasets*, basándose en un producto comercial [11] para garantizar la aplicación a futuros en la industria y obtener resultados lo más cercanos a la implementación física real, en el receptor se ha optado por una implementación simulada para no añadir distorsiones, debidas a limitaciones de Hardware.

La evaluación objetiva se realiza con el uso de 5 métricas, es conocido que centrar el análisis a los resultado en una única métrica puede no ser relevante o concluyente como se expone en [13], se utilizarán técnicas *Full Reference*, las cuales tienen como resultado un valor o valores que miden la diferencia entre el vídeo de origen (referencia) y el vídeo transmitido (distorsionado), las métricas utilizadas son Vídeo Multi-Method Assessment Fusion (VMAF)[14], Peak Signal-to-Noise Ratio (PSNR) [15], Peak Signal-to-Noise Ratio - Human Visual System [16], Structural Similarity Index Measure (SSIM)[17][18], Multiscale Structural Similarity Index (MS-SSIM)[19] para evaluar los distintos vídeos.

Los resultados se obtienen a partir de experimentos utilizando la plataforma de pruebas Video Quality Metrics Toolkit (VQMTK)[20], está herramienta ofrece diferentes métricas de análisis de vídeo integradas en una imagen *Docker*, lo que agiliza su uso, además se han desarrollado diferentes *Scripts* en *Python* que dan soporte a cálculos adicionales.

IV. Resultados

En este trabajo se han evaluado los vídeos contenidos en el *dataset* presentado en la sección anterior del documento, con el objetivo de obtener las métricas de calidad de vídeo *Full-Reference* de forma individual, a modo de resumen, unificamos los resultados registrados en la Tabla I, donde se compara el resultado de las diferentes métricas al ser aplicadas a los tres vídeos que conforman el dataset vídeo Sin TS-TDM, vídeo Con TS-TDM y vídeo TS-TDM Hardwareless.

Como se puede observar en todas las métricas analizadas la solución presentada en este trabajo aporta mejoras (tonos verdes) tanto en la propuesta implementada (TS-TDM) como en el escenario ideal (TS-TDM Hardwareless). En conjunto, los resultados confirman que TS-TDM eleva la calidad percibida sin aumentar el *bitrate* ni la complejidad del códec, gracias a la asignación temporal de slots y a la segmentación en tiles. Además, el escenario *hardware-less* revela que optimizaciones futuras en la capa analógica/FEC podrían acercar el rendimiento real al límite teórico, potenciando aún más la viabilidad de vídeo de alta calidad sobre el par bifilar legado.

Tabla I
Resumen de resultados

Métricas	Vídeo	Sin TS-TDM	TS-TDM	TS-TDM Hardwareless
VMAF (0-100)	1	45,54	48,89	68,53
	2	28,686	36,701	61,148
	3	30,27	37,17	56,00
	4	44,69	54,58	64,37
PSNR AVG (dB)	1	32,02	32,17	34,23
	2	26,93	30,74	33,50
	3	29,95	22,59	31,93
	4	23,62	33,71	34,18
PSNR-HVS (dB)	1	33,40	39,63	41,13
	2	23,44	32,05	34,78
	3	21,19	30,46	32,56
	4	18,24	30,02	30,69
SSIM (%)	1	18,04	19,65	21,48
	2	13,03	16,87	18,93
	3	11,90	15,20	18,22
	4	10,65	13,68	15,82
MS-SSIM	1	0,930	0,940	0,944
	2	0,871	0,936	0,955
	3	0,816	0,906	0,930
	4	0,810	0,926	0,931

1. *VMAF*

Vídeo Multi-Method Assessment Fusion (VMAF) [14], es una métrica utilizada para evaluar la calidad de los vídeos. Combina varios métodos de evaluación de calidad de vídeo para proporcionar una puntuación que refleja la percepción humana de la calidad del vídeo sobre pantallas 1080p (FHD), lo cual será una métrica muy existente para nuestro experimento ya que nuestro video parte de una resolución 720p (HD). La comparación de las métricas VMAF obtenidas la podemos ver en detalle en la Tabla II, donde el índice pasa de 45,5 a 48,9 puntos; una ganancia de +3,3 pts (~7 %) en el vídeo 1, de 28,68 a 36,70 puntos, una ganancia de +8,2 pts

ISBN

(~27 %) en el vídeo 2, de 30,27 a 37,17 puntos, una ganancia de +6,9 pts (~23 %) en el vídeo 3 y de 44,69 a 54,58 puntos, una ganancia de 9,89 pts (~22 %). En todos los resultados es suficiente para producir una *just noticeable difference* (JND) percibida por el usuario final. El techo marcado por el caso hardware-less (68,5 puntos) confirma que el canal físico junto a las limitaciones del SoC son hoy la principal fuente de degradación y deja un margen de mejora significativo para la electrónica de línea.

La mediana del vídeo con TS-TDM Hardwareless es superior en un incremento del 50,49%, 113,16%, 85% y 44,04% para los vídeos 1, 2, 3, y 4 respectivamente en comparación a los vídeos de referencia, lo que hace destacar que son los más nítidos seguidos de los vídeos con TS-TDM. La proximidad de la $\bar{x}$ y Mediana en los cuatros vídeos y sus procesados indica que no hay valores atípicos fuertes en las tres muestras. El procesado con TS-TDM Hardwareless es claramente el de mayor calidad, tanto en promedio como en valores individuales.

Tabla II
RESUMEN DE RESULTADOS VMAF

Métrica VMAF	Vídeo	Sin TS-TDM	TS-TDM	TS-TDM Hardwareless
$\bar{x}$	1	45,54	48,89	68,53
	2	28,686	36,701	61,148
	3	30,27	37,17	56,00
	4	44,69	54,58	64,37
Mediana	1	45,8020	49,0296	69,0588
	2	30,258	36,98	60,96
	3	30,89	37,86	56,59
	4	44,28	54,93	65,42
σ	1	0,9611	0,8902	1,0734
	2	4,768	3,52	1,36
	3	1,43	2,16	1,87
	4	1,63	2,02	2,48

2. *PSNR*

Peak Signal-to-Noise Ratio (PSNR) [15], evalúa la calidad de un vídeo comprimido en comparación del vídeo original, con esta comparación se puede determinar la degradación que sufre un vídeo al ser comprimido y transmitido por el canal.

Para ello se evalúan los componentes de luminancia (Y) y crominancia (U y V) por separado, y luego se obtiene un promedio para determinar la calidad general del vídeo. La comparación entre los cuatro vídeos y entre los tres procesados, sin TS-TDM (A), con TS-TDM (B) y TS-TDM Hardwareless (C) la podemos encontrar en la Tabla III. En *PSNR* el procesado (C) supera en Y en más de +2dB, +2,96dB, +9dB y +1,14dB los vídeos 1, 2, 3, y 4 respectivamente, a los vídeos procesados con (A) y (B), lo que indica que sombras y contornos más nítidos al ojo humano.

En U y V tiene una mejora respecto a el procesado (A) y (B), lo que indica colores más fieles al original. Por último, el promedio de los vídeos procesados en (C) 34,22 dB, es de +2,21 dB, en vídeo 1, 33,50dB, es de +6,57 dB, en el vídeo 2, 31,95, es de +1,98 en el vídeo 3, 34,18dB, es de +10,56 dB en el vídeo 4, respecto a el procesado (A) lo cual indica una mejora que es superior a los +0,5 dB con una media de mejora de +5,33dB, que sería un umbral perceptible por el ojo humano.

Aunque el incremento absoluto es modesto, el nivel supera en ambos todos los casos los 32 dB, considerado aceptable para SD a tasa moderada. El salto limitado sugiere que la ventaja de TS-TDM no reside tanto en reducir el error medio global como en redistribuirlo de forma que resulte menos molesto, algo que PSNR no refleja con fidelidad.

En resumen, el procesado (C) ofrece unos mejores resultados en comparación a los procesados (A) y (B) ofreciendo una imagen más estable, con bordes definidos y menor grano en los fondos, transiciones suaves, ausencia de macrobloques visibles y nitidez constante.

Tabla III
RESUMEN DE RESULTADOS PSNR

Métrica PSNR	Vídeo	Y	U	V	AVG
Sin TS-TDM	1	30,59	37,85	39,56	32,02
	2	25,43	33,84	35,00	26,93
	3	28,06	35,41	35,97	29,95
	4	22,01	34,92	32,38	23,62
TS-TDM	1	30,79	37,55	39,02	32,17
	2	29,55	34,04	35,85	30,74
	3	20,94	33,75	34,00	22,59
	4	33,62	35,52	32,71	33,71
TS-TDM Hardwareless	1	32,88	39,63	40,26	34,23
	2	32,51	36,52	36,99	33,50
	3	30,76	36,13	36,10	31,93
	4	34,75	34,51	32,17	34,18

3. *PSNR-HVS*

Peak Signal-to-Noise Ratio - Human Visual System (PSNR-HVS)[16], es una mejora respecto a PSNR que tiene en cuenta las características del sistema visual humano (HVS) para proporcionar una evaluación más precisa de la calidad de las imágenes y vídeos. La comparación entre los tres procesados, sin TS-TDM (A), con TS-TDM (B) y TS-TDM Hardwareless (C) la podemos encontrar en la Tabla IV.

Tabla IV
RESUMEN DE RESULTADOS PSNR-HVS

Métrica PSNR-HVS	Vídeo	Y	U	V	AVG
Sin TS-TDM	1	35,75	40,71	41,46	33,40
	2	25,37	36,36	36,37	23,44
	3	20,02	35,10	34,58	21,19
	4	20,21	33,50	28,56	18,24
TS-TDM	1	36,58	40,90	41,37	39,63
	2	34,58	38,40	39,13	32,05
	3	32,80	38,18	38,38	30,46
	4	36,86	33,65	28,58	30,02
TS-TDM Hardwareless	1	36,97	43,54	42,88	41,13
	2	37,44	40,07	40,86	34,78
	3	35,02	40,03	39,21	32,56
	4	40,24	33,16	28,48	30,69

El procesado (C) alcanza el valor más alto en los cuatro vídeos siendo de +7,73dB, +11,34dB, +11,37db y +12,45dB en los vídeos 1, 2, 3, y 4 respectivamente mostrando un equilibrio óptimo entre nitidez de luminancia y fidelidad cromática según la métrica de PSNR-HVS. Los resultados reflejan

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

incrementos del (40,93%) de media entre el procesado (A) y (B) y del (48,93%) de media entre el procesado (A) y (C) en los cuatros vídeos., destacando beneficios perceptuales. El Impacto visual del procesado (C), tiene una mejor calidad perceptual, bordes definidos con color y tiene una probabilidad baja de artefactos. El procesado (B): Muy buena calidad (casi en el nivel del procesado (C)) pero con crominancia ligeramente inferior. El procesado (A): resultados óptimos, pero inferiores en comparación con el procesado (B) y el procesado (C).

PSNR-HVS, muestra valores más altos que PSNR, en los procesados (B) y (C) donde aparecen mejoras perceptuales respecto al procesado (A).

4. SSIM

Structural Similarity Index Measure (SSIM) [16] [18] compara estructuras en ventanas locales (brillo, contraste y correlación) y va de 0 (sin similitud) a 100 (idéntico). Valores más altos implican mayor preservación de la percepción estructural. A diferencia del PSNR, que se basa en diferencias absolutas de píxeles, SSIM evalúa la similitud estructural, esto hace que la métrica se parezca más a la percepción visual humana. La comparación de los vídeos entre los tres procesados, sin TS-TDM (A), con TS-TDM (B) y TS-TDM Hardwareless (C) la podemos encontrar en la Tabla V. El resultado del análisis con la métrica SSIM, refleja un aumento de media del (+23,64 %) con el procesado (B) respecto al procesado (A) y un aumento de media del (+41,51 %) con el vídeo (C) respecto al procesado del vídeo (A).

Tabla V
RESUMEN DE RESULTADOS SSIM

Métrica SSIM	Vídeos	Y	U	V	AVG
Sin TS-TDM	1	20,07	15,25	16,00	18,04
	2	14,13	11,19	11,70	13,03
	3	11,05	14,44	14,16	11,90
	4	10,08	12,93	11,42	10,65
TS-TDM	1	20,69	17,76	18,49	19,65
	2	18,37	14,60	15,12	16,87
	3	17,47	12,66	12,61	15,20
	4	20,11	10,35	9,00	13,68
TS-TDM Hardwareless	1	22,68	19,85	19,69	21,48
	2	21,32	16,33	16,17	18,93
	3	19,80	16,41	15,93	18,22
	4	21,56	12,64	11,25	15,82

5. MS-SSIM

En cambio, Multiscale Structural Similarity Index (MS-SSIM) [19]. Se basa en el supuesto de que el sistema visual humano está muy adaptado para extraer información estructural de la escena y, por tanto, una medida de la similitud estructural puede proporcionar una buena aproximación a la calidad de imagen percibida. Un valor cercano a 1 indica mejor calidad de imagen y un valor cercano a 0 indica peor calidad.

La comparación de los cuatro vídeos entre los procesados de vídeo, sin TS-TDM (A), con TS-TDM (B) y TS-TDM Hardwareless (C) la podemos encontrar en la Tabla VI. MS-SSIM mantiene un resultado positivo de (+8,37 %) con el procesado (B) respecto al procesado (A) y (+14,25 %) con el procesado (C) respecto al procesado (A), dando como resultado mejor preservación de detalles.

El vídeo con TS-TDM Hardwareless presenta la mejor calidad global según MS-SSIM, con la media y media armónica más altas y el mayor valor máximo. El vídeo con TS-TDM es una mejora clara frente al vídeo sin TS-TDM. A pesar de tener el valor mínimo más bajo, el vídeo con canal simulado muestra una calidad superior de manera consistente.

Tabla VI
RESUMEN DE RESULTADOS MS-SSIM

Métrica MS-SSIM	Vídeos	Sin TS-TDM	TS-TDM	TS-TDM Hardwareless
Valor Mínimo	1	0,8163	0,8413	0,8288
	2	0,5881	0,5890	0,9429
	3	0,6544	0,6676	0,9134
	4	0,7353	0,8019	0,9070
Valor Máximo	1	0,9474	0,9550	0,9671
	2	0,9460	0,9539	0,9684
	3	0,8769	0,9351	0,9516
	4	0,9129	0,9356	0,9501
$\overline{x}$	1	0,9313	0,9404	0,9454
	2	0,8737	0,9377	0,9559
	3	0,8176	0,9069	0,9307
	4	0,8113	0,9269	0,9315
Media armónica	1	0,9309	0,9401	0,9449
	2	0,8711	0,9366	0,9559
	3	0,8160	0,9065	0,9307
	4	0,8107	0,9268	0,9315

V. CONCLUSIONES Y TRABAJOS FUTUROS

Los resultados obtenidos demuestran que la solución propuesta TS-TDM constituye una solución viable y de bajo coste para transmitir vídeo H.264 sobre el par bifilar sin apantallar típico de los sistemas de videoportero. Con la configuración evaluada (720p, 25fps y 380kbps) se observaron mejoras consistentes respecto al flujo de referencia:

- VMAF aumentó de 45,5 a 48,9 puntos (+7 %), de 28,68 a 36,70 puntos (+26%), de 30,27 a 37,17 (+22%) y de 44,69 a 54,58 (+22%) en los vídeos que componen el dataset superando el umbral de *just noticeable difference* y evidenciando una ganancia perceptual clara.
- PSNR-HVS registró una subida de +10,72 dB de media en los vídeos del dataset, señal de que el esquema preserva con mayor fidelidad los contornos y las texturas finas, justo donde la sensibilidad del ojo humano es mayor.
- SSIM creció un 23,64 %, mientras que la latencia extremo-a-extremo se mantuvo por debajo de 40 ms, compatible con la interacción en tiempo real propia de un videoportero.

La propuesta "TS-TDM Hardwareless", que elimina las imperfecciones del front-end analógico, elevó VMAF hasta 68,5 y PSNR hasta 34 dB. Ello revela que el margen de mejora restante depende esencialmente de la electrónica de línea y no de los principios de multiplexación propuestos. De aplicarse TS-TDM en los videoporteros actuales, tendría como resultado una mejora la calidad de vídeo sin realizar cambios en las instalaciones actuales ya que se mantendría tanto el canal de transmisión como el videoportero instalado. Una mejora calidad de video, en un sistema critico de seguridad como el videoportero, siendo este, el acceso principal a nuestros hogares repercute en una mejora en QoE por parte del

ISBN

usuario. Teniendo en cuenta el actual parque instalado de videoporteros que implementan el sistema bajo estudio, el impacto de la aplicación de TS-TDM tendría un gran alcance al ser aproximadamente más del 50% del parque instalado actual en España.

Como trabajos futuros se quiere realizar más pruebas en instalaciones in-situ, para ver los efectos que aportan las diferentes longitudes de cable, la diafonía entre pares, el ruido impulsivo y las variaciones térmicas a la solución propuesta. Por último, para complementar las métricas objetivas, se llevarán a cabo ensayos subjetivos con usuarios (métodos ACR y DSCQS) que validen la mejora percibida en escenarios reales de uso y ayuden a calibrar los umbrales de calidad alcanzados por TS-TDM.

VI. Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el programa de Doctorandos y doctorandas empresariales (Innodocto) 2023 dentro de la convocatoria de Promoción del talento: Agentes de innovación, incorporación de investigadores y tecnólogos para proyectos de innovación en empresas y doctorandos empresariales de la Comunidad Valenciana y en colaboración con la empresa FERMAX TECHNOLOGIES SL.

Referencias

[1] C. F. Zaera, E. S. Peris, V. G. Millan, and J. S. Gimeno, "TWO-WIRE MULTICHANNEL VIDEO DOOR SYSTEM," 2015. [Online]. Available: https://patents.google.com/patent/US9215410B2/e

[2] "H.264 'zero' latency video encoding and decoding for time-critical applications - EE Times." Accessed: May 27, 2025. [Online]. Available: https://www.eetimes.com/h-264-zero-latency-video-encoding-and-decoding-for-time-critical-applications/

[3] "G.993.2 : Very high speed digital subscriber line transceivers 2 (VDSL2)." Accessed: May 27, 2025. [Online]. Available: https://www.itu.int/rec/T-REC-G.993.2

[4] L. Yonge *et al.*, "An overview of the homeplug AV2 technology," 2013. doi: 10.1155/2013/892628.

[5] "G.9960 : Unified high-speed wireline-based home networking transceivers - System architecture and physical layer specification." Accessed: May 27, 2025. [Online]. Available: https://www.itu.int/rec/T-REC-G.9960

[6] "DS-KAD706-S - Hikvision Europe." Accessed: May 27, 2025. [Online]. Available: https://www.hikvision.com/europe/products/Video-Intercom-Products/2-wire-Series/Pro-Series/DS-KAD706-S/

[7] "ADAM-M | RTS Intercoms." Accessed: May 27, 2025. [Online]. Available: https://products.rtsintercoms.com/na/en/adam-m/

[8] C. Ozcinar, A. De Abreu, and A. Smolic, "Viewport-aware adaptive 360° video streaming using tiles for virtual reality," in *Proceedings - International Conference on Image Processing, ICIP*, 2017. doi: 10.1109/ICIP.2017.8296667.

[9] Z. Pang, "VATP360: Viewport Adaptive 360-Degree Video Streaming based on Tile Priority," Jul. 2023, Accessed: May 27, 2025. [Online]. Available: https://arxiv.org/pdf/2307.15984

[10] "GStreamer: open source multimedia framework." Accessed: May 18, 2025. [Online]. Available: https://gstreamer.freedesktop.org/

[11] FERMAX, "AMP. VID. DUOX PLUS SKYLINE CONNECT." Accessed: May 15, 2025. [Online]. Available: https://www.fermax.com/spain/productos/videoporteros/placas/placa-skyline/f73910-amp-vid-duox-plus-skyline-connect-wv/

[12] "FFmpeg." Accessed: May 18, 2025. [Online]. Available: https://ffmpeg.org/

[13] M. Shahid, A. K. Singam, A. Rossholm, and B. Lövström, "Subjective quality assessment of H.264/AVC encoded low resolution videos," *2012 5th International Congress on Image and Signal Processing, CISP 2012*, pp. 63–67, 2012, doi: 10.1109/CISP.2012.6470021.

[14] C. G. Bampis, Z. Li, and A. C. Bovik, "Spatiotemporal Feature Integration and Model Fusion for Full Reference Video Quality Assessment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, 2019, doi: 10.1109/TCSVT.2018.2868262.

[15] W. contributors, "Peak signal-to-noise ratio," 2025.

[16] "(PDF) A NEW FULL-REFERENCE QUALITY METRICS BASED ON HVS." Accessed: May 27, 2025. [Online]. Available: https://www.researchgate.net/publication/251229783_A_NEW_FULL-REFERENCE_QUALITY_METRICS_BASED_ON_HVS

[17] Z. Wang and A. C. Bovik, "Mean squared error: Lot it or leave it? A new look at signal fidelity measures," *IEEE Signal Process Mag*, vol. 26, no. 1, pp. 98–117, 2009, doi: 10.1109/MSP.2008.930649.

[18] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1398–1402, 2003, doi: 10.1109/ACSSC.2003.1292216.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[20] "Description of Video Quality Metric (VQM) Software - ITS." Accessed: May 16, 2025. [Online]. Available: https://tinyurl.com/dzxv945

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Proposal of Architecture for Extended Reality Network Functions in Digital Twin Networks

Xavier Calle-Heredia, Xavier Hesselbach
Departament of Networking Engineering
Universitat Politècnica de Catalunya
Jordi Girona 1-3, E-08034.
xavier.calle@upc.edu, xavier.hesselbach@upc.edu

The emergence of Extended Reality (XR) has provided the path towards the deployment of novel networking applications that enhances both the way users interact in digital environments and the network experience. On the other hand, Digital Twin Networks (DTNs) are envisioned as a promising paradigm to revolutionize the conventional lifecycle management of networks by means of real-time simulations, network prediction, resilience, orchestration, among others. The integration of previously mentioned technologies enable the deployment of novel and immersive network functions operating within a Digital Twin Network environment; the proposed synergy is envisioned to offer promising solutions to be exploited in multiple use cases including: networking, telemedicine, space exploration, and autonomous driving. To support this integration, an architecture must be provided with the aim of offering a robust infrastructure able to cover cost demands for XR services while offloading processing capabilities from client devices such as smartphones, laptops, IoT devices, etc.

***Keywords*—Extended Reality Network Functions; Digital Twin Networks; Network Virtualization, Architectural Entity**

I. INTRODUCTION

This research focuses on developing a robust and reliable architectural framework that seamlessly integrates the Digital Twin Network (DTN) paradigm and the Extended Reality (XR) technology. In this work, the term DTN is used as defined by ITU-T in [1], although Network Digital Twin (NDT) is emerging in the community (including IETF) to avoid confusion with Delay/Disruption Tolerant Networking). The concept of Extended Reality refers to any type of immersive technology that is able to merge or alter our perception of the real world by integrating digital content. This includes Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR). In Virtual Reality, virtual environments are created to provide an immersive experience. The Augmented Reality provides digital content overlayed onto the deployed scenario. Finally, Mixed Reality allows for the interaction of AR with the VR substrate. The latest technological advancements have introduced digital environments, such as the metaverse, as scenarios in which users can interact and collaborate in a more intuitive and satisfying way. While DTNs can be exploited to optimize network resources, enhance network management, provide real-time monitoring, real-time network simulations, among others; XR network services can be deployed to provide immersive, intuitive, and realistic network functions, then optimizing the Quality of Experience (QoE). Therefore, the proposed framework, denominated Digital Twin Network plus Extended Reality Functions (DTN+XRF), is envisioned to open a wide range of relevant use cases including telemedicine, industry 4.0, space exploration, in which, a replica of the Original Network (ON) is created to enable users to immerse themselves in the application, manipulating it in real time, while resources are optimized.

To cover the deployment of DTN+XRF, the architectural framework is the cornerstone. The architecture has to be robust to support expensive processing processes related to DTN-ON synchronization, XR services creation, XR demands, etc; resilient to overcome failures and maintain a continuous operation; and lightweight from the user perspective to enable devices that typically lack of huge processing units to use the services offered by the DTN+XRF.

The remainder of this document is organized as follows: section II describes the background of this research, section III introduces the proposed architecture. While the Section IV describes briefly some applications for the proposed system. Then in section V some challenges for a real deployment of the proposed system are addressed, and finally section VI presents the conclusions of this research.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

II. BACKGROUND

For realizing the DTN+XRF, three primary technologies have been identified: Digital Twins (DTs), Extended Reality (XR) and Network Virtualization (NV).

Digital Twin Networks [2] defined as a replica of an existing network approached with Digital Twin (DT) technology, is an emerging paradigm introduced to enhance bidirectional interaction and communication between physical and virtual objects, then binding the gap between both domains. In [3], the role of DTNs is analyzed in optimizing network life cycle management, real-time monitoring, data visualization and manipulation, etc. The paper [4] highlights the ability of DTN to simulate different scenarios in order to predict future behaviors and then optimize processes in the ON such as routing, management, etc. The article [5] uses a slightly different term: Network Digital Twin (NDT) and proposes its construction based on a digital mapping system that exploits YANG models to describe the structure, status and organization of network elements, then those models are standardized based on Labeled Property Graphs, finally uses an API REST that exposes the network information from the NDT in order to make it available for clients that will interact with the NDT elements.

In [6], XR technology comprised of Augmented Reality (AR), which adds virtual elements to the physical world; Virtual Reality (VR), which creates simulated scenarios; and Mixed Reality (MR), which combine AR and VR; is analyzed along with challenges for covering XR applications including high data transmission speed, low latency, reliability. The paper [7] discusses the capability of XR to create immersive and intelligent scenarios with physical and virtual agents that enable the manipulation of physical and virtual objects in environments such as the metaverse.

The article [8] describes the role of two NV technologies: Network Function Virtualization (NFV) and Software Defined Network (SDN) in the deployment of flexible, programmable, interoperable, dynamic, and easy orchestration networks. Although the paper [9] describes the role of Network Slicing (NS) in the development of tailored networks that satisfy specific requirements and needs. The paper [10] proposes a holistic virtualization of networks approached with DT technology to minimize the latency of offered services; while it introduces pervasive intelligence to optimize the deployment of services, resource utilization and orchestration, and enhanced simulations.

In [11] the integration of DT and XR to collaborative environments is addressed in which users can interact with immersive gestures. In [12], we describe how the synergy of XR and DTN is exploited to improve user interfaces for novel smart services. In [13] the utilization of DT and XR is used to enhance the programming of industrial robots for industry 4.0.

III. PROPOSED ARCHITECTURE

In this section, a first approach of the architectural framework to cover the deployment of the DTN + XRF system is presented.

Taking the existing Extended Reality definition to communication networks, the concept of Extended Reality Network Functions is defined as network functions based on the properties of Extended Reality.

The proposed system is intended to integrate a DTN and an XRF module; then, the system has to perform multiple tasks including DTN deployment, DTN-ON synchronization, DTN monitoring, XR service deployment, XR service maintenance, and so forth. Therefore, the architectural framework must be conceived with enough robustness to guarantee the achievement of all the desired tasks; on the other hand, the architecture should not increase in complexity to facilitate system deployment and management. Additionally, the DTN+XRF can be realized as a system composed of two domains: the Original Network (ON) domain and the Digital Twin Network (DTN) domain. The tasks to be performed in the ON domain are as follows. ON information collecting, ON modification event handler, ON information transmission. On the other hand, the DTN domain can be subdivided into DTN tasks that include DTN modeling, deployment, synchronization, simulation, monitoring, self-maintenance (autonomous); and the XR tasks including DTN model processing, XR service deployment, XR service maintenance.

Then, the architectural framework is composed of three architectural entities (AEs) in charge of overseeing the whole system operation; such AEs are the Original Network Manager (ONM), the Digital Twin Manager (DTM), and the Extended Reality Manager (XRM). AEs are presented in Figure 1

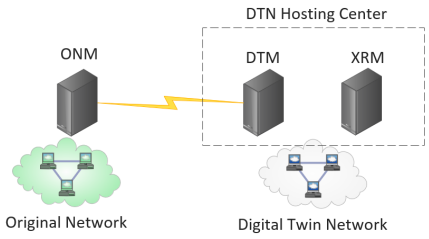


Fig. 1. Proposed architectural entities.

In the previous figure, note that DTM and XRM are supported by a DTN hosting center that can be conceived as a robust data center with enough processing, storage, and networking capabilities to cover the operations of DTN+XRF. On the other hand, the ONM is hosted by the existing ON infrastructure.

Since the objective is to maintain the simplicity of the AEs, the AEs must be conceived in a modular approach. Every module requires to perform specific tasks. The proposed DTN+XRF, which consists of an instance of an existing network, can be deployed anywhere, such as research facilities or even space stations; then, it has been identified that a critical functionality of AEs is communication between the ON and the DTN to create a DTN or synchronize it. Figure 2 depicts an initial approach to the modularity of AEs; notice that the initial modules in both DTM and ONM are Edge Routers that are in charge

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

of the DTN-ON information transmission. It is important to emphasize that in future research, the complete modular architecture of AEs, as well as the data flow between every module, will be analyzed in more detail.

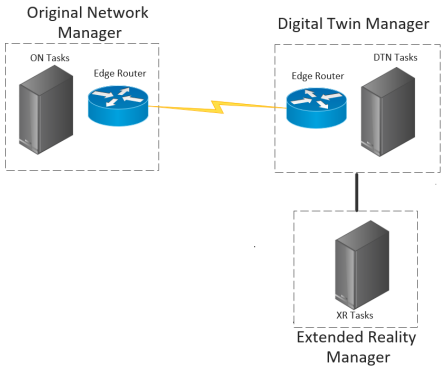


Fig. 2. AEs initial modular definition.

Implementing these modules requires high computing capacity, ultra-low latency, efficient virtualized resource management, and massive connectivity. While 5G networks can only support low-complexity scenarios, 6G enablers are expected to overcome these limitations. These include advanced virtualization technologies and high-capacity data centers that enable functionalities such as SDN, NFV, cloud-native functions, AI, and OSM, thereby enhancing network softwarization and programmability. Enabling envisioned technologies for Digital Twins are deeply discussed in [14] and [15]. For DTN-XRF, the study of the specific framework is part of future work.

IV. Use Cases

This section introduces potential applications of DTN+XRF. The architecture has to be robust and flexible to guarantee continuous operation and to be able to adapt to any type of application.

A. Use Case 1: Base space

This is a novel application, the Original Network (ON) from Earth is modeled and replicated as a DTN, and deployed in a new established space base (Moon, Mars or satellites). The original network can be the one owned by an space agency, or the network in devices (rockets, satellites, space vehicles, appliances, life system support, etc.). The DTN+XRF architecture offers the possibility to extend the ON network with new services addressed with XR to manipulate rovers and all kinds of robotic systems. For this use case, the communication latency, to communicate the ON and the space base to create/synchronize the DTN, might be in the range from seconds to minutes considering the large communication distances and space channel conditions. Therefore, robust architectures are required to support the challenges in communication, including local content repositories and proxies to maintain services even with synchronism being lost.

B. Use Case 2: Telemedicine and telesurgery

DTN networks provide a revolutionary strategy to address remote surgery and health care procedures. The DTN+XRF system allows the interaction in an immersive and intuitive manner with the Digital Twin of a surgery robot or health care system and applies enhanced security and control operations, taking advantage of local data centers that enhance the computing capabilities, to guarantee patient safety.

C. Use Case 3: Nano-networks

Nano-networks are networks of nano-devices. Nano-devices are basically nano-machines, in some cases made of cells (hybrid nano-devices). Main issues in nano-networks include limited computational capacity,communication, mobility, location and energy consumption (energy harvesting is a relevant issue). Considering accessibility, control operation, motorization, etc. the access to the nano-network devices is challenging making difficult the operation. Creating instances of nano-networks with the proposed DTN+XRF architecture can help to statistically access the swarm of nano-devices, with real-time monitoring and able to activate functionalities in an immersive metric visualization.

V. Challenges

In this section, challenges are addressed to be able to deploy an initial proof-of-concept experimental testbed of the proposed Extended Reality Network Functions.

- Computing resources: To support DTN and XRF processes, extremely cutting-edge performance will be required to handle DTN models. Among the computing resources to deploy the system are included: pure bare metal servers in order to deploy a Digital Twin Network with physical hardware; in a different approach, virtualized resources are more suitable for deploying the DTN+XRF and are composed of: Virtual Machines (Vms) that provide isolation and complete stack of resources for covering complex tasks for constructing different modules of the AEs (DTM, XRM and ONM); on the other hand are Containers, which are lighter and suitable for deploying micro services and complementing functionalities of VMs; along with the virtualized resources, an interesting approach can be the orchestration of such virtualized resources with structures such as Kubernetes that takes care of resources scheduling, automated deployment tasks, load balancing, etc.
- Networking resources: The communication channel between the ON and the DTN is critical to ensure robust networking resources; according to research [16] the data rate can be about 2.3 Gbps. This estimation can help to understand the nature of the requirements.
- DTN modeling: Efficient DTN deployment requires robust modeling techniques. Approaches such as adjacency matrices or adjacency lists can be applied,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

but context-aware and resilient structures are needed to meet specific use cases.

- XR services construction: Building XR network services in digital environments is still an open challenge. Candidate technologies include OpenStack, Open Source MANO (OSM), and Network Function Virtualization (NFV), among others.
- Transmission delay: The DTN+XRF is able to be deployed everywhere, even far from the ON in terms of delay. Transmission latency can be significant, as in space station scenarios. To address this, the interface between DTM and ONM edge routers should use adaptive transmission protocols, buffering and caching to handle disruptions, compression for efficient data transfer, and Delay Tolerant Networking (DTN) techniques.
- Lack of standards: Since this is a new disruptive proposal, no standards covering the full context exist yet. New concepts, scenarios, use cases and metrics are required, such as Accuracy (the differences between the ON and the DTN) or Time To Update (the time required to synchronize).

VI. CONCLUSIONS AND FUTURE WORK

This work has proposed an architecture to integrate network functions based on Extended Reality (XR) technology within a Digital Twin Network (DTN) scenario. The system is envisioned as a new generation of services for a wide range of use cases, in which users can immerse themselves in the application and interact in a natural and realistic manner. So, 3 architectural entities are defined: ONM in the physical network, and DTM and XRM in the Digital Twin Network. The DTM and XRM must be hosted by robust data center infrastructures with enough processing, storage, and networking capabilities to support the requested DTN+XRF functions. The paper has presented 3 uses cases. The applicability of this architecture is only limited by the number of connected devices.

The envisioned future works for achieving the deployment of the proposed system is the following:

- Deepen the analysis of the modules for each AE, as well as the data flow and interaction between each module.
- Analyze the transmission protocols to establish a robust communication interface between both edge routers in the ONM and DTM.
- To study different modeling techniques for deploying the DTN; while the XR services construction has to be studied.
- Define metrics to standardize and validate the performance of the system.
- Study the architecture of the data center that requires hosting the DTN + XRF system.
- To carry out simulation tests to achieve an initial idea of the capabilities of the system.

ACKNOWLEDGMENTS

This work is part of the I+D+i project PID2022-137329OB-C41, supported by MICIU/AEI/10.13039/501100011033 and FEDER, UE.

REFERENCES

[1] ITU-T, "Digital twin network – requirements and architecture," International Telecommunication Union - Telecommunication Standardization Sector (ITU-T), Geneva, Switzerland, Tech. Rep. Y.3090, Feb. 2022, recommendation ITU-T Y.3090. [Online]. Available: https://handle.itu.int/11.1002/1000/14852

[2] Y. Wu, K. Zhang, and Y. Zhang, "Digital twin networks: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13 789–13 804, 2021.

[3] J. Zhao, X. Xiong, and Y. Chen, "Design and application of a network planning system based on digital twin network," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 900–904, 2022.

[4] Z. Wei, S. Wang, D. Li, F. Gui, and S. Hong, "Data-driven routing: A typical application of digital twin network," in *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, 2021, pp. 1–4.

[5] D. González-Sánchez, L. Bellido, I. D. Martinez-Casanueva, D. Martínez-García, D. Fernández, and D. R. Lopez, "Toward building a digital twin for network operations and management," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 2583–2598, 2025.

[6] M. Jian, B. Long, and H. Liu, "A survey of extended reality in 3gpp release 18 and beyond," p. 542, 2023.

[7] J. Guan, J. Irizawa, and A. Morris, "Extended reality and internet of things for hyper-connected metaverse environments," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2022, pp. 163–168.

[8] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "Nfv and sdn—key technology enablers for 5g networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2468–2478, 2017.

[9] S. Zhang, "An overview of network slicing for 5g," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 111–117, 2019.

[10] X. Shen, J. Gao, W. Wu, M. Li, C. Zhou, and W. Zhuang, "Holistic network virtualization and pervasive network intelligence for 6g," *IEEE Communications Surveys Tutorials*, vol. 24, no. 1, pp. 1–30, 2022.

[11] G. Schroeder, C. Steinmetz, C. E. Pereira, I. Muller, N. Garcia, D. Espindola, and R. Rodrigues, "Visualising the digital twin using web services and augmented reality," in *2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, 2016, pp. 522–527.

[12] C. Yang, X. Tu, J. Autiosalo, R. Ala-Laurinaho, J. Mattila, P. Salminen, and K. Tammi, "Extended reality application framework for a digital-twin-based smart crane," *Applied Sciences*, vol. 12, no. 12, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/12/6030

[13] Y. Liang, G. Li, G. Zhang, J. Guo, Q. Liu, J. Zheng, and T. Wang, "Latency reduction in immersive systems through request scheduling with digital twin networks in collaborative edge computing," *ACM Trans. Sen. Netw.*, Oct. 2024, just Accepted. [Online]. Available: https://doi.org/10.1145/3701562

[14] S. Mihai, M. Yaqoob, D. V. Hung, W. Davis, P. Towakel, M. Raza, M. Karamanoglu, B. Barn, D. Shetve, R. V. Prasad, H. Venkataraman, R. Trestian, and H. X. Nguyen, "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Communications Surveys Tutorials*, vol. 24, no. 4, pp. 2255–2291, 2022.

[15] M. Sanz Rodrigo, D. Rivera, J. I. Moreno, M. Àlvarez Campana, and D. R. López, "Digital twins for 5g networks: A modeling and deployment methodology," *IEEE Access*, vol. 11, pp. 38 112–38 126, 2023.

[16] A. G. Wermann and J. A. Wickboldt, "Ktwin: A serverless kubernetes-based digital twin platform," *Computer Networks*, vol. 259, p. 111095, Mar. 2025. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2025.111095

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Migración en vivo de microservicios con estado para alta disponibilidad en redes asistidas por drones

Sergio Frejo-Martín, Jaime Galán-Jiménez
Departamento de Ingeniería de Sistemas Informáticos y Telemáticos,
Universidad de Extremadura, España
sfrejom@unex.es, jaime@unex.es

Las redes asistidas por drones proporcionan una solución rápida y flexible para el despliegue de conectividad y servicios en entornos con acceso limitado o nulo a la infraestructura convencional. Sin embargo, estas redes están sujetas a restricciones por la autonomía de sus nodos aéreos, que impone reemplazos frecuentes de los mismos, con la consecuente disrupción de los servicios desplegados. Este artículo introduce un modelo de migración en vivo para microservicios con estado en redes asistidas por drones. La propuesta hace uso de contenedores Docker y el mecanismo de precopia para maximizar la disponibilidad, así como una arquitectura compuesta por un Gemelo Digital para monitorización en tiempo real, una Función de Gestión de Enjambre para orquestar reemplazos de drones y un Servidor Central de Imágenes. El modelo propuesto consigue continuidad de servicio y transiciones suaves durante los reemplazos de drones. A fin de probar la viabilidad y rendimiento de la solución presentada, se ha desarrollado un emulador representativo de la misma, cuya implementación y resultados también se presentan.

***Palabras Clave*—Vehículos Aéreos No Tripulados, Arquitectura de Microservicios, Migración en Vivo, Tiempo de inactividad, Computación en el Borde**

I. INTRODUCCIÓN

Las redes asistidas por Vehículos Aéreos No Tripulados (*Unmanned Aerial Vehicles*, UAVs), también conocidos como drones, han emergido como solución innovadora y efectiva para afrontar los desafíos inherentes al despliegue rápido y flexible de infraestructuras de comunicación en áreas donde las redes convencionales resultan imprácticas o están enteramente ausentes. Tales áreas incluyen entornos rurales, regiones aisladas y zonas afectadas por catástrofes en las que las infraestructuras convencionales se han visto comprometidas. El despliegue de redes de comunicaciones en estas circunstancias impone elevados costes, retos logísticos y tiempos de ejecución prolongados. En contraste, la ágil movilidad y los eficientes costes del despliegue de drones convierten a estos en plataformas versátiles y eficaces para proporcionar servicios esenciales de comunicación [1].

Las nuevas capacidades de cómputo y comunicación inalámbrica con las que estos vehículos pueden ser equipados incrementan el potencial de esta solución en aún mayor medida. Los drones pueden ahora dar soporte a soluciones sofisticadas en campos como la computación en el borde (*Edge Computing*) o las redes definidas por software (*Software-Defined Networking*, SDN). Estas capacidades habilitan nuevas oportunidades en dominios muy diversos, entre los que cabe mencionar la gestión de redes y servicios, la monitorización ambiental o la coordinación de respuestas a emergencias. El despliegue de arquitecturas basadas en microservicios (*Microservice-Based Architectures*, MSAs) en redes ad-hoc aéreas también ha atraído el interés de la comunidad investigadora, pues permite explotar la modularidad y escalabilidad de estas soluciones [2], [3]. Sin embargo, las implementaciones prácticas de este concepto implican una serie de limitaciones que no pueden ser pasadas por alto, principalmente relacionadas con la continuidad y fiabilidad del servicio.

La más significativa de estas limitaciones es el tiempo de vuelo de los drones, derivada de la autonomía proporcionada por sus baterías. Este factor impone reemplazos frecuentes de los dispositivos desplegados a fin de poder mantener la funcionalidad de red. Tales reemplazos de nodos de la red pueden darse también como consecuencia de otros problemas técnicos. Estas operaciones de reemplazo no son triviales, puesto que interrumpen los microservicios desplegados y generan potenciales problemas, como pérdida de datos, interrupción del servicio y otros inconvenientes que degradan la experiencia de usuario. Si bien los servicios sin estado, o en los que este no resulta crítico, pueden restaurarse con relativa facilidad, los mi-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

croservicios con estado (esto es, aquellos que conservan información a lo largo del tiempo y entre sesiones) son particularmente vulnerables a las interrupciones que tienen lugar durante los reemplazos de nodos.

Asegurar una operación fluida y continua de microservicios con estado en redes asistidas por drones sigue siendo una tarea pendiente. A fin de resolver esta problemática, este artículo presenta un nuevo *framework* para habilitar la migración con conservación de estado en microservicios desplegados en este tipo de redes.

La importancia de este esfuerzo investigativo reside en su capacidad para materializar el potencial teórico de las redes ad-hoc aéreas y sus implementaciones prácticas. Habilitar el soporte para aplicaciones con estado no solo permite mejorar la experiencia de usuario en tales aplicaciones, sino también ampliar el rango de casos de uso que las redes ad-hoc aéreas son capaces de soportar. En este caso, no solo supone una mejora cuantitativa derivada de la disminución del tiempo de inactividad de los servicios, sino también una mejora cualitativa que habilita el despliegue de servicios críticos y dependientes de la conservación de su estado de ejecución.

Las siguientes secciones incluirán una revisión del estado del arte en lo que refiere a técnicas de migración en vivo (Sec. II), explorando las aproximaciones existentes y su relevancia en el contexto concreto de las redes asistidas por drones. La Sección III ofrece una descripción de la arquitectura propuesta, detallando sus componentes y su flujo de operación. Los mecanismos que habilitan la preservación de estado durante los reemplazos de drones se exponen en la Sección IV. En la misma se describe el proceso de migración en vivo y las estrategias seguidas para minimizar el tiempo de inactividad de los servicios. La siguiente Sección V muestra detalles de la implementación del emulador dedicado a evaluar la viabilidad y el rendimiento de la solución propuesta. A continuación, en la Sección VI se expondrán los resultados obtenidos en lo referido a tiempos de inactividad, tiempos de migración, volumen de tráfico de control generado, etc. Las limitaciones y retos pendientes del sistema serán analizadas en la Sección VII, identificando áreas susceptibles de mejora y refinamiento. Para concluir, la Sección VIII cerrará el artículo con un resumen de los hallazgos y conclusiones extraídos de las pruebas de sistema, así como comentando futuros cauces de investigación orientados a la mejora del rendimiento y la escalabilidad de la solución propuesta.

II. ESTADO DEL ARTE

El concepto de migración en vivo describe la transferencia de procesos o instancias de servicios desde un *host* origen a otro, causando durante la misma la mínima interrupción del servicio. Un estudio recientemente publicado [4] ofrece una revisión exhaustiva de diferentes aproximaciones relevantes a la migración en vivo. Las primeras aproximaciones hacían uso de máquinas virtuales para alojar los servicios a migrar [5]. Por un lado, este enfoque proporciona un entorno de ejecución aislado y reproducible que no depende del sistema operativo o sus bibliotecas. Pero el uso de máquinas virtuales presenta limitaciones como su gran tamaño, del orden de gigabytes, o la dependencia de su hipervisor asociado para migraciones en vivo, que merma la interoperabilidad que tal solución puede ofrecer.

Los contenedores se presentaron como una alternativa ligera a las máquinas virtuales para conseguir entornos de ejecución aislados que pudieran ser guardados y restaurados a conveniencia [6]. El trabajo presentado en [7] implementa una aproximación a la migración de servicios en redes ad-hoc aéreas basada en contenedores sin estado con el fin de optimizar el consumo energético. Si bien es dependiente del sistema operativo y las bibliotecas de las máquinas anfitrionas, este enfoque de migración en vivo reduce drásticamente el volumen de datos a migrar, así como el tiempo requerido para llevar a cabo un arranque en caliente (*warm start*), es decir, para desplegar los servicios en un nuevo anfitrión.

La estructura interna de almacenamiento basada en capas que implementan los contenedores es una parte fundamental de su conveniencia para migraciones en vivo. El sistema de archivos de un contenedor Docker está compuesto por capas de solo lectura jerárquicamente apiladas en las que se describen diferencias en el sistema de archivos. Tales capas se conocen como **capas de imagen base**. Cuando se montan mediante montaje por superposición (`union mount`), el sistema de archivos es restaurado al estado descrito por las capas. En adición a estas, la **capa modificable del contenedor** actúa como una capa adicional con permisos de lectura y escritura que recoge las modificaciones acontecidas durante la ejecución actual del contenedor [8].

La estructura de sistema de archivos basada en capas descrita es especialmente conveniente para refinar el proceso de migración en vivo. Múltiples técnicas de optimización para reducir los tiempos de migración se han convertido en aproximaciones de referencia. Destacan especialmente los métodos de precopia, postcopia y el enfoque híbrido entre ambos. En el método de precopia, la idea es transferir la mayor parte posible del estado del servicio mientras este sigue en ejecución. Con el uso de postcopia se procura transferir los recursos mínimos necesarios para el despliegue del servicio, y después transferir progresivamente el resto durante su ejecución en el nuevo anfitrión. El enfoque híbrido combina ambas técnicas a conveniencia [9].

Si bien la transferencia del sistema de archivos habilita la migración de servicios, esta no es suficiente para llevar a cabo una migración con estado. Los datos restantes requeridos para ello incluyen, pero no se limitan a, el estado de la memoria del proceso (áreas de memoria virtual, segmentos compartidos de memoria y archivos mapeados en memoria), registros de CPU, descriptores de archivos abiertos (incluyendo banderas y posiciones de archivos), conexiones de red activas (como *sockets* TCP o dominios UNIX), credenciales de procesos (UIDs, GIDs, permisos...), temporizadores y manejadores de señales [10], [11]. Estos datos generados en tiempo de ejecución

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

pueden ser sistemáticamente empaquetados y transferidos usando CRIU (*Checkpoint/Restore in Userspace*)[12]. CRIU se ha convertido en la herramienta de referencia para la obtención y restauración de puntos de control (*checkpoints*) de contenedores. Originalmente un proyecto de Virtuozzo, ha pasado a estar integrado en los principales gestores de contenedores y motores de virtualización, tales como OpenVZ, LXC/LXD, Docker o Podman, así como distribuido como paquete para diversas distribuciones de Linux [4].

La migración en vivo se consigue mediante la transmisión de todas las capas del sistema de archivos y un punto de control (*checkpoint*) del anfitrión origen al destino, que monta las capas y carga el *checkpoint*, restaurando el contenedor desde el punto exacto en el que fue guardado. Si bien no existe un consenso claro con respecto al mejor medio para llevar a cabo migraciones en vivo, los contenedores muestran una mayor prevalencia en estudios recientes, principalmente debido a su reducido tamaño y la estandarización de herramientas de gestión de los mismos. Una reciente contribución de Dinh-Tuan y Jiang [13] realiza un análisis de las técnicas de migración en vivo y el manejo granular de capas de contenedores para efectuar migraciones con estado en Kubernetes.

Los enjambres de drones están adquiriendo relevancia como plataforma para el despliegue de arquitecturas de microservicios, en especial entre la comunidad investigadora. Esto se debe en buena parte a ventajas como la alta movilidad, versatilidad y despliegues ágiles y rápidos que ofrecen. El trabajo de investigación sobre estas arquitecturas se ha centrado en buena parte en la optimización del despliegue y localización de microservicios en el enjambre. El trabajo de [7] recurre a un enfoque basado en MILP (Programación Lineal Entera Mixta) para optimizar la localización de microservicios, balanceando dinámicamente el consumo energético en los drones y adaptando la distribución a la demanda de servicios en tiempo real. Pese a implementar la migración de servicios con este fin, el estado de ejecución de los mismos no se toma en consideración. La migración en vivo con conservación de estado es una característica deseable en el contexto de las redes asistidas con drones con soporte para el despliegue de microservicios, pues los reemplazos de drones son frecuentes y llevan consigo interrupciones de servicio. Pese a ello, no se han encontrado trabajos de investigación previos relacionados con este tema.

III. DESCRIPCIÓN DE LA ARQUITECTURA PROPUESTA

El presente artículo propone un modelo para redes asistidas por drones con soporte para alojamiento de microservicios con estado. Los propósitos del sistema son maximizar la disponibilidad de los servicios desplegados, lo que se consigue minimizando el tiempo de inactividad de los mismos que tiene lugar durante las migraciones, y mantener el estado de ejecución a través de dichas migraciones. En conjunto, el fin de estas metas es proporcionar una experiencia consistente y continua, comparable a la de las redes terrestres convencionales con capacidades de computación en el borde.

Para conseguir esto, nuestra propuesta se inspira en el modelo de migración en vivo propuesto por Ma et al. en [8]. Este modelo solventa los problemas de la migración en vivo aprovechando el anteriormente mencionado esquema interno del sistema de archivos de los contenedores para implementar mecanismos de precopia. Su aproximación minimiza la cantidad de datos a transferir durante las migraciones sincronizando capas de imagen base inmutables en múltiples servidores en el borde, y aislando la capa modificable del contenedor. Los autores demuestran que su método reduce las transferencias de servicio en hasta un 80% en comparación a soluciones previamente existentes. Su profundo y sistemático estudio del sistema de almacenamiento de Docker es también destacable y constituye la base para optimizar la gestión y almacenamiento en caché de las capas.

Una adaptación eficaz de este mecanismo de migración a redes asistidas por drones demanda identificar los componentes y procesos involucrados, y alinear los mismos con los procesos y recursos habituales de este tipo de redes. Una representación esquemática del modelo propuesto se muestra en la figura 1. Pueden distinguirse en la misma dos áreas diferenciadas; el **Área de Interés**, que constituye la región en la que los drones proveen conectividad y servicios a sus usuarios, y la **Estación Base**, que concentra la infraestructura de carga para los drones y el **Centro de Inteligencia** (IC). El IC provee capacidades de computación avanzadas para alojar tareas pesadas que no pueden ser soportadas por los drones.

El IC concentra tres componentes, que abordan funciones relacionadas con la monitorización y gestión de la red y los servicios desplegados en la misma:

- **Gemelo Digital (Digital Twin, DT):** Un modelo virtual que representa a cada dron del enjambre. Está diseñado para reflejar con precisión las características técnicas y estado en tiempo real de cada nodo. Los drones envían regularmente actualizaciones de estado al DT, que integra los parámetros técnicos recibidos en su modelo para mantener una representación actualizada de los drones y las condiciones generales de la red [14], [15].
- **Función de Gestión del Enjambre (Swarm Management Function, SMF):** Responsable de orquestar operaciones de los drones en el enjambre, asegurando eficiencia en la coordinación, localización de tareas y distribución óptima de los nodos. Hace uso del DT como fuente de datos para monitorizar el estado de los drones y detectar anomalías como baterías cerca de agotarse o errores técnicos en los drones. Una vez el problema es identificado, el SMF coordina las acciones correctivas, como orquestar el reemplazo del dron afectado, necesarias para mantener el correcto funcionamiento de las operaciones de red.
- **Servidor Central de Imágenes (Central Image Server, CIS):** Un repositorio responsable de gestionar y distribuir imágenes base a la red de drones. Cada

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

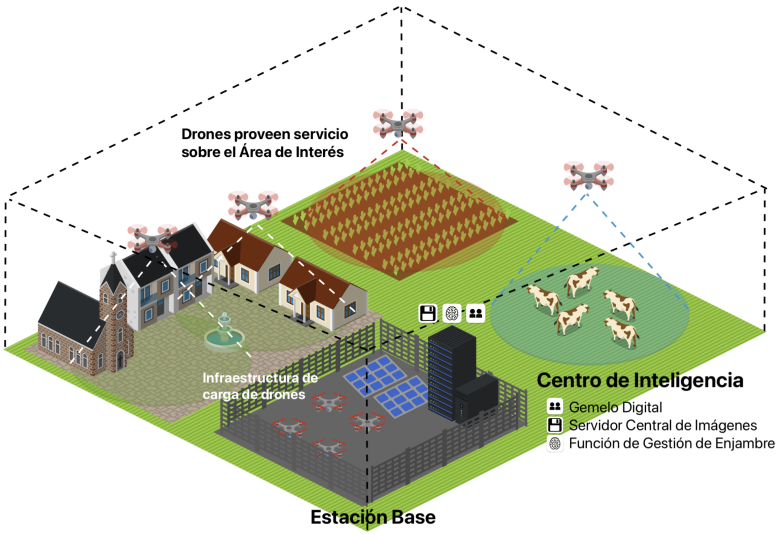


Fig. 1. Diagrama del modelo propuesto

imagen base está asociada a un Rol, que se ve definido por unas coordenadas de localización y un conjunto de servicios a desplegar. Estas imágenes contienen las dependencias y configuraciones de software necesarias para llevar a cabo las funciones asignadas a su rol. El CIS mantiene una versión actualizada de la imagen base de cada rol. Adicionalmente, dado que cada capa del almacenamiento recibe un identificador aleatorio `cacheID` en cada dron, el CIS mantiene una relación entre estos identificadores y el hash SHA256 que identifica unívocamente la capa. Este mecanismo de mapeo contribuye a la sincronización y recuperación eficiente de capas, facilitando el proceso de gestión de imágenes.

Los componentes del IC cooperan con el fin común de gestionar la red. Al principio, los drones son desplegados sobre el Área de Interés designada y los microservicios definidos son puestos a disposición de los usuarios. Desde este momento, los drones comienzan a enviar actualizaciones periódicas de estado al DT, que almacena los datos en la contraparte digital de cada dispositivo. La SMF comprueba regularmente los datos del DT en busca de incidencias. Cuando se detecta un dron acercándose al fin de su batería o experimentando un problema técnico, se inicia el proceso de reemplazo. El SMF y el CIS cooperan para replicar en el dron entrante el mismo estado en que se encontraba el dron saliente cuando este dejó de prestar servicio.

El proceso de reemplazo de los drones está orientado a minimizar el tiempo de inactividad en los servicios alojados. Para conseguir esto, hace uso de la técnica de precopia frecuentemente usada en escenarios de migración en vivo. Antes de abandonar la Estación Base, el dron de reemplazo descarga del CIS la imagen correspondiente al rol a asumir. La imagen del nodo aéreo saliente se sincroniza con el CIS y con el nodo aéreo entrante. Para el momento en el que el reemplazo físico tiene lugar, el volumen de datos a transferir se reduce a la información generada en tiempo de ejecución y a la capa de memoria modificable del sistema de archivos que refleja los cambios realizados en este durante el mismo periodo. De este modo, el periodo de inactividad se ve reducido al tiempo que toma obtener un *checkpoint* y transferirlo junto con la capa modificable del contenedor al dron entrante mediante comunicación dron-dron. Una explicación más detallada de este mecanismo se presenta en la siguiente sección.

IV. MECANISMO DE PRESERVACIÓN DE ESTADO

Las redes asistidas por drones requieren del reemplazo periódico de sus nodos aéreos. Este es un proceso crítico que consume cantidades sustanciales de energía y recursos computacionales y causa la interrupción de los servicios desplegados en el dron. En la mayoría de arquitecturas de redes asistidas por drones propuestas, los servicios no son solo interrumpidos, sino también terminados, requiriendo que estos vuelvan a lanzarse en el dron de reemplazo. El hecho de que este proceso tenga lugar periódicamente para cada dron en el enjambre subraya la necesidad de optimización del mismo. El modelo de red propuesto en este artículo incorpora un mecanismo de reemplazo de drones concebido para maximizar la disponibilidad del servicio al minimizar el tiempo de inactividad y aprovechar la arquitectura de red para brindar soporte para la migración de servicios con estado.

El proceso está dividido en fases, dado que algunos pasos tienen lugar de forma simultánea en diferentes puntos de la arquitectura. Las interacciones entre los diferentes actores involucrados pueden ser vistas en el diagrama de secuencia de la figura Fig 2. Las etapas del proceso son:

1) **Fase 0 - Comienzo del proceso de reemplazo:** En el estado por defecto del sistema, los drones

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

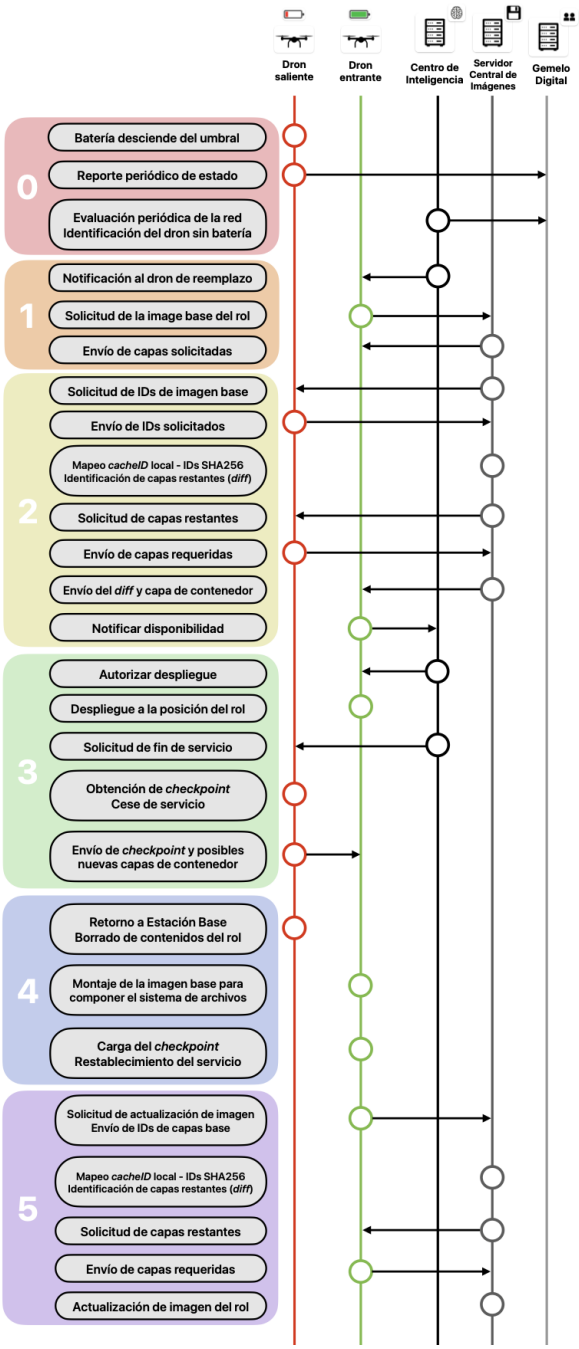


Fig. 2. Diagrama de secuencia del proceso de reemplazo de un dron.

proveen servicio a los usuarios finales y dispositivos IoT (*Internet of Things*) sobre el Área de Interés designada. Mientras, en el Centro de Inteligencia, la SMF consulta periódicamente el DT para monitorizar el estado de la red. Cuando la batería de un dron cae por debajo del umbral mínimo o reporta un problema técnico, el sistema comienza el proceso de reemplazo y preservación de estado para asegurar la operación ininterrumpida.

2) **Fase 1 - Configuración pre-despliegue del dron de reemplazo:** la Función de Gestión de Enjambre designa a uno de los drones disponibles en el IC como reemplazo del dron comprometido. El dispositivo de reemplazo es informado del rol que ha de desempeñar, lo cual especifica la posición que ha de ocupar y los microservicios que alojará. Este procede a solicitar la imagen base asociada a su rol al CIS, que incluye las capas que constituyen el estado más reciente conocido del sistema de archivos del dron a reemplazar.
3) **Fase 2 - Sincronización de Imagen Base:** el CIS obtiene del dron saliente los identificadores locales (IDs) de las capas que componen su imagen base. Después, mapea estos IDs de caché local a los IDs unívocos derivados del hash SHA256 e identifica las capas ausentes en el dron de reemplazo. El CIS solicita al dron saliente las capas restantes junto con su capa modificable del contenedor. Para transferir la capa modificable del contenedor, el contenedor asociado debe realizar un `commit` que transforme dicha capa en una nueva imagen base y cree una nueva capa modificable del contenedor. Cuando las capas son recibidas en el CIS, este actualiza la versión local de la imagen y envía las nuevas al dron de reemplazo. Una vez todas las capas del sistema de archivos han sido recibidas, el dron de reemplazo pasa a estar listo para el despliegue, lo cual notifica al SMF.
4) **Fase 3 - Cesión del servicio:** el SMF autoriza el despliegue del dron de reemplazo una vez el estado de la imagen base es uniforme en ambos drones involucrados y el CIS. El dron entrante vuela al área designada. Una vez este ha despegado, el dron saliente recibe la orden de cese de servicio. Para preservar su estado de ejecución, este dron crea un *checkpoint*, lo prepara para la transmisión y detiene el contenedor. Esto marca un momento crítico en el proceso durante el cual se produce una breve interrupción del servicio, inevitable para garantizar la consistencia de su estado. Si el contenedor no fuera detenido, cualquier cambio en el estado de ejecución que se diese durante la transmisión del *checkpoint* llevaría a un estado inconsistente. Es importante apreciar que el *checkpoint* representa una pequeña fracción del volumen total de datos que compone el estado del servicio. La mayor parte de la transferencia de datos tuvo lugar en la fase previa, mientras el nuevo dron seguía aún cargando en el IC. Una vez se consigue establecer conexión entre el dron saliente y entrante, el primero transfiere el *checkpoint* y la capa modificable del contenedor, que es de nuevo transferida como capa de imagen base.
5) **Fase 4 - Restablecimiento del servicio:** el dron entrante contiene ahora los recursos necesarios para relanzar el servicio desde el mismo punto previo al despliegue. Para conseguir esto, inicia el `daemon` de Docker, monta las capas recibidas para restaurar el sistema de archivos y carga el *checkpoint*. Los

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

microservicios son entonces restablecidos al mismo punto en el que el contenedor anterior fue detenido. Mientras, el dron saliente borra los archivos asociados con su rol previo y vuelve a la estación base, donde se le proveerá carga y, de requerirla, asistencia técnica.

6) **Fase 5 - Sincronización con el CIS:** una vez el proceso de despliegue ha sido completado y los microservicios han sido restaurados en el dron de reemplazo, se requiere de un paso final para garantizar la consistencia del sistema. Para ello, el CIS actualiza su versión de la imagen del rol integrando las capas restantes generadas durante la transferencia, de haberlas. Esto garantiza que la copia de seguridad almacenada es una representación precisa del estado operacional más reciente, de modo que futuros drones a desplegar reciban la configuración y entorno de ejecución más recientes conocidos.

V. EMULADOR DE LA ARQUITECTURA

Para evaluar la viabilidad y rendimiento de la arquitectura propuesta en un entorno controlado y reproducible, se ha procedido a la implementación de un emulador haciendo uso de Python 3.10 y Kathará [1] en su versión 3.7.7. El uso de Kathará como herramienta de emulación de redes ha sido clave para obtener un alto grado de fidelidad, pues es capaz de ofrecer un comportamiento fidedigno de los dispositivos de red emulados y de proveer entornos aislados de ejecución gracias a su enfoque basado en contenedores [16]. En lo que respecta al despliegue de microservicios, se ha hecho uso de Flask para la implementación de servidores web sencillos.

El esquema de red emulado está constituido por dos subredes. En primer lugar, una red de gestión que contiene la totalidad de los drones involucrados y los dispositivos de gestión (DT, SMF y CIS). Esta concentra el tráfico de control y la transferencia de capas de imagen base y *checkpoints*. Por otra parte, la red de servicio, compuesta por los drones desplegados y los usuarios, aloja el tráfico relacionado con el consumo de microservicios por parte de los usuarios de la red. Kathará representa cada dispositivo de red mediante un contenedor Docker. De igual manera, cada dron desplegado inicia un contenedor que se conecta a la red de servicio y aloja en un entorno aislado los microservicios a desplegar.

El emulador permite la configuración de múltiples parámetros para dar forma a la topología a desplegar, como el número de drones en la flota, la cantidad de roles a cubrir, la cantidad de usuarios y las imágenes asociadas a cada rol. También permite establecer los tiempos de duración y carga de la batería, para poder agilizar las pruebas. La herramienta está disponible para su uso y modificación por parte de la comunidad investigadora[2].

Sobre este emulador se han ejecutado diversos escenarios de prueba. El propósito de estas pruebas ha sido la recolección de métricas relacionadas con el tiempo de inactividad de los servicios y el volumen de tráfico derivado de la introducción del mecanismo de preservación de estado. Se ha procurado mantener las condiciones de las pruebas constantes entre experimentos, a fin de que los resultados sean comparables. Las pruebas contemplan en todos los casos cuatro roles a cubrir, una flota de 20 drones y un usuario para comprobar la continuidad del servicio. Para cada experimento, la ejecución se prolonga durante dos ciclos de reemplazo de cada rol, a fin de obtener una muestra de 8 migraciones en cada una. Los experimentos difieren entre sí en el tipo de microservicio desplegado en los drones, mostrando estos complejidad y carga crecientes para comprobar la correcta migración con estado de los mismos.

Tabla I
RESULTADOS DE LOS TRES EXPERIMENTOS

Métrica	Experimento 1	Experimento 2	Experimento 3
Tiempo promedio de inactividad (s)	14	18	22
Volumen de tráfico en red de gestión (MB)	3356.7	5780.6	6262.3
Volumen de tráfico excluyendo CIS (MB)	0.75	0.68	0.70

VI. EVALUACIÓN DE RESULTADOS

La ejecución de distintas rondas de prueba sobre el emulador implementado constata la viabilidad del enfoque propuesto. El modelo de red demuestra capacidad para sustentarse a sí mismo y prolongar su funcionamiento continuo. El promedio total de tiempo de inactividad de los servicios durante migraciones es de 18 segundos, tras el cual la ejecución de estos se reanuda desde el mismo punto en el nuevo dron. Esto es un margen de tiempo asumible para garantizar una calidad de servicio adecuada, y constituye una fracción del tiempo total de transferencia, carga y despliegue de la imagen en el dron de reemplazo.

La Tabla I muestra las métricas más relevantes asociadas a cada ronda de prueba ejecutada sobre el emulador. Entre cada experimento varía únicamente el servicio alojado en los drones desplegados. En el Experimento 1 se ejecuta la transmisión de una secuencia numérica a los usuarios, en el Experimento 2, se transmiten imágenes acompañadas de un número de secuencia, simulando un servicio de monitorización, y en el Experimento 3 se despliega un servicio de *streaming* de vídeo. De este modo, una ejecución correcta muestra una salida continua en el lado usuario, sin saltos en la secuencia de números recibidos, el conjunto de imágenes ni los fotogramas del vídeo retransmitido.

Examinando los datos de la Tabla I y la Figura 3 sobre el tiempo promedio de inactividad, puede apreciarse una tendencia al alza. Cabe mencionar que esta no responde a incrementos en el tamaño de las imágenes contenidas en los drones, que es respectivamente de 280MB, 487MB y 500MB para los Experimentos 1, 2 y 3. En su lugar, el tiempo promedio de inactividad crece en función del volumen de datos generados en tiempo de ejecución, de ahí el incremento en más de un 50% apreciable en el Experimento 3 (*streaming* de vídeo) con respecto al Experimento 1 (envío de secuencia numérica). La misma conclusión puede extraerse del gráfico de la figura 4, en

[1]Kathará - Network Emulation System https://www.kathara.org

[2]Herramienta LiveMigrationSwarm en GitHub https://github.com/sfrejom/LiveMigrationSwarm

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

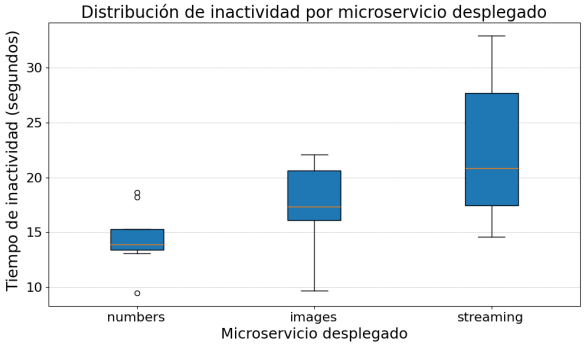


Fig. 3. Diagrama de cajas de tiempo de inactividad por microservicio.

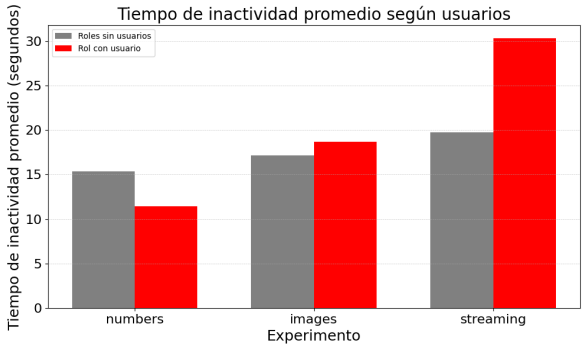


Fig. 4. Gráfica comparativa del tiempo de inactividad entre roles con y sin usuarios.

la cual se compara el tiempo promedio de migración de los roles con usuario asignado y sin usuario asignado.

Si consideramos el volumen de tráfico reportado, la progresión ascendente sí es proporcional al tamaño de las imágenes transferidas. Por eso resulta interesante excluir del análisis el volumen de tráfico del CIS y agregar solo el procedente de la SMF y el DT, que es el tráfico estrictamente de control. Esto nos permite ver que el volumen de tráfico de control se mantiene constante y es marginal con respecto al total. En el caso del CIS, el elevado volumen de tráfico soportado puede acarrear problemas en términos de escalabilidad. Cuando el sistema comienza a funcionar, todos los drones a desplegar solicitan sus respectivas imágenes simultáneamente, causando un gran pico de tráfico en el CIS. Una vez completado el despliegue inicial, las variaciones en carga computacional, distancias a recorrer y condiciones ambientales hacen que las baterías en los drones se descarguen a ritmos distintos, reduciéndose la cantidad de migraciones simultáneas a satisfacer por el CIS.

VII. DESVENTAJAS Y LIMITACIONES

La implementación de la continuidad de servicio en redes asistidas por drones introduce ciertos desafíos a tener en cuenta. Múltiples áreas susceptibles de mejora han sido identificadas durante el diseño e implementación del sistema. Esta sección realizará una revisión crítica de las limitaciones encontradas.

Una contrapartida del soporte de servicios con estado es el incremento en el tráfico de control que soporta la red como consecuencia del mecanismo de preservación de estado. El proceso de sincronizar las capas que componen una imagen base entre el dron anfitrión, el CIS y, eventualmente, el dron de reemplazo, requiere de tráfico de control dedicado, junto con transmisión frecuente de capas de almacenamiento en la red. Derivado de esto cabe esperar incrementos en el consumo energético, utilización del ancho de banda y carga de CPU en el conjunto de la red. Si bien un decrecimiento moderado del rendimiento y eficiencia de la red es un compromiso esperable por la implementación de la preservación de estado, ciertos esfuerzos de optimización en el proceso de transmisión pueden ayudar a mitigar el impacto.

Por otra parte, el Centro de Inteligencia constituye un punto único de fallo, lo cual es potencialmente preocupante. Este componente incluye elementos vitales para la coordinación de la red, como la instancia de la Función de Gestión de Enjambres o el Servidor Central de Imágenes, haciendo críticos los fallos en cualquiera de estos componentes. Incluso cuando funcionan de forma correcta, tales componentes pueden constituir cuellos de botella que limiten la escalabilidad del modelo, como es el caso del CIS durante los picos de tráfico que experimenta. Esto podría solventarse haciendo uso de las capacidades computacionales de los drones para implementar versiones distribuidas de los componentes del IC. Una Función de Gestión de Enjambres descentralizada requeriría de mecanismos de consenso para coordinar decisiones de gestión de la red. Del mismo modo, una versión distribuida del Servidor Central de Imágenes (es decir, Servidores Distribuidos de Imágenes) podría utilizar el enjambre para almacenar imágenes de drones vecinos a modo de copia de seguridad. Sin embargo, la introducción de redundancia a este nivel incrementaría la complejidad de mantener las copias de la imagen base actualizadas. Finalmente, para aumentar la robustez del Digital Twin podría implementarse como solución en la nube, o explorar esquemas descentralizados del mismo.

Una limitación clave final a destacar del modelo propuesto es su susceptibilidad a fallos repentinos en los drones. Si un dispositivo sufre un fallo crítico y cae, el estado de ejecución de los servicios se pierde, interrumpiendo el servicio. Si bien esto podría mitigarse mediante copias de seguridad del estado frecuentes, el incremento en el tráfico y consumo energético hace de esta una solución impracticable. Asegurar la continuidad de los servicios en estos casos sigue siendo un reto sin resolver, subrayando la necesidad de desarrollar estrategias más resilientes con tolerancia a fallos. El mantenimiento predictivo puede reducir la incidencia de estos fallos críticos, pero anticiparse a cada problema a tiempo para llevar a cabo un reemplazo preventivo no siempre será factible.

En términos generales, si bien el modelo cumple con los objetivos definidos, los retos y problemas reseñados en esta sección evidencian que existe margen de mejora.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Abordar estas limitaciones requiere una combinación de mejoras estructurales y optimizaciones estratégicas. Por suerte, las redes asistidas por drones destacan por su versatilidad y adaptabilidad, lo que permite una amplia variedad de recursos interesantes, como los basados en computación en el borde o en redes definidas por software. Estos recursos son clave para la implementación de soluciones como la descentralización de componentes, funciones optimizadas de gestión de red, mantenimiento preventivo, o el mecanismo de preservación de estado en sí mismo.

VIII. CONCLUSIONES Y TRABAJO FUTURO

Las redes asistidas por drones se han consolidado como una solución ágil para el despliegue rápido de redes ad-hoc. Sin embargo, estas redes aún afrontan limitaciones intrínsecas, principalmente relacionadas con la necesidad de reemplazar periódicamente los nodos de la red cuando su batería se agota.

Para hacer frente a esto, este artículo presenta un modelo de redes asistidas por drones ajustado para minimizar el impacto de los reemplazos de nodos. El *framework* propuesto garantiza la continuidad de servicio aprovechando un mecanismo de migración en vivo que hace uso de precopia y una estrategia de preservación de estado basada en entornos de ejecución en contenedores. Este enfoque mitiga eficazmente el tiempo de inactividad y asegura que los microservicios mantengan un estado consistente durante el proceso de migración. Esto constituye una mejora en las capacidades de la red en su conjunto, habilitando el despliegue de servicios en los que la conservación de estado es crítica.

Las pruebas ejecutadas sobre el emulador basado en la arquitectura propuesta arrojan resultados adecuados que constatan su viabilidad. La transferencia de los servicios y el estado de ejecución de los mismos se integra de forma óptima en el proceso de reemplazo de los drones, gracias a la cooperación de los distintos agentes de gestión. El resultado son tiempos de inactividad reducidos en reemplazos que no perjudican la experiencia de uso de los servicios.

A pesar de estos avances, varios desafíos persisten aún. El principal reto a afrontar pasa por mejorar la solución propuesta en términos de escalabilidad y resiliencia. Para esto se propone el diseño de alternativas distribuidas de los componentes de gestión de la red para su despliegue en los nodos aéreos. Además, el mecanismo de migración y preservación de estado genera tráfico adicional en la red, haciendo convenientes futuras labores de optimización que equilibren eficiencia y rendimiento. Finalmente, si bien este trabajo asume un entorno controlado, los fallos repentinos en los drones siguen siendo un desafío para la continuidad del servicio. El desarrollo de soluciones a estos problemas que aprovechen la versatilidad y adaptabilidad de los drones es un área abierta para futuras etapas de esta investigación.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades (proyecto PID2021-124054OB-C31) y por la Unión Europea, Fondo Europeo de Desarrollo Regional, y la Junta de Extremadura. Autoridad de Gestión: Ministerio de Hacienda. Número de expediente de la ayuda: GR24099.

REFERENCIAS

[1] J. Galán-Jiménez, E. Moguel, J. García-Alonso, and J. Berrocal, "Energy-efficient and solar powered mission planning of uav swarms to reduce the coverage gap in rural areas: The 3d case," *Ad Hoc Networks*, vol. 118, p. 102517, 2021. [Online]. Available: https://doi.org/10.1016/j.adhoc.2021.102517

[2] S. García-Gil, J. G. de-la Hiz, A. García-López, J. M. Murillo, and J. Galán-Jiménez, "Qos evaluation of edge computing microservice-based applications in uav ad hoc networks," *IEEE*, 2024, https://doi.org/10.21203/rs.3.rs-4817788/v1.

[3] A. García-López, D. Ramos-Ramos, S. García-Gil, J. Gómez-delaHiz, J. M. Murillo, and J. Galán-Jiménez, "Microservices adaptability and self-organization of uav-assisted wireless networks in failure rural scenarios," *Research Square*, 2024. [Online]. Available: https://doi.org/10.21203/rs.3.rs-4817788/v1

[4] W. Soussi, G. Gür, and B. Stiller, "Democratizing container live migration for enhanced future networks - a survey," *ACM Computing Surveys*, vol. 57, no. 4, pp. 1–37, 2024. [Online]. Available: https://doi.org/10.1145/3704436

[5] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proceedings of the 2nd Symposium on Networked Systems Design & Implementation (NSDI '05)*. Berkeley, CA, USA: USENIX Association, 2005, pp. 273–286. [Online]. Available: https://www.usenix.org/conference/nsdi-05/live-migration-virtual-machines

[6] R. Morabito, "Power consumption of virtualization technologies: An empirical investigation," in *2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC)*. IEEE, 2015, pp. 522–527. [Online]. Available: https://doi.org/10.1109/UCC.2015.93

[7] S. García-Gil, D. Ramos-Ramos, J. Berrocal, J. M. Murillo, and J. Galán-Jiménez, "Microservices migration: A pathway to improved energy efficiency in uav networks," *Internet of Things*, vol. 30, p. 101463, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660524004049

[8] L. Ma, S. Yi, N. Carter, and Q. Li, "Efficient live migration of edge services leveraging container layered storage," *IEEE Transactions on Mobile Computing*, vol. 18, no. 9, pp. 2020–2031, 2019. [Online]. Available: https://doi.org/10.1109/TMC.2018.2871842

[9] A. Choudhary, M. C. Govil, G. Singh, L. K. Awasthi, E. S. Pilli, and D. Kapil, "A critical survey of live virtual machine migration techniques," *Journal of Cloud Computing: Advances, Systems and Applications*, vol. 6, no. 23, 2017.

[10] CRIU Project, "Checkpoint/restore in userspace (criu) documentation," 2024, accessed: 30-Jan-2025. [Online]. Available: https://criu.org/Checkpoint/Restore

[11] ——, "Criu wiki: Image files and checkpoint storage format," 2024, accessed: 30-Jan-2025. [Online]. Available: https://criu.org

[12] R. A. Addad, D. L. C. Dutra, M. Bagaa, T. Taleb, and H. Flinck, "Towards a fast service migration in 5g," in *2018 IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2018, pp. 1–6. [Online]. Available: https://doi.org/10.1109/CSCN.2018.8581836

[13] H. Dinh-Tuan and J. Jiang, "Optimizing stateful microservice migration in kubernetes with ms2m and forensic checkpointing," in *2025 28th Conference on Innovation in Clouds, Internet and Networks (ICIN)*, 2025, pp. 83–90.

[14] L. Lei, G. Shen, L. Zhang, and Z. Li, "Toward intelligent cooperation of uav swarms: When machine learning meets digital twin," *IEEE Network*, vol. 35, no. 1, pp. 386–392, 2021.

[15] B. Li, W. Liu, W. Xie, N. Zhang, and Y. Zhang, "Adaptive digital twin for uav-assisted integrated sensing, communication, and computation networks," *arXiv preprint*, vol. 2310.17470, 2023. [Online]. Available: https://arxiv.org/abs/2310.17470

[16] G. Bonofiglio, V. Iovinella, G. Lospoto, and G. Di Battista, "Kathará: A container-based framework for implementing network function virtualization and software defined networks," in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, 2018, pp. 1–9.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Sistema de Reconocimiento Facial Federado en Conjuntos Abiertos basado en OpenMax

Ander Galván, Marivi Higuero, Jorge Sasiain, Eduardo Jacob
Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), 48013 Bilbao, España
{ander.galvan, marivi.higuero, jorge.sasiain, eduardo.jacob}@ehu.eus

El reconocimiento facial impulsado por Inteligencia Artificial ha demostrado una alta precisión en algunos escenarios y aplicaciones. Sin embargo, presenta desafíos relacionados con la privacidad y la identificación de personas, especialmente considerando que pueden aparecer sujetos desconocidos para el sistema que lo implementa. En este trabajo, se propone el diseño, la implementación y la evaluación de un sistema de reconocimiento facial en un escenario de aprendizaje federado, orientado a conjuntos abiertos. Concretamente, se diseña una solución basada en el algoritmo OpenMax para escenarios de aprendizaje federado. La propuesta emplea el intercambio de los vectores de activación promedio y las distancias locales para identificar de manera eficaz tanto personas conocidas como desconocidas. Los experimentos realizados demuestran la implementación efectiva de la solución propuesta.

***Palabras Clave*—Inteligencia Artificial, Deep Learning, Reconococimiento Facial, Aprendizaje Federado, Conjuntos Abiertos**

I. INTRODUCCIÓN

En los últimos años, la Inteligencia Artificial, específicamente el Deep Learning, ha adquirido gran relevancia en diversos sectores. Entre sus aplicaciones más destacadas, los sistemas de reconocimiento facial se han consolidado como herramientas clave para la identificación y verificación biométrica. Gracias al Deep Learning, estos sistemas alcanzan altos niveles de precisión, lo que ha favorecido su implementación en ámbitos como el control de acceso, la autenticación biométrica y la videovigilancia.

No obstante, su despliegue presenta desafíos asociados a la privacidad de datos en escenarios de aprendizaje colaborativo. En este sentido, el aprendizaje federado surge como una estrategia efectiva al evitar la centralización de datos, manteniendo la información en los dispositivos locales, lo que contribuye al cumplimiento de normativas como el Reglamento General de Protección de Datos [1].

Por otro lado, los sistemas de reconocimiento facial deben adaptarse a contextos donde no todas las identidades están disponibles en el entrenamiento. En estos casos, resulta esencial no solo reconocer personas conocidas, sino también identificar imágenes de individuos desconocidos. Este reto se aborda mediante el enfoque de reconocimiento en conjuntos abiertos, que exige la detección de identidades no vistas previamente.

En este contexto, el presente estudio contribuye mediante el diseño, la implementación y la evaluación de un sistema de reconocimiento facial en un escenario de aprendizaje federado orientado a conjuntos abiertos, garantizando la precisión y privacidad de los datos. Específicamente, se emplea un escenario de aprendizaje federado a lo largo de todo el proceso, en el cual, en primer lugar, se entrena y evalúa un modelo utilizando imágenes de un conjunto de personas conocidas. Posteriormente, se lleva a cabo el diseño de la solución propuesta basada en el algoritmo OpenMax [2] con el objetivo de ajustar el modelo para la identificación y clasificación de imágenes de personas no incluidas en el conjunto de entrenamiento previo.

El resto del artículo se organiza así: la sección II presenta el contexto del estudio; la sección III revisa estudios sobre reconocimiento facial y aprendizaje federado; la sección IV detalla el diseño y la sección V la implementación del sistema; la sección VI presenta los resultados; y la sección VII resume las conclusiones y líneas futuras.

II. CONTEXTO TECNOLÓGICO

El Deep Learning se ha consolidado como una técnica poderosa inspirada en el funcionamiento del cerebro humano, capaz de abordar tareas como la detección de objetos, el reconocimiento de voces y la traducción automática, destacando por su capacidad de aprendizaje sin supervisión directa [3]. Entre sus modelos más relevantes se encuentran las redes neuronales convolucionales, especializadas en el procesamiento de imágenes mediante capas convolucionales, capas de pooling y funciones de activación no lineales que permiten la extracción de características. En tareas de clasificación, estas redes suelen incluir capas completamente conectadas y emplear la función de activación SoftMax para interpretar las salidas como distribuciones de probabilidad sobre las clases [4].

Por su parte, el aprendizaje federado permite entrenar modelos de forma colaborativa sin necesidad de centralizar los datos, preservando así la privacidad de los clientes [5]. En este enfoque, un servidor inicializa y distribuye el

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

modelo a los clientes para su entrenamiento local, recopila los parámetros actualizados y los agrega, repitiendo este ciclo hasta alcanzar la convergencia.

En reconocimiento facial, la mayoría de modelos se entrenan en escenarios de conjunto cerrado, donde las imágenes de prueba corresponden solo a identidades conocidas. Esto provoca que rostros de personas no vistas durante el entrenamiento se clasifiquen erróneamente como conocidas con alta confianza [6]. Para resolver este problema, OpenMax [2] se presenta como una alternativa a la función de activación SoftMax, diseñada para adaptar el comportamiento de los modelos a conjuntos abiertos, permitiendo la clasificación de imágenes de personas con las que no se ha entrenado el modelo, denominadas personas desconocidas, sin asignarlas erróneamente a alguna de las clases conocidas.

Más específicamente, este algoritmo implica calcular un vector de activación promedio (Mean Activation Vector, MAV) para cada clase conocida, obtenido a partir de las activaciones (las salidas de la última capa totalmente conectada antes de SoftMax) de las muestras correctamente clasificadas durante el entrenamiento. A continuación, se calcula la distancia entre cada muestra correctamente clasificada y su MAV correspondiente, y se ajusta una distribución de Weibull utilizando las η mayores distancias (tail size). Esta distribución estadística permite estimar la probabilidad de que una nueva muestra no pertenezca a una clase conocida.

Durante la fase de inferencia, el vector de activación de la muestra a clasificar se reajusta en función de estas probabilidades, considerando únicamente las α clases con mayores niveles de activación (alpha rank). Además, se calcula una activación para la clase desconocida, la cual se integra al conjunto de clases conocidas. Finalmente, se calculan las probabilidades de las clases conocidas y de la clase desconocida. Si la clase desconocida obtiene la mayor probabilidad, o si la probabilidad de las clases conocidas cae por debajo de un umbral ϵ, la muestra se clasifica como desconocida.

III. ESTADO DEL ARTE

El aprendizaje federado emergió como una estrategia efectiva para entrenar modelos de forma colaborativa, sin necesidad de centralizar los datos, preservando la privacidad de los datos de los clientes involucrados. En el contexto del reconocimiento en conjuntos abiertos mediante aprendizaje federado, FedPD [7] propuso un desentrelazado de parámetros que permitió separar el aprendizaje de clases conocidas y desconocidas, mientras que FedOSS [8], enfocado en el ámbito médico, generó muestras virtuales de clases desconocidas mediante discrepancias y colaboración entre clientes, mejorando así la delimitación de las fronteras de decisión.

En el contexto del reconocimiento facial mediante aprendizaje federado, diferentes estudios abordaron aspectos relacionados con la eficiencia y la privacidad durante el entrenamiento. Por ejemplo, FedFace [9] se enfocó en la colaboración entre dispositivos que contenían imágenes de identidad única, mientras que FedGC [10] utilizó técnicas de corrección de gradientes para mejorar la convergencia y preservar la confidencialidad de los datos. FedFR [11] combinó la optimización de modelos globales y locales mediante un módulo de personalización desacoplada. Más recientemente, AdaFedFR [12] introdujo un aprendizaje adaptativo de representaciones entre clases para mejorar la capacidad de generalización y la eficiencia en la comunicación, y FedFS [13] empleó técnicas de aprendizaje autosupervisado para optimizar las representaciones faciales personalizadas. Asimismo, trabajos como [14] integraron estrategias de agregación segura y otras medidas adicionales para garantizar la privacidad de los datos durante el proceso de entrenamiento federado.

A pesar de estos avances, no se ha realizado un estudio que integre de manera conjunta el aprendizaje federado y el reconocimiento facial en conjuntos abiertos, áreas en las que se enmarca el presente trabajo.

IV. DESCRIPCIÓN DE LA PROPUESTA

El presente trabajo aborda el diseño, la implementación y la evaluación de un sistema de reconocimiento facial orientado a conjuntos abiertos. Dicho trabajo se desarrolla sobre un escenario de aprendizaje federado con el objetivo de garantizar la privacidad de las imágenes faciales de los clientes, sin comprometer la precisión del sistema de reconocimiento facial.

El sistema consiste en un modelo de reconocimiento facial que se entrena con un conjunto de personas conocidas. El propósito es entrenar un modelo capaz de clasificar dichas personas conocidas. Para analizar dicha capacidad de clasificación, se lleva a cabo una fase de inferencia en conjuntos cerrados, utilizando imágenes no previamente vistas por el modelo de dichas personas conocidas.

Seguidamente, se lleva a cabo una fase de inferencia en conjuntos abiertos, en la que se emplean imágenes tanto de personas conocidas como desconocidas. Además, para realizar una correcta clasificación, en lugar de utilizar la función de activación SoftMax, se implementa la solución propuesta basada en el algoritmo OpenMax [2], la cual permite identificar con mayor precisión a las personas conocidas y etiquetar como desconocidas aquellas que no forman parte del conjunto de entrenamiento.

Sin embargo, tal y como se describe en la sección II, OpenMax hace uso de las imágenes utilizadas durante el entrenamiento para determinar si una imagen debe clasificarse como perteneciente a una persona conocida o como desconocida. Esto requiere que el servidor tendría que tener acceso a las bases de datos de los clientes, lo cual quebranta los principios del aprendizaje federado, que evita el intercambio de datos sensibles. Alternativamente, cada cliente podría calcular las activaciones correspondientes a las imágenes correctamente clasificadas de cada persona conocida, delegando al servidor la responsabilidad de completar el resto del proceso de OpenMax. Sin embargo, dado que el servidor no tiene por qué ser una entidad confiable, que posee el modelo ya entrenado y que recibe activaciones correspondientes a imágenes de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

diferentes clientes, la privacidad de las imágenes de estos se puede ver comprometida.

En consecuencia, la solución planteada hace uso de una propuesta basada en OpenMax para escenarios de aprendizaje federado, como se muestra en la Figura 1. En este caso, una vez entrenado el modelo, cada cliente realiza las siguientes operaciones: (1) calcula las activaciones de las imágenes correctamente clasificadas de cada persona conocida; (2) calcula el MAV correspondiente para cada persona conocida, utilizando las activaciones previamente calculadas; (3) calcula las distancias entre las activaciones de las imágenes correctamente clasificadas y su MAV correspondiente; y (4) envía al servidor tanto el MAV como las distancias calculadas para cada persona conocida. Finalmente, (5) el servidor agrupa la información recibida por persona conocida, calculando la media de los MAV y concatenando las distancias asociadas de los diferentes clientes, con el propósito de completar el proceso de OpenMax de manera centralizada.

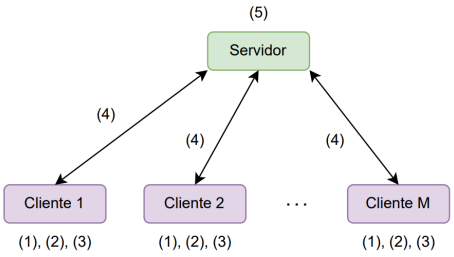


Fig. 1. Solución propuesta basada en OpenMax para escenarios de aprendizaje federado.

En este escenario, el servidor (o las entidades capaces de captar el intercambio de información entre los clientes y el servidor) recibe y almacena únicamente el MAV y las distancias asociadas a cada persona conocida, calculadas de manera local por cada cliente. La transmisión desde cada cliente al servidor del MAV asociado a cada persona conocida, calculado como la media de múltiples activaciones, proporciona un nivel de privacidad superior y disminuye la posibilidad de reconstrucción de características específicas de las imágenes originales, manteniendo así la privacidad de las mismas. Además, las distancias transmitidas no contienen información sensible, ya que son valores derivados de las activaciones. Por consiguiente, compartir dichas distancias no compromete la privacidad de las imágenes.

V. IMPLEMENTACIÓN

La presente trabajo se ha desarrollado en un servidor Dell PowerEdge R760, equipado con las siguientes especificaciones técnicas: 80 CPU Intel Xeon Silver 4416+, memoria RAM de 128 GB y dos GPU NVIDIA A40 de 48 GB de memoria cada una de ellas. En cuanto al software, el trabajo se ha desarrollado en el lenguaje de programación Python 3.10.12.

Con el propósito de validar y añadir robustez a los resultados obtenidos, se ha replicado el escenario del presente trabajo utilizando dos bases de datos diferentes: MS1M ArcFace [15] y VGGFace2 [16]. Por un lado, MS1M ArcFace es una base de datos de imágenes faciales derivada de la base de datos original MS-Celeb-1M [17], la cual ha sufrido procesos de limpieza y preprocesamiento. Esta versión mejorada contiene aproximadamente 5.8 millones de imágenes correspondientes a 85.000 personas. Por otro lado, VGGFace2 se trata de otra base de datos de imágenes faciales publicada en el año 2018, que engloba aproximadamente 3.3 millones de imágenes de 9.000 personas diferentes.

Para este trabajo se han seleccionado 70 individuos conocidos con el fin de mantener un número manejable de clases que haya permitido optimizar el entrenamiento en un entorno federado compuesto por 5 clientes, evitando así la sobrecarga en la capacidad computacional del servidor utilizado. Cada cliente ha dispuesto de múltiples imágenes de cada individuo (60 para entrenamiento y 50 adicionales para inferencia en conjuntos cerrados), las cuales han cubierto variaciones en poses, iluminación y expresiones, simulando condiciones realistas en el reconocimiento facial. Además, en la fase de inferencia en conjuntos abiertos, se han incluido imágenes de 8.000 personas desconocidas para evaluar la capacidad del modelo de identificar correctamente individuos no presentes durante el entrenamiento.

El algoritmo de agregación del servidor combina los modelos locales de los clientes, integrando su conocimiento y preservando su privacidad. En el presente trabajo, se ha optado por el algoritmo FedAvg [18] debido a su eficacia y amplia aceptación en escenarios de aprendizaje federado, el cual promedia los parámetros de los modelos locales para generar un modelo global.

Para evaluar el rendimiento, en las fases de entrenamiento e inferencia en conjuntos cerrados se ha utilizado la precisión (accuracy), debido a la distribución equilibrada de imágenes por clase. En la fase de inferencia en conjuntos abiertos, dado el desequilibrio por la inclusión de 8.000 imágenes de personas desconocidas, se ha hecho uso de la métrica F1 macro, que balancea la evaluación considerando todas las clases por igual.

Para concluir con esta sección, cabe destacar que se ha utilizado el modelo InceptionResnetV1, previamente preentrenado con la base de datos Casia-WebFace [19]. Este modelo fue elegido por su capacidad para generar representaciones faciales robustas y discriminativas, respaldada tanto por los resultados presentados en [20] como por su implementación en Python, que incluye el preentrenamiento con dicha base de datos.

VI. RESULTADOS

A continuación, se exponen los resultados obtenidos en las tres fases del trabajo: entrenamiento, inferencia en conjuntos cerrados e inferencia en conjuntos abiertos.

La fase de entrenamiento de los clientes empleando ambas bases de datos ha resultado satisfactoria, alcanzando valores de precisión de entrenamiento cercanos al 100%. Asimismo, cabe destacar la rápida convergencia del modelo para la clasificación de las 70 personas conocidas, lograda en apenas 5 rondas globales de aprendizaje

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

federado. Esto se debe, en primer lugar, a que el modelo se ha iniciado con pesos pre-entrenados, como se ha mencionado anteriormente, y, en segundo lugar, a que cada cliente ha realizado un entrenamiento local durante 3 rondas por cada ronda global de aprendizaje federado.

Seguidamente, en la fase de inferencia en conjuntos cerrados, se ha evaluado la capacidad del modelo para clasificar a las 70 personas conocidas, utilizando imágenes no vistas por el modelo durante su entrenamiento. El modelo ha demostrado capacidad para clasificar y distinguir dichas personas, alcanzando valores de precisión del 99.6% y 98.83% con las bases de datos MS1M ArcFace y VGGFace2, respectivamente.

Para concluir con los resultados, al implementar la solución propuesta basada en OpenMax y al incluir imágenes de personas desconocidas para el modelo, los valores de F1 macro han experimentado una ligera disminución en comparación con la fase de inferencia en conjuntos abiertos, manteniéndose por encima del 90%, alcanzando exactamente valores de 93.58% y 90.8% para ambas bases de datos. Esto indica que el modelo, generalmente, ha sido capaz de clasificar correctamente a las personas conocidas durante el entrenamiento, así como distinguir a las personas desconocidas.

VII. CONCLUSIONES

Este trabajo presenta un sistema de reconocimiento facial en escenarios de aprendizaje federado, basado en OpenMax para la clasificación en conjuntos abiertos. Los resultados muestran que el modelo es capaz de clasificar con alta precisión a las personas conocidas y distinguir a las desconocidas, manteniendo la privacidad de los datos de los clientes. La propuesta alcanzó valores de precisión superiores al 90%, demostrando su efectividad en escenarios de conjuntos abiertos.

Asimismo, cabe destacar que el presente trabajo resulta relevante para aplicaciones como sistemas distribuidos de control de acceso biométrico, donde se garantiza la privacidad de los usuarios sin sacrificar la precisión. Además, aporta a la investigación en los ámbitos del aprendizaje federado y reconocimiento facial, integrando ambos enfoques con precisión y privacidad. En el futuro, se explorará el uso de técnicas de protección de datos, como encriptación homomórfica y privacidad diferencial.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Ciencia de España a través del proyecto AIBioSurv-Tech (Biosurveillance through Artificial Intelligence (AI) in the post-COVID era: Implications for architecture and cybersecurity), con referencia TED2021-129975B- C22AEI/ AEI/10.13039/501100011033/ Unión Europea NextGenerationEU/PRTR.

REFERENCIAS

[1] Parlamento Europeo y Consejo de la Unión Europea, "Reglamento General de Protección de Datos (RGPD)," 2016.

[2] A. Bendale and T. E. Boult, "Towards Open Set Deep Networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1563–1572.

[3] S. Madakam, T. Uchiya, S. Mark, and Y. Lurie, "Artificial Intelligence, Machine Learning and Deep Learning (Literature: Review and Metrics)," *Asia-Pacific Journal of Management Research and Innovation*, vol. 18, no. 1-2, pp. 7–23, 2022. [Online]. Available: https://doi.org/10.1177/2319510X221136682

[4] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 53, Mar 2021. [Online]. Available: https://doi.org/10.1186/s40537-021-00444-8

[5] Q. Li, Z. Wen, Z. Wu, S. Hu, N. Wang, Y. Li, X. Liu, and B. He, "A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 3347—-3366, Apr. 2023. [Online]. Available: http://dx.doi.org/10.1109/TKDE.2021.3124599

[6] M. Barcina-Blanco, J. L. Lobo, P. Garcia-Bringas, and J. D. Ser, "Managing the unknown: a survey on Open Set Recognition and tangential areas," 2024. [Online]. Available: https://arxiv.org/abs/2312.08785

[7] C. Yang, M. Zhu, Y. Liu, and Y. Yuan, "FedPD: Federated Open Set Recognition with Parameter Disentanglement," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 4859–4868.

[8] M. Zhu, J. Liao, J. Liu, and Y. Yuan, "FedOSS: Federated Open Set Recognition via Inter-Client Discrepancy and Collaboration," *IEEE Transactions on Medical Imaging*, vol. 43, no. 1, pp. 190–202, 2024.

[9] D. Aggarwal, J. Zhou, and A. K. Jain, "FedFace: Collaborative Learning of Face Recognition Model," in *2021 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2021, pp. 1–8.

[10] Y. Niu and W. Deng, "Federated Learning for Face Recognition with Gradient Correction," in *AAAI Conference on Artificial Intelligence*, vol. 36, no. 2, 2022, pp. 1999–2007.

[11] C.-T. Liu, C.-Y. Wang, S.-Y. Chien, and S.-H. Lai, "FedFR: Joint Optimization Federated Framework for Generic and Personalized Face Recognition," in *AAAI Conference on Artificial Intelligence*, vol. 36, no. 2, 2022, pp. 1656–1664.

[12] D. Qiu, X. Lin, K. Wang, X. Chu, and P. Yan, "AdaFedFR: Federated Face Recognition with Adaptive Inter-Class Representation Learning," *arXiv preprint arXiv:2405.13467*, 2024.

[13] H. Kim, H. Choi, and Y. Kwak, "Federated Learning for Face Recognition via Intra-subject Self-supervised Learning," *arXiv preprint arXiv:2407.16289*, 2024.

[14] A. Woubie, E. Solomon, and J. Attieh, "Maintaining Privacy in Face Recognition Using Federated Learning Method," *IEEE Access*, vol. 12, pp. 39 603–39 613, 2024.

[15] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 5962—5979, Oct. 2022. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2021.3087709

[16] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A Dataset for Recognising Faces across Pose and Age," in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 2018, pp. 67–74.

[17] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition," 2016. [Online]. Available: https://arxiv.org/abs/1607.08221

[18] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *20th International Conference on Artificial Intelligence and Statistics*, vol. 54. PMLR, Apr 2017, pp. 1273–1282. [Online]. Available: https://proceedings.mlr.press/v54/mcmahan17a.html

[19] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning Face Representation from Scratch," *CoRR*, 2014. [Online]. Available: http://arxiv.org/abs/1411.7923

[20] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815–823.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Arquitectura para la gestión de offerings en un mercado de activos descentralizado

Víctor González, Juan Ramón Santana, Luis Sánchez, Pablo Sotres, Johnny Choque, Jorge Lanza
Grupo de Ingeniería Telemática,
Departamento Ingeniería de Comunicaciones,
Universidad de Cantabria
Edificio Ingeniería de Telecomunicación, Plaza de la Ciencia, 39005 Santander
{vgonzalez,jrsantana,lsanchez,psotres,jchoque,jlanza}@tlmat.unican.es

Resumen

La aparición de mercados de datos distribuidos plantea nuevos desafíos en cuanto a la interoperabilidad, la gestión segura de activos y la coordinación entre participantes heterogéneos. Este trabajo aborda estos retos mediante el diseño de una arquitectura para la gestión de descripciones estructuradas de los activos digitales ofertados (*offerings*) en un mercado de activos descentralizado. Al estar construido sobre la tecnología DLT de IOTA, se garantiza la seguridad, trazabilidad y descentralización de las operaciones. La arquitectura propuesta se estructura en torno a dos componentes principales: un gestor de activos digitales (*Offering Manager*), responsable del descubrimiento y publicación de *offerings*; y el conector, mediante el cual cada participante (ya sea proveedor o consumidor) interactúa con el ecosistema y con los otros participantes. Este enfoque modular busca facilitar la interoperabilidad entre participantes heterogéneos y garantizar una gestión eficiente y segura de las *offerings*. Se discuten las decisiones de diseño, la implementación actual y los desafíos identificados en su despliegue.

***Palabras Clave*—mercado de datos, DLT, arquitectura, offering, descentralización, interoperabilidad**

I. INTRODUCCIÓN

En los últimos años, la Comisión Europea ha promovido activamente el desarrollo de los Espacios de Datos como una estrategia clave para habilitar el intercambio confiable, soberano e interoperable de datos entre organizaciones [1]. Esta visión impulsa la creación de infraestructuras federadas que permitan la compartición de activos, tanto datos como servicios, sin comprometer la autonomía de los participantes y permitiendo a los proveedores mantener la soberanía de sus activos.

Al mismo tiempo, en un entorno caracterizado por la creciente heterogeneidad de sistemas, formatos y políticas de acceso a la información, resulta fundamental avanzar hacia arquitecturas que no solo aseguren la descentralización y la trazabilidad, sino que también promuevan un marco operativo común. La consolidación de infraestructuras distribuidas para el intercambio de datos y servicios exige un replanteamiento integral de los mecanismos tradicionales de descubrimiento, negociación y gestión de activos, necesarios para hacer realidad un mercado de activos como el que se presenta en este trabajo. En este sentido, el uso de tecnologías de registro distribuido (DLT) se posiciona como un habilitador clave para garantizar la inmutabilidad y auditabilidad de las transacciones, sin renunciar a la flexibilidad requerida por los distintos perfiles de los participantes [2].

La interoperabilidad, entendida no solo como la capacidad de comunicación entre sistemas, sino como una propiedad sistémica que abarca la semántica, la gobernanza y la lógica de negocio compartida, se convierte en un elemento estratégico dentro de ecosistemas descentralizados. Diseñar componentes que puedan operar de manera desacoplada pero coordinada, manteniendo una visión unificada del estado del mercado, constituye un desafío que demanda tanto una arquitectura modular como un enfoque orientado a servicios escalables. La modularidad de las soluciones propuestas permite una adaptación progresiva a las necesidades del dominio de aplicación, sin comprometer los principios de seguridad, soberanía y transparencia que rigen el marco de los Espacios de Datos [3] [4].

Así, el diseño de una arquitectura no puede abordarse únicamente desde una perspectiva técnica, sino que requiere la consideración de las dinámicas de colaboración entre consumidores y proveedores (o, en general, actores) distribuidos, sus incentivos y sus restricciones. El valor de una infraestructura como la que se propone en este artículo reside no sólo en su capacidad de ejecución, sino en su potencial para orquestar flujos de información fiables en tiempo real, integrando capacidades locales

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

con garantías globales. Este enfoque holístico representa un paso hacia la construcción de ecosistemas de activos resilientes, sostenibles y orientados a la creación de valor compartido.

En este contexto se sitúa el proyecto SEDIMARK [5], cuyo objetivo es construir un mercado descentralizado de datos y servicios, basado en DLT y catálogos distribuidos, que permita a los actores intercambiar activos de manera segura, trazable y sin intermediarios centralizados. La arquitectura del mercado de SEDIMARK se estructura en dos niveles principales: una infraestructura base, común a todos los participantes, que incluye la DLT (en este caso, *IOTA stardust*) y un catálogo descentralizado; y un conjunto de herramientas habilitadoras (llamado *toolbox*) modular que cada participante puede desplegar localmente. Esta *toolbox* ofrece una colección de componentes reutilizables, como el conector, el *Offering Manager* (ambos necesarios para interactuar con el resto de participantes), y módulos de procesado como formateo, curado u otros mecanismos de enriquecimiento de datos, así como herramientas de entrenamiento de modelos de inteligencia artificial. Los módulos que cada participante decida utilizar permiten adaptar sus interacciones a su rol específico dentro del ecosistema (proveedor, consumidor o ambos).

Este trabajo se centra en el diseño e implementación de la arquitectura necesaria para la gestión de las descripciones estructuradas de los activos que se desean ofrecer o adquirir (*offerings*). En particular, se aborda el rol del *Offering Manager* y el conector como elementos fundamentales en el descubrimiento, publicación, negociación y transferencia de *offerings*. Además, se describe cómo se aprovecha la DLT para asegurar la trazabilidad, autenticidad e integridad de las operaciones realizadas, contribuyendo así a un mercado transparente y descentralizado. La arquitectura propuesta pone especial énfasis en la interoperabilidad entre participantes heterogéneos, facilitando su integración en un ecosistema distribuido sin necesidad de confianza previa.

Los objetivos principales de este trabajo son: (i) presentar una propuesta de arquitectura que permita la gestión descentralizada de *offerings* en un entorno basado en DLT; (ii) describir los componentes clave involucrados en esta arquitectura y la interacción entre ellos; y (iii) analizar el grado de implementación alcanzado, así como los retos técnicos pendientes en el camino hacia un despliegue completo.

El resto del artículo se organiza como se describe a continuación. En la Sección 2 se introducen los conceptos fundamentales relacionados con los Espacios de Datos, la tecnología DLT y el modelo de mercado descentralizado de activos. La Sección 3 describe en detalle la arquitectura propuesta, incluyendo sus componentes principales, su diseño modular y el estado actual de implementación. Finalmente, la Sección 4 presenta las conclusiones y discute las líneas futuras de trabajo.

II. MARCO TEÓRICO Y TECNOLÓGICO

A la hora de diseñar una arquitectura que permita la compartición de datos de forma independiente, no es suficiente con un enfoque puramente tecnológico; requiere comprender los fundamentos que sustentan esta nueva generación de infraestructuras digitales. La transformación hacia una economía del dato distribuida, en la que múltiples actores colaboran en un entorno de confianza sin intermediarios centrales, se apoya en una combinación de principios organizativos, normativos y tecnológicos que definen las reglas del juego en estos ecosistemas emergentes.

En este contexto, resulta esencial revisar los pilares sobre los que se construye el modelo de mercado propuesto. Esta sección presenta los conceptos clave que fundamentan el diseño de la arquitectura: por un lado, el paradigma de los Espacios de Datos, impulsado por diversas iniciativas europeas, que establece las bases para la soberanía y la interoperabilidad en el intercambio de datos; y por otro, las tecnologías de registro distribuido, que proporcionan una capa de confianza, trazabilidad y descentralización sobre la que se articulan las operaciones del sistema. Finalmente, se introduce el proyecto SEDIMARK, marco en el que se sitúa este trabajo, cuya visión y enfoque técnico sirven como referencia concreta para la implementación de los conceptos previamente descritos.

A. Espacios de Datos

El concepto de Espacios de Datos ha emergido como un marco mediante el cual abordar la creciente complejidad asociada al intercambio, compartición y gobernanza de datos en un mundo cada vez más digital e interconectado. A medida que los sectores productivos intensifican su digitalización y cooperación, la necesidad de gestionar los flujos de datos de forma segura, eficiente y ética se vuelve cada vez más crítica. Los Espacios de Datos buscan proporcionar un entorno estandarizado y confiable en el que organizaciones, individuos y dispositivos puedan intercambiar y acceder a datos y servicios, garantizando al mismo tiempo la privacidad, la soberanía y la propiedad de los mismos [7]. Este modelo innovador permite el desarrollo de ecosistemas digitales conocidos como *data-driven*, favoreciendo la economía del dato y la innovación colaborativa.

Uno de los principales objetivos de los Espacios de Datos es abordar la fragmentación actual existente en el ecosistema de los datos, caracterizado por silos privados de organizaciones que limitan la reutilización, reducen la eficiencia e impiden explotar el potencial transformador de la información. Frente a ello, la misión de los Espacios de Datos se articula en torno a dos ejes principales.

En primer lugar, empoderar a las organizaciones con soberanía sobre sus datos. A diferencia de los modelos centralizados tradicionales, en los que el control sobre los datos suele transferirse a plataformas intermediarias, los Espacios de Datos proponen un modelo federado que permite a las entidades mantener el control sobre con quién, en qué condiciones y con qué propósito se comparten los datos. Este enfoque resulta especialmente relevante en contextos regulados o sensibles, pero se extiende a cualquier dominio donde la protección de los activos sea una prioridad.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En segundo lugar, fomentar la colaboración y la innovación entre actores distribuidos. Al proporcionar un entorno seguro para el intercambio de datos, los Espacios de Datos facilitan el acceso a fuentes de información externas sin comprometer la confidencialidad ni la integridad de los datos propietarios. Esta dinámica habilita nuevos modelos de negocio y acelera la creación de valor conjunto a partir de datos heterogéneos y dispersos.

El sustrato tecnológico de los Espacios de Datos es esencial para materializar estos principios. Se basa en una arquitectura descentralizada que permite a múltiples participantes interactuar sin necesidad de una autoridad central. Las infraestructuras distribuidas garantizan la integridad, trazabilidad y transparencia de los intercambios, y permiten aplicar políticas de acceso y uso previamente negociadas entre proveedores y consumidores.

Otro elemento clave es la definición de estándares, protocolos e interfaces comunes que aseguren la interoperabilidad tanto a nivel técnico como semántico [3] [4]. Esta interoperabilidad es fundamental no solo para el intercambio fluido de datos, sino también para preservar su significado y contexto en entornos heterogéneos. Además, se presta especial atención a las capas de control del sistema, donde se gestionan procesos previos al intercambio efectivo de datos, tales como el descubrimiento de recursos, la negociación de contratos o la definición de mecanismos de gobernanza.

Finalmente, dado el crecimiento exponencial del volumen y complejidad de los datos, los Espacios de Datos también contemplan la integración de tecnologías avanzadas como la inteligencia artificial y el *machine learning* [8]. Estas herramientas permiten extraer valor de grandes conjuntos de datos distribuidos, optimizar procesos de compartición y mejorar la toma de decisiones, manteniendo en todo momento las garantías de seguridad, control y trazabilidad.

B. DLT e IOTA

Las DLT, o tecnologías de registro distribuido, se han consolidado en los últimos años como un pilar clave en el diseño de sistemas digitales que requieren descentralización, inmutabilidad y resistencia frente a manipulaciones o puntos únicos de fallo. Desde una perspectiva técnica, una DLT puede entenderse como un sistema distribuido en el que múltiples nodos mantienen de manera sincronizada un historial compartido y verificable de eventos o transacciones, sin necesidad de una autoridad central que actúe como intermediaria. Para lograrlo, las DLT combinan tres principios fundamentales: replicación, que asegura que todos los participantes conserven una copia idéntica del registro; consenso, mediante el cual los nodos acuerdan colectivamente qué transacciones son válidas y en qué orden deben registrarse; e inmutabilidad, que garantiza que la información una vez almacenada no pueda ser modificada sin el acuerdo de la red. Estos mecanismos convierten a las DLT en una rama particular de los sistemas distribuidos, diseñados para operar de forma fiable en entornos abiertos y potencialmente adversos, ofreciendo tolerancia a fallos bizantinos y transparencia en la gestión de la información.

La adopción de DLT se ha extendido más allá del ámbito financiero para abarcar sectores como la trazabilidad en cadenas de suministro, la gestión de identidad, o más recientemente, el intercambio de datos y servicios en entornos distribuidos. En el contexto de los Espacios de Datos, la DLT tiene el potencial de jugar un papel fundamental como mecanismo de soporte para la confianza, la transparencia y el cumplimiento de las políticas pactadas entre las partes. Su capacidad para registrar de forma segura el ciclo de vida de los activos, desde su publicación hasta su uso, permite reforzar las garantías de auditabilidad, prevenir fraudes y habilitar nuevas formas de gobernanza descentralizada. Además, el uso de DLT contribuye a automatizar interacciones complejas mediante contratos inteligentes, que obligan al cumplimiento de las negociaciones de forma autónoma y verificable.

Sin embargo, no todas las soluciones DLT presentan las mismas propiedades ni el mismo grado de adecuación a escenarios de intercambio masivo de información. Muchos enfoques tradicionales, como las cadenas de bloques (o *blockchains*) basadas en prueba de trabajo u otros mecanismos de consenso, presentan limitaciones en cuanto a escalabilidad, latencia y consumo energético. Estas restricciones dificultan su adopción en arquitecturas orientadas a servicios o en flujos de datos continuos, lo que ha impulsado el surgimiento de nuevas generaciones de DLT diseñadas para superar estas barreras y adaptarse a las necesidades de sistemas distribuidos de alta eficiencia.

En este sentido, la tecnología DLT desarrollada por IOTA [9] representa una alternativa que está perfectamente alineada con los principios de los Espacios de Datos. A diferencia de las cadenas de bloques tradicionales, la DLT de IOTA se basa en una estructura de grafo acíclico dirigido (DAG) denominada *Tangle*. En este modelo, cada nueva transacción debe validar al menos dos previas, lo que elimina la necesidad de mineros, reduce significativamente los costes energéticos y habilita un procesamiento paralelo y asíncrono de transacciones. Esta arquitectura mejora la escalabilidad y disminuye la latencia, lo que convierte a IOTA en una plataforma adecuada para aplicaciones distribuidas que requieren manejar grandes volúmenes de interacciones en tiempo real, como mercados de datos, entornos de IoT o infraestructuras de intercambio descentralizadas.

Además, IOTA no se limita a registrar transacciones básicas, sino que incorpora mecanismos adicionales que enriquecen sus posibilidades. Entre ellos se incluyen contratos inteligentes ligeros, sistemas de identidad digital soberana y canales de comunicación privados, que permiten desplegar lógica compleja de gobernanza sobre los activos compartidos. Estas capacidades amplían el horizonte de uso más allá de la mera compraventa de datos, abarcando aspectos como la trazabilidad, el control de acceso y el cumplimiento de políticas en entornos distribuidos.

En el marco del proyecto SEDIMARK, IOTA se ha adoptado como DLT base sobre la cual construir un mercado de activos totalmente descentralizado. Su integración permite registrar de manera segura los intercambios de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

activos, asegurar la integridad de las transacciones realizadas entre participantes, y dotar al sistema de un núcleo de confianza distribuido que no depende de intermediarios centralizados. La elección de IOTA responde no solo a criterios técnicos de eficiencia, escalabilidad y sostenibilidad, sino también a principios estratégicos de soberanía digital, consolidando su papel como tecnología fundacional para la próxima generación de sistemas distribuidos en el ámbito de los Espacios de Datos.

C. *El proyecto SEDIMARK*

SEDIMARK [5] es un proyecto europeo cuyo objetivo es el desarrollo de un mercado de datos y servicios fundamentado en los principios de descentralización, confianza, inteligencia, interoperabilidad y calidad de los datos. Estos pilares son igualmente esenciales para permitir que los usuarios extraigan el máximo valor posible de los datos, al tiempo que conservan su privacidad y mantienen un control total sobre los activos que comparten. Dentro del ecosistema de SEDIMARK, cualquier usuario que interactúe con el mercado recibe el nombre de Participante, y los elementos que se intercambian en el sistema (ya sean datos, servicios o resultados analíticos) se denominan *assets*. En este documento, hablar de un activo es equivalente a hablar de *asset*.

Los participantes pueden asumir dos roles principales: (i) Proveedores, quienes comparten sus activos con otros actores del mercado y cuentan con herramientas específicas para gestionarlos, definir políticas de acceso y publicarlos como *offerings* (descripciones estructuradas) disponibles para el resto de participantes; y (ii) Consumidores, quienes exploran el mercado, descubren *offerings* relevantes y solicitan el acceso a los activos asociados, ya sea en forma de conjuntos de datos en distintos formatos o como servicios analíticos. Un mismo participante puede desempeñar ambos roles de forma simultánea, actuando tanto como proveedor como consumidor, según sus necesidades.

La visión de SEDIMARK consiste en ofrecer una *toolbox* completa que permita a cualquier participante adquirir y procesar datos, evaluar y mejorar su calidad, construir modelos de aprendizaje automático (*machine learning*) [11], y compartir datos o servicios con otras organizaciones dentro del espacio europeo. El diseño del sistema evita por completo la necesidad de servidores centrales para el almacenamiento o gestión de activos, promoviendo una descentralización total conforme a los principios de los Espacios de Datos. De este modo, los participantes conservan el control absoluto sobre sus activos y pueden definir con precisión quién puede acceder a ellos, en qué condiciones y durante cuánto tiempo.

Cada participante puede descargar e instanciar una *toolbox* personalizada, compuesta por un conjunto modular de componentes *software* que se adaptan dinámicamente al rol y a los requisitos de uso definidos por el usuario. Esta naturaleza dinámica no solo permite una configuración inicial específica, sino también la ampliación posterior de funcionalidades mediante la incorporación de nuevos módulos.

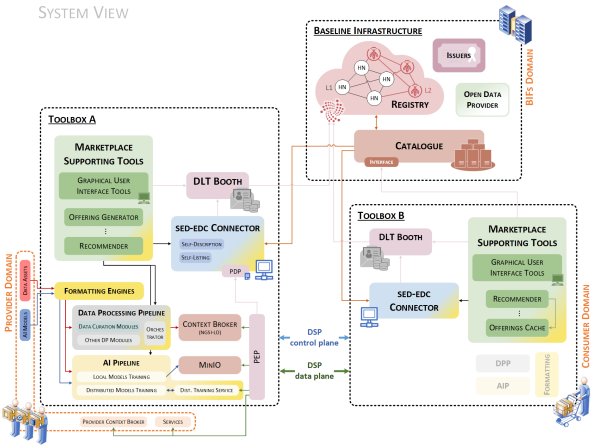


Figura 1. Arquitectura del mercado de activos de SEDIMARK.

La arquitectura de alto nivel de SEDIMARK, que se muestra en la Figura 1, distingue entre dos dominios principales: el dominio de los participantes y el dominio de los facilitadores de la infraestructura base (BIFs). El primero está constituido por las instancias locales de la *toolbox*, que son desplegadas por cada participante de manera autónoma. El segundo, en cambio, proporciona las funcionalidades necesarias para el funcionamiento del mercado a través de un conjunto de servicios distribuidos que constituyen la infraestructura base. Esta división refuerza el carácter descentralizado del sistema, que prescinde de cualquier componente centralizado y se apoya en una red *peer-to-peer* basada en tecnología DLT (en particular, la plataforma IOTA) como núcleo de confianza y coordinación entre los distintos actores del ecosistema.

Entre los componentes identificados en la Figura 1, aquellos directamente involucrados en la gestión de *offerings* (y que constituyen el núcleo del presente trabajo) son fundamentalmente el *Registry*, el *DLT Booth* y el *Connector*. El *Registry*, basado en una DLT compuesta por nodos distribuidos entre múltiples participantes, constituye la base para la generación de confianza en todo el ecosistema del mercado, al garantizar la integridad y trazabilidad de las operaciones. El *DLT Booth* es un componente crítico diseñado para la interacción segura entre los componentes de la *toolbox* y el *Registry*, principalmente a través de los contratos inteligentes desplegados en la DLT, preservando los niveles de seguridad y trazabilidad requeridos. Por su parte, el *Connector* (que en este documento llamamos simplemente conector) habilita el intercambio de información punto a punto entre participantes de forma fiable, y es considerado un componente cuyo despliegue es obligatorio dentro de la *toolbox*. Además de actuar como intermediario en la mayoría de las interacciones del plano de control, su papel resulta fundamental en la creación de un entorno de confianza para el intercambio de datos y servicios. En particular, es el principal responsable de coordinar la gestión y negociación de *offerings*, funciones que se analizan en profundidad en las secciones siguientes de este artículo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Una característica diferenciadora de la arquitectura de SEDIMARK es la incorporación de una infraestructura DLT como componente nuclear para garantizar la confianza mediante el registro de identidades, contratos y transacciones entre participantes. Este enfoque lo distingue de conectores de referencia como el IDS Connector o el Eclipse Dataspace Connector (EDC), en los que la confianza se apoya principalmente en autoridades de certificación, servicios OAuth2 y auditoría mediante *logs* centralizados. Mientras que en el IDS Connector y el EDC la verificación de políticas y el cumplimiento contractual se gestionan en el plano de control de los conectores, en SEDIMARK dichos procesos quedan además anclados en la DLT (IOTA Tangle), lo que garantiza inmutabilidad, no repudio y trazabilidad compartida sin necesidad de intermediarios centralizados. Todo ello, se orienta a permitir una descentralización completa del mercado de datos soportado por SEDIMARK en el que la soberanía de los activos digitales se mantiene, en todo momento, en el proveedor/dueño de dichos activos.

Asimismo, el conector extensible adoptado en SEDIMARK está estrechamente integrado con el DLT Booth, lo que permite verificar identidades descentralizadas (DID) y registrar acuerdos en tiempo real durante el proceso de negociación, como se describe en detalle en la Sección 3. Esta visión lo acerca a las propuestas emergentes de Gaia-X o los conectores de FIWARE, que buscan compatibilidad con credenciales verificables y descentralizadas, pero va un paso más allá al hacer de la infraestructura DLT un componente estructural y no un añadido opcional. Como consecuencia, la arquitectura de SEDIMARK se orienta a mercados de activos digitales en los que la auditabilidad y la soberanía digital son requisitos esenciales, complementando la visión más centrada en interoperabilidad y certificación de los conectores tradicionales.

III. ARQUITECTURA PARA LA GESTIÓN DE OFFERINGS

El diseño de la infraestructura de SEDIMARK se basa en un enfoque modular y descentralizado, orientado a facilitar la gestión segura, interoperable y eficiente de activos en un entorno distribuido. Dentro de este ecosistema, la gestión de *offerings* constituye un componente crítico para el descubrimiento, la negociación y la transferencia de datos y servicios entre participantes.

Esta sección presenta una descripción detallada de la arquitectura, comenzando con una visión general del sistema, y seguida de la caracterización de los componentes funcionales implicados en la gestión de *offerings*.

A. *Visión general*

La arquitectura de gestión de *offerings* en SEDIMARK se apoya en un modelo distribuido que combina componentes locales desplegados por cada participante, que conforman su *toolbox*, con una infraestructura DLT. Tal como se ilustra en la Figura 2, cada participante del mercado tiene su propio conector, elemento indispensable de la *toolbox*, ya que actúa como interfaz principal para

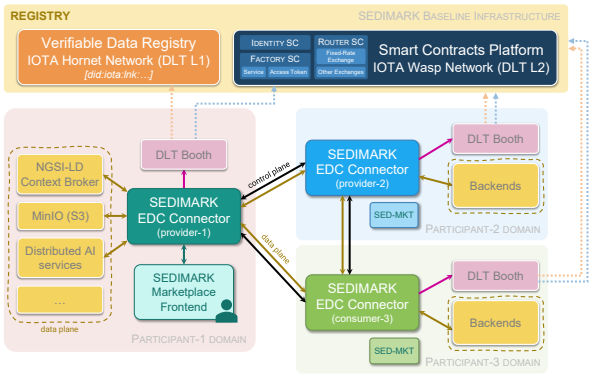


Figura 2. Arquitectura del conector y componentes relacionados.

todas las interacciones de control y comunicación entre participantes.

El conector soporta la integración con distintos servicios de almacenamiento de datos, como *brokers* NGSI-LD o sistemas de almacenamiento distribuido como *MinIO* [12]. Esta flexibilidad fomenta la heterogeneidad entre participantes, permitiendo que proveedor y consumidor puedan utilizar tecnologías diferentes mientras la transferencia permanece transparente para ambos extremos. Además, cada conector incluye como subcomponente un *Offering Manager*, encargado de la gestión de *offerings* en cuanto a su publicación y descubrimiento. En el proceso de publicación por parte de un proveedor, el *Offering Manager* registra formalmente la *offering* en la DLT (en el *Registry*) mediante el almacenamiento de su *hash*, lo que asegura su trazabilidad. Posteriormente, cuando un consumidor realiza el descubrimiento de *offerings*, se comprueba que el *hash* de la *offering* que se está obteniendo del proveedor coincide con aquella que se registró en la DLT. A través de este mecanismo, los datos nunca son publicados (solamente las descripciones) y la transferencia de activos tiene lugar directamente entre los participantes que ya han negociado un contrato basándose en la información disponible en la *offering*.

Para facilitar la interacción con el conector, que tecnológicamente se sustenta sobre una interfaz de programación de aplicaciones (API) basada en el modelo *Representational State Transfer* (REST), cada usuario también cuenta con una interfaz web de usuario (*front-end*), de forma que esta interacción sea independiente del conocimiento tecnológico de dicho usuario.

Por otra parte, el *DLT Booth* provee un servicio de gestión de credenciales, que permite al conector verificar la identidad de otros participantes en la DLT y realizar operaciones utilizando contratos inteligentes desplegados en la capa 2 (L2) de dicha DLT, que se basa en IOTA Stardust. De este modo, la gestión de identidades, los contratos inteligentes y el registro verificable de operaciones se realiza de forma segura y trazable.

Así, la combinación de conectores distribuidos, infraestructura DLT y servicios de almacenamiento federados permite a SEDIMARK construir un mercado completa-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

mente descentralizado, donde cada participante conserva el control sobre sus activos, al mismo tiempo que se garantiza la confianza entre partes mediante mecanismos criptográficamente robustos y verificables.

Tras esta presentación de la arquitectura general del sistema y de las funcionalidades básicas de sus componentes, en las siguientes secciones se detallan los principales componentes que permiten la gestión descentralizada de *offerings* en SEDIMARK. Se comienza analizando la interacción de los participantes con la DLT, clave para garantizar la trazabilidad y la confianza en el ecosistema. A continuación, se describen en profundidad el *Offering Manager* y el conector, responsables de orquestar el descubrimiento, la negociación y la transferencia de activos.

B. Interacción con la DLT

La infraestructura DLT desempeña un papel central en SEDIMARK como núcleo de confianza para el registro de las operaciones entre participantes. A través de ella se gestionan las identidades descentralizadas (DID), se publican las transacciones realizadas en el mercado y se ejecutan contratos inteligentes que regulan las condiciones de acceso y uso de los activos. Para facilitar esta interacción, se introduce el componente *DLT Booth*, que actúa como intermediario entre el conector del participante y la DLT subyacente.

El *DLT Booth* expone un API basado en servicios web que permite interactuar con contratos inteligentes desplegados sobre una *Ethereum Virtual Machine* (EVM), como es el caso de la L2 de IOTA Stardust utilizada en SEDIMARK, con el fin de gestionar identidades autosoberanas (Self-Sovereign Identity, SSI). En este ámbito, el *DLT Booth* permite crear y publicar documentos de identidad descentralizada (DID), solicitar la emisión o revocación de credenciales verificables (VC) por parte de autoridades emisoras, y generar presentaciones verificables (VP) para autenticarse ante otros participantes del ecosistema. Estas capacidades son esenciales durante las fases iniciales de incorporación al mercado (*onboarding*) y para la verificación de identidades durante procesos de negociación.

En cuanto a las operaciones sobre la EVM, el *DLT Booth* permite realizar firmas de mensajes en formato EIP-191, así como ejecutar transacciones en cadenas configuradas, lo que habilita la integración con contratos inteligentes que gestionan el acceso a ciertos tipos de activos.

Desde el punto de vista de la arquitectura, el *DLT Booth* es un componente clave para la integración del conector en SEDIMARK, que hace uso de los servicios ofrecidos por su API para comunicarse con la DLT. Durante una negociación entre un proveedor y un consumidor, el conector utiliza el *DLT Booth* para verificar la identidad del interlocutor, comprobando credenciales registradas en la DLT, y para registrar el resultado de la transacción final una vez acordado el intercambio del activo. De este modo, se garantiza no solo la trazabilidad de las operaciones, sino también su no repudio.

En la Figura 2 se puede observar que todas las flechas con transparencia, que representan una interacción con el *Registry*, son operaciones realizadas a través del *DLT Booth*. En términos generales, este componente permite desacoplar la integración con la DLT del resto de la arquitectura, facilitando en gran medida su implementación y su modularidad.

C. Offering Manager

El *Offering Manager* es el componente responsable de la gestión local de *offerings* en la *toolbox* de cada participante. Su función principal es proporcionar una interfaz para la creación, consulta, actualización y eliminación (CRUD) de *offerings*. Estas *offerings* son descripciones estructuradas (en formato JSON-LD) de los *assets* que un proveedor desea publicar en el mercado, y que posteriormente podrán ser descubiertos, negociados y adquiridos por otros participantes. En la Figura 2, el *Offering Manager* se considera parte del conector y por ello no se representa como un bloque individual.

El *Offering Manager* también expone un API REST que soporta operaciones CRUD. Cada *offering* se identifica mediante un *Universally Unique IDentifier* (UUID) generado automáticamente, lo que evita colisiones y garantiza unicidad en entornos distribuidos. A través de esta API, el participante (en este caso, el proveedor) puede registrar nuevas *offerings*, consultar las ya existentes, modificarlas o eliminarlas. Además, el componente ofrece una interfaz que permite listar todas las *offerings* disponibles localmente en la instancia de la *toolbox*.

Más allá de su funcionalidad como gestor de *offerings*, el *Offering Manager* también está integrado con el conector para llevar a cabo los procesos de creación y publicación de *offerings*, y que éstas estén alineadas con los subsiguientes procesos de negociación e intercambio de los *assets* contenidos en dichas *offerings*. De esta manera, cuando se recibe una petición de creación, el *Offering Manager* inicia una serie de llamadas al conector para llevar a cabo operaciones esenciales: la creación del *asset* correspondiente, la definición de las políticas de acceso asociadas, y la generación del contrato necesario para regular las futuras negociaciones. Esta interacción asegura que cada *offering* esté vinculada a un *asset* registrado formalmente en el sistema, con condiciones de uso definidas mediante políticas de acceso y uso descritas a través del estándar *Open Digital Rights Language* (ODRL).

Del mismo modo, cuando se consulta una *offering*, parte de la información se obtiene directamente del almacenamiento interno del *Offering Manager*, mientras que el resto se solicita dinámicamente al conector en función del *asset* previamente creado. Este modelo híbrido permite mantener una capa local de control y eficiencia, y al mismo tiempo asegurar la consistencia a nivel de sistema.

En resumen, el *Offering Manager* actúa como punto de entrada para la publicación de activos y sus respectivas *offerings* en el mercado. Su diseño permite a los proveedores gestionar sus *offerings* de forma sencilla y dinámica, mientras se asegura la correcta vinculación con los *assets* y las políticas de acceso definidas a través del conector. Esta separación de responsabilidades favorece la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

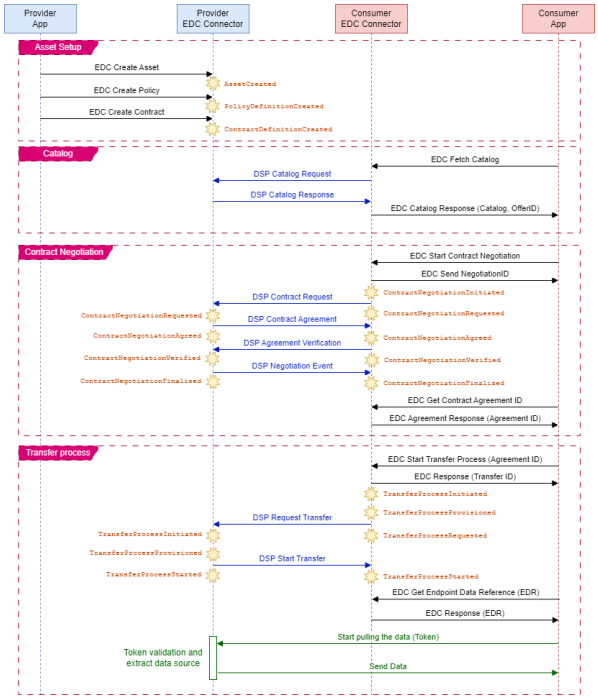


Figura 3. Eventos internos en el conector.

modularidad, con la que cada componente cumple un rol específico dentro del ciclo de vida de las *offerings*.

D. Conector

El conector es el componente central en la arquitectura de SEDIMARK para la interacción entre participantes. Actúa como interfaz única entre las *toolbox* distribuidas, asegurando una comunicación segura y estandarizada. Todas las operaciones de descubrimiento de datos, negociación de contratos y transferencia de activos entre participantes pasan, necesariamente, por este componente.

En función del rol del participante dentro del mercado, el conector puede adoptar diferentes configuraciones. Por ejemplo, un consumidor de datos no necesita exponer ninguna interfaz, mientras que un proveedor debe habilitar al menos dos:

- La interfaz dedicada a la comunicación interna entre conectores, donde se utiliza el protocolo *Data Space Protocol (DSP)* sin requerir intervención del usuario.
- Una interfaz pública, que permite la obtención de activos ya negociados. Internamente, funciona como un *proxy* seguro hacia el almacenamiento privado del proveedor.

Fundamentalmente, el conector se organiza como una serie de máquinas de estados que controlan los distintos procesos implicados en la publicación y consumo de *assets*. La Figura 3 ilustra los flujos de eventos generados durante el ciclo de vida completo de un *asset*, desde la creación del mismo y sus políticas asociadas hasta la negociación del contrato y la transferencia del *asset*. Cada evento representa una transición entre estados en el conector.

Una de las principales características del conector en SEDIMARK es su extensibilidad. La implementación de referencia seguida en el proyecto es la promovida por la *Eclipse Foundation*, que ofrece un conector [13] adaptable mediante la inclusión de *plugins* en Java. Esto permite ajustar y ampliar sus funcionalidades sin alterar el núcleo del sistema. Aprovechando esta capacidad, en el ámbito de SEDIMARK se ha desarrolado un *plugin* que permite interrumpir y modificar el flujo predeterminado para ejecutar lóigca adicional ante ciertos eventos. Este mecanismo ha sido aprovechado para integrar operaciones con la infraestructura DLT mediante el *DLT Booth*, adaptando la funcionalidad del conector a las necesidades particulares de SEDIMARK. Por ejemplo, durante la negociación de un contrato, es posible verificar en tiempo real la identidad del participante remoto y registrar de forma verificable el acuerdo alcanzado en la DLT. De este modo, el conector no solo orquesta el flujo de mensajes, sino que también garantiza la confianza y la trazabilidad de las operaciones.

Además de su papel como coordinador del ciclo de vida de los *assets*, el conector en SEDIMARK está diseñado con un alto nivel de interoperabilidad. Esto le permite gestionar e intercambiar activos residentes en distintas tecnologías sin requerir cambios personalizados en su funcionamiento. Por ejemplo, es capaz de exponer y consumir datos a través de interfaces REST, incluyendo estándares semánticos como *Next Generation Service Interface with Linked Data* (NGSI-LD), así como integrarse con flujos más complejos como los que requieren modelos de aprendizaje distribuido o federado (*federated learning*, FL). Asimismo, puede actuar como intermediario para el intercambio de activos almacenados en infraestructuras cloud, como *Amazon Web Services* (AWS) S3. De esta forma, se facilita la federación de recursos heterogéneos, manteniendo la coherencia del proceso de publicación, negociación y acceso a los datos, independientemente del tipo de tecnología subyacente.

E. Despliegue y validación

El diseño modular de la arquitectura SEDIMARK facilita su despliegue en múltiples escenarios, desde configuraciones mínimas hasta entornos de prueba más complejos. Para simplificar el proceso de uso e incrementar su reusabilidad en la comunidad, se han desarrollado distintos ficheros *docker-compose* que permiten desplegar:

- Un único conector, útil para pruebas locales o integración básica.
- Dos conectores en la misma red, configurados para simular la interacción entre participantes del mercado.
- Una infraestructura más completa que incluye, además del conector, componentes como el *DLT Booth* y el *Offering Manager*.

Estas configuraciones permiten evaluar el comportamiento del sistema en distintos contextos sin requerir una instalación manual compleja. Aunque los repositorios donde se alojan estas configuraciones son actualmente

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

privados [14], han servido como base para validar el funcionamiento completo del sistema.

La validación se ha realizado mediante el despliegue de los componentes en varias máquinas virtuales independientes con Ubuntu 22.04.5 LTS, configuradas con 2 núcleos de CPU a 2.40 GHz y 16 GB de memoria RAM. Este entorno ha permitido comprobar la interoperabilidad de los componentes, la persistencia de las *offerings* y los *assets*, y la correcta interacción con la infraestructura DLT en condiciones realistas.

Durante las pruebas descritas, se han monitorizado también algunos indicadores básicos de consumo de recursos. Cada instancia de conector desplegada en una máquina virtual con las especificaciones descritas anteriormente mostró un uso medio de 411 MB de RAM y una carga de CPU del 5 % en escenarios de interacción entre dos participantes. El DLT Booth presentó un consumo prácticamente despreciable (hasta dos órdenes de magnitud inferior al del conector), y la base de datos Postgres utilizada para la persistencia de los datos tuvo un consumo medio de 48 MB de RAM y un 3 % de uso de CPU. Estos valores confirman la viabilidad del enfoque incluso en entornos con recursos moderados. No obstante, un análisis exhaustivo de métricas de rendimiento que incluya latencia extremo a extremo, escalabilidad en presencia de múltiples participantes y consumo energético asociado al uso de la infraestructura DLT se considera fuera del alcance de este trabajo. Este análisis se plantea como línea prioritaria de investigación futura. Asimismo, se plantea un despliegue en dispositivos de bajas capacidades de cómputo (como Raspberry Pi) para evaluar la eficiencia y robustez de los componentes en entornos de edge computing y bajo consumo.

IV. CONCLUSIONES

Este trabajo ha presentado en detalle la arquitectura de gestión de *offerings* desarrollada en el contexto del proyecto SEDIMARK, destacando su enfoque distribuido y modular, basado en tecnologías abiertas y componentes interoperables. La solución propuesta permite a los participantes del mercado mantener control sobre sus activos, al tiempo que garantiza la confianza entre las partes mediante mecanismos descentralizados como identidades autosoberanas y contratos inteligentes sobre infraestructura DLT.

Entre las contribuciones principales de este trabajo se encuentra el diseño e implementación del *Offering Manager*, un componente ligero y extensible que permite a cada participante gestionar su propio catálogo de ofertas de forma local, dinámica y autónoma. Además, se ha detallado el papel central del conector, la pieza clave que orquesta tanto la publicación, como el descubrimiento, negociación y transferencia de *assets* entre entidades. Su arquitectura interna, basada en máquinas de estados y enriquecida mediante un sistema de *plugins*, ha permitido integrar capacidades avanzadas como la verificación de identidades y el registro de acuerdos en la DLT de forma flexible.

Finalmente, se ha verificado la capacidad del sistema para operar en entornos reales, gracias al desarrollo de escenarios de despliegue reproducibles mediante contenedores y a la validación funcional en máquinas virtuales independientes. Todo ello demuestra la viabilidad técnica de la solución y sienta las bases para su adopción en entornos reales de intercambio de activos como el mercado de datos y servicios de SEDIMARK.

AGRADECIMIENTOS

Este trabajo ha sido financiado principalmente por el proyecto SEDIMARK, subvencionado por la Comisión Europea en el marco del programa Horizonte Europa (Grant Agreement nº 101070074). También ha contado con el apoyo del proyecto THROTTLE (TED2021-131988B-I00), financiado por el Ministerio de Ciencia, Innovación y Universidades (MICIU) / Agencia Estatal de Investigación (AEI) y por la Unión Europea a través de los fondos NextGenerationEU / PRTR.

REFERENCIAS

[1] Comisión Europea, Dirección General de Redes de Comunicación, Contenido y Tecnologías, "A European strategy for data", COM/2020/66, [Online]
[2] S. Girish Savadatti, S. Krishnamoorthy and R. Delhibabu, "Survey of Distributed Ledger Technology (DLT) for Secure and Scalable Computing", in IEEE Access, vol. 13, pp. 8393-8415, 2025, doi: 10.1109/ACCESS.2025.3528211.
[3] B. Otto, S. Steinbuss, A. Teuschery and S. Lohmann, "IDS Reference Architecture Model". Zenodo, abr. 01, 2019. doi: 10.5281/zenodo.5105529.
[4] Gaia-X European Association for Data and Cloud AISBL, "Gaia-X Architecture Document", [Online]
[5] SEDIMARK Consortium, "SEDIMARK Project website", [Online]
[6] IOTA Foundation, "IOTA Stardust Upgrade", [Online]
[7] E. Curry, S. Scerri and T. Tuikka, "Data Spaces", Springer International Publishing, Cham, 2022. doi: 10.1007/978-3-030-98636-0.
[8] Z. Boukhers, C. Lange, and O. Beyan, "Enhancing Data Space Semantic Interoperability through Machine Learning: a Visionary Perspective", in Companion Proceedings of the ACM Web Conference 2023. doi: 10.1145/3543873.3587658.
[9] W. F. Silvano and R. Marcelino, "Iota Tangle: A cryptocurrency to communicate Internet-of-Things data", Future Generation Computer Systems, Volume 112, 2020, Pages 307-319, doi: 10.1016/j.future.2020.05.047.
[10] S. Wan, H. Lin, W. Gan, J. Chen and P. S. Yu, "Web3: The Next Internet Revolution", IEEE Internet of Things Journal, vol. 11, no. 21, pp. 34811-34825, 1 Nov.1, 2024, doi: 10.1109/JIOT.2024.3432116.
[11] C. Mazzocca, N. Romandini, R. Montanari and P. Bellavista, "Enabling Federated Learning at the Edge through the IOTA Tangle", Future Generation Computer Systems, Volume 152, 2024, Pages 17-29, doi: 10.1016/j.future.2023.10.014.
[12] MinIO, "MinIO website", [Online]
[13] Eclipse Foundation, "Eclipse Data Component Connector", GitHub, [Online]
[14] V. González, J. R. Santana and P. Sotres, "Toolbox UC in Github", GitHub, [Online]

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Enfoque práctico simplificado de integraciones IT-OT seguras en entornos industriales

Álvaro Martínez Martínez, Jesús Algar Fernández, Juan Pedro Muñoz Gea, Alejandro Martínez Sala, Maria-Dolores Cano
Departamento de Tecnologías de la Información y las Comunicaciones
Universidad Politécnica de Cartagena
TRUST Lab, Edificio ELDI, O3-10, 30202, Cartagena
alvaro.mmartinez@upct.es, jesus.algar@upct.es, juanp.gea@upct.es, alejandros.martinez@upct.es, mdolores.cano@upct.es

La creciente digitalización del sector industrial exige integrar de forma segura las tecnologías de la información (*Information Technologies*, IT) con las tecnologías operativas (*Operational Technologies*, OT), especialmente en entornos de infraestructuras críticas como la gestión del agua. Este artículo presenta una arquitectura simplificada de integración IT-OT que cumple simultáneamente el Esquema Nacional de Seguridad y la familia de normas IEC-62443, evitando conexiones directas y privilegiando enlaces controlados (*conduits*). La propuesta se implementa y valida en un laboratorio virtual basado en GNS3 con componentes reales desplegados en contenedores: sensores IIoT (caudal y presión), *broker* MQTT Eclipse-Mosquitto, pasarela Node-RED, *firewall* OPNSense, base de datos InfluxDB y plataforma de visualización Grafana. Los resultados demuestran la viabilidad de transmitir datos industriales en tiempo real hacia la red corporativa sin exponer la zona de control: los mensajes MQTT atraviesan la DMZ bajo reglas de *firewall* restrictivas hasta almacenarse en InfluxDB, desde donde se visualizan de forma remota en Grafana. El trabajo aporta (i) un procedimiento paso a paso para reproducir la integración con costes reducidos, (ii) un mapeo explícito entre los controles técnicos implementados y los requisitos ENS/IEC-62443, y (iii) una discusión de buenas prácticas y limitaciones para su escalado a planta real.

***Palabras Clave*- IoT, IIoT, Ciberseguridad Industrial, Integración IT-OT, ENS, GNS3.**

I. Introducción

La digitalización industrial se ha convertido en un eje estratégico para mejorar la eficiencia y la toma de decisiones en sectores críticos como el del agua. Sin embargo, la interconexión de sistemas tradicionalmente aislados, conocidos como tecnologías operativas (*Operational Technologies*, OT), con las tecnologías de la información (*Information Technologies*, IT) introduce nuevos vectores de amenaza que pueden comprometer tanto la producción como la seguridad de las personas [1]. Garantizar una convergencia segura entre ambos mundos exige abordar retos de segmentación de redes, autenticación, monitorización y mantenimiento, sin menoscabar la disponibilidad que caracteriza a los entornos industriales.

En España, el Esquema Nacional de Seguridad (ENS) [2]y la familia de normas IEC-62443 [3]se han consolidado como marcos de referencia para proteger infraestructuras críticas. El ENS establece los principios y requisitos mínimos para los sistemas que manejan información de interés público, mientras que IEC-62443 profundiza en la protección de sistemas de control industrial mediante un enfoque por capas y conductos (*conduits*) seguros [4] Estudios recientes abogan por incorporar la verificación continua de requisitos IEC-62443 en pipelines DevSecOps o replantear el diseño organizativo mediante gemelos de procesos [5]. Cumplir simultáneamente ambos estándares supone un reto adicional para muchas pymes industriales, que a menudo carecen de metodologías prácticas y replicables.

Este trabajo aborda esa brecha mediante una arquitectura simplificada de integración IT-OT validada en un laboratorio virtual GNS3. El caso de uso se centra en una planta de gestión de aguas y emplea componentes reales desplegados en contenedores: sensores *Industrial Internet of Things* (IIoT), *broker* MQTT (Eclipse-Mosquitto), pasarela o *gateway* IoT Node-RED, cortafuegos o *firewall* OPNSense, base de datos InfluxDB y plataforma Grafana; todos ellos interconectados a través de una zona desmilitarizada (*Demilitarized Zone*, DMZ) industrial (véase Figura 1). La propuesta se ha concebido para ser económica, reproducible y alineada con buenas prácticas de la norma IEC-62443. Las contribuciones principales del artículo son:

1) Diseño de una arquitectura de red en tres zonas (OT, DMZ, IT) que aúna los requisitos esenciales de ENS e IEC-62443 sin necesidad de hardware especializado.
2) Procedimiento paso a paso para desplegar y configurar cada componente con imágenes Docker, facilitando su adopción y enseñanza.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

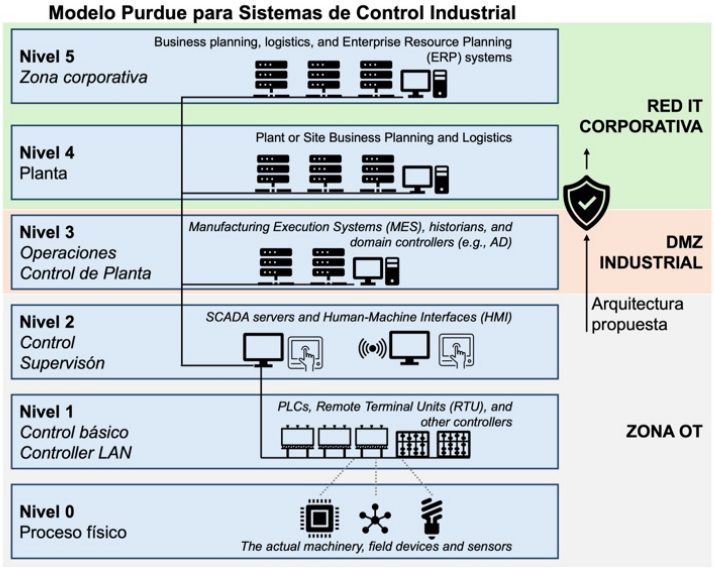


Fig. 1. Modelo Purdue para sistemas de control industrial OT conectado a red IT, el escudo señala la integración segura entre la red OT y la red corporativa IT a través de la DMZ industrial.

3) Mapa de cumplimiento que relaciona los controles técnicos implementados con los requisitos ENS e IEC-62443, proporcionando una guía práctica de autoevaluación.
4) Resultados de validación basados en indicadores ya disponibles (latencia MQTT→DB, carga CPU *gateway* IoT, reglas de *firewall*), que demuestran la viabilidad de la solución en tiempo real.
5) Discusión de limitaciones y escalabilidad, identificando mejoras futuras como la incorporación de Zero-Trust.

El resto del artículo se organiza como sigue. La Sección II revisa el estado del arte y los marcos normativos. La Sección III describe detalladamente la arquitectura propuesta. La Sección IV presenta la metodología de validación y los resultados obtenidos, analizando los hallazgos, limitaciones y posibles extensiones. Por último, se resumen las conclusiones y las líneas de trabajo futuro.

II. Estado del arte y marco normativo

A. Trabajos relacionados

Durante los últimos años, la investigación sobre ciberseguridad industrial ha evolucionado desde la segmentación de redes clásica hacia modelos de gobernanza continua y validación automatizada. Bhole *et al*. analizan en [6] 46 soluciones de seguridad comercial y académica. Los autores concluyen que, aunque las herramientas de supervisión en tiempo real han madurado, persisten lagunas en la detección de anomalías OT y en la medición objetiva de riesgo residual. En paralelo, el trabajo desarrollado en [4] demuestra que los requisitos de la serie IEC-62443 pueden mapearse al ecosistema IIoT mediante una tabla exhaustiva de trazabilidad tecnología-estándar, pero advierten de la ausencia de métricas de coste y de estudios comparativos entre marcos nacionales y la norma internacional

El problema de incorporar ese cumplimiento normativo al ciclo de vida del software se aborda en el trabajo de Göttel *et al*. [7] que propone integrar la verificación de IEC-62443 en los pipelines CI/CD. Su análisis revela que la cobertura automatizable es aún parcial y depende en gran medida de protocolos IT, no OT. En conjunto, la literatura evidencia tres carencias. En primer lugar, la falta de laboratorios reproducibles que faciliten la transferencia tecnológica [8]. En segundo lugar, hay pocos trabajos que combinen marcos internacionales y nacionales. Por último, sería relevante poder disponer de indicadores operativos sencillos que permitan cuantificar el impacto de la integración IT-OT en pymes industriales.

B. Marcos normativos aplicables

El ENS, actualizado por el Real Decreto 311/2022 [9], fija los principios de protección (confidencialidad, integridad, disponibilidad, autenticidad y trazabilidad) que deben cumplir los sistemas de las administraciones públicas españolas, así como un régimen de auditoría cada dos años para verificar su cumplimiento.

Complementariamente, la serie IEC-62443 [3] establece controles técnicos y organizativos específicos para los sistemas de automatización y control industrial (*Industrial Automation and Control Systems*, IACS). Los documentos 62443-3-3 (requisitos de sistema) y 62443-4-2 (requisitos de componente) promueven la filosofía conocida como zonas y conductos (*zones and conduits*) y se articulan de forma natural con el modelo Purdue, favoreciendo la creación de DMZ industriales y la defensa en profundidad. Aunque ambos marcos normativos comparten fundamentos, la literatura revisada muestra pocos trabajos que los integren de forma operativa y práctica.

De la revisión anterior se desprende que persisten vacíos en la validación cuantitativa de arquitecturas IT-OT en la convergencia de marcos regulatorios internacionales y nacionales. Para cubrirlos, este trabajo presenta una arquitectura simplificada y reproducible que (i) mapea de forma explícita los controles del ENS y los requisitos IEC-62443-3-3, (ii) se despliega íntegramente en un laboratorio virtual GNS3 con componentes *open-source* de bajo coste y (iii) aporta métricas objetivas (latencia de mensajes MQTT,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

carga de CPU en el *gateway* IoT y número de reglas del cortafuegos) que es aplicable y de base para la adecuación y mejora de infraestructuras críticas en general y del sector del agua en particular.

III. Arquitectura Propuesta

A. *Visión global*

La solución se organiza en tres zonas lógicamente segregadas: OT, DMZ industrial e IT corporativa. Estas reproducen el modelo Purdue y materializan la filosofía de zonas y conductos antes mencionada de IEC-62443. Cada zona reside en una VLAN distinta y está interconectada por un cortafuegos OPNSense que aplica las políticas de filtrado y NAT exigidas por el ENS y la serie IEC-62443-3-3. La Figura 2 ofrece una panorámica compacta de esta topología, resaltando la ubicación exacta del conducto seguro que enlaza la DMZ con la red corporativa.

B. *Diseño de red y flujo de datos*

El entramado de comunicaciones se muestra en las Figuras 2 y 3, donde se distinguen las subredes 10.10.1.0/24 (OT), 10.20.1.0/24 (DMZ) y red IT corporativa (DHCP). En cada área de red se utiliza un switch administrable (simulado con imágenes Cisco IOU en GNS3), esenciales para segmentar correctamente la red. La Tabla I resume la asignación de VLAN y la función de cada área, mientras que la Figura 4 detalla el flujo desde los sensores industriales hasta la base de datos en IT (sensor -> *broker* -> pasarela -> base de datos) estableciendo una arquitectura de datos tipo MING [10], subrayando que los únicos puertos y protocolos abiertos son MQTT/TLS, HTTPS. Así se garantiza que ningún tráfico OT acceda directamente a red IT, cumpliendo la separación de dominios establecida en el ENS.

C. *Servicios desplegados*

El diagrama de servicios (Figura 4) ilustra la pila software: (1) sensores IIoT simulados en Python/*paho*-MQTT, (2) *broker* Eclipse-Mosquitto con TLS 1.2, (3) Node-RED como pasarela con ACLs que añade marcas temporales ISO-8601 y reenvía los mensajes a (4) InfluxDB y (5) Grafana para la visualización.

El *broker* MQTT es el servicio de mensajería *publish-subscriber* central de la arquitectura. Se escogió el protocolo MQTT debido a su ligereza y rapidez, esencial en sistemas industriales y dispositivos IoT. El *broker* Mosquitto, desplegado en Docker, actúa como mediador de datos: los sensores industriales (clientes MQTT) publican sus mediciones (publish) en topics concretos y los sistemas consumidores o subscribers (como el gateway IoT Node-RED) se suscriben para recibir esos datos en tiempo real.

El *gateway* IoT con Node-RED es una herramienta de programación visual muy empleada en IIoT por su rapidez para crear flujos de integración sin necesidad de desarrollar código complejo. En esta arquitectura, actuará como puerta de enlace IIoT. Recibirá los datos proporcionados por los sistemas IIoT mediante un flujo que se suscribe a los *topics* de los sensores, para procesarlos, y redistribuirlos a IT en formato JSON (véase la figura 5), estableciendo un túnel de comuni-

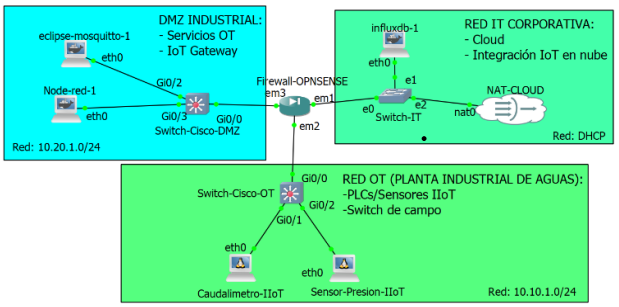


Fig. 2. Arquitectura simplificada en maqueta GNS3 red OT-DMZ-IT.

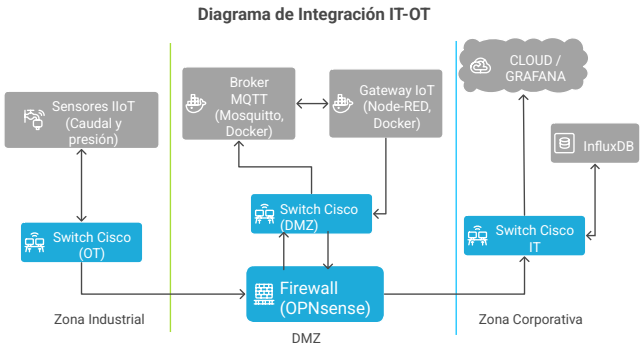


Fig. 3. Esquema de red y comunicaciones de integración IT-OT.

Tabla I
VLANS SEGÚN ÁREA

VLAN	Áreas de red	Direcciones IP
10	Área industrial	10.10.1.0/24
20	Área DMZ	10.20.1.0/24
30	Área corporativa	DHCP (192.168.163.0/24)

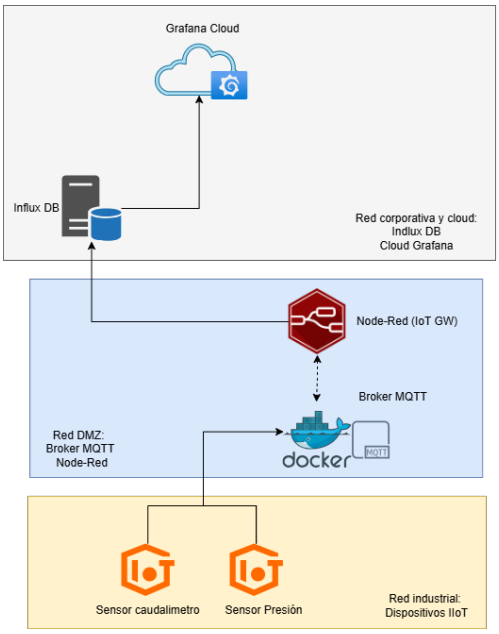
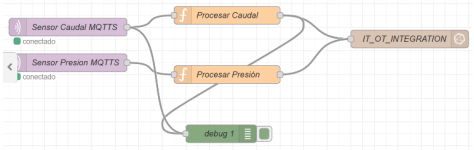
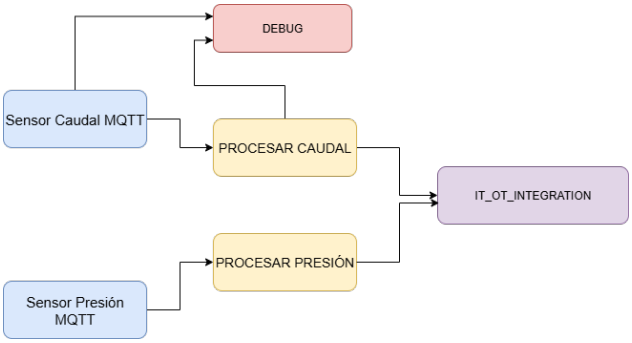


Fig. 4. Diagrama de servicios de Integración IT-OT.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) Screenshot desde Node Red



(b) Representación del flujo de datos mostrado en (a)

Fig. 5. Captura flujo de datos Gateway IIoT en Node Red.

cación indirecta entre la zona industrial OT y la corporativa IT. Como se observa en las figuras 2 y 3, el módulo Node-Red ha sido desplegado en un *docker* conectado a la DMZ industrial [11]. Todos los componentes corren en contenedores *docker* gestionados manualmente en GNS3, de modo que la plataforma puede reproducirse sin hardware especializado y escalarse mediante *docker-compose* en los trabajos futuros.

D. Cumplimiento normativo

La arquitectura alinea los cinco principios del ENS (confidencialidad, integridad, disponibilidad, autenticidad y trazabilidad) con los requisitos de seguridad (Security Requirements, SC) de IEC-62443-3-3. El control de acceso (SR 3.1) se implementa mediante VLAN y reglas OPNSense, dando cumplimiento directo a las medidas MP.COM.4 (Separación de flujos de información en la red), lo que en entornos OT garantiza la aislación de segmentos de control y evita la interferencia en bucles de control en tiempo real, y OP.PL.2 (Arquitectura de seguridad), que en OT simplifica el diseño de red minimizando puntos de fallo y facilitando la gestión de cambios. La integridad de las comunicaciones (SR 5.3) se asegura con TLS, protegiendo los datos de telemetría y comandos de control industrial frente a manipulaciones, y la detección de eventos (SR 7.1) se materializa centralizando logs en Node-RED, alineándose con la medida OP.EXP.8 (Registro de actividad) y permite correlacionar sucesos operativos con alarmas de planta para una respuesta rápida ante incidentes. El mapa-calor incluido en la Figura 6 demuestra que se cubre más del 50 % de los requisitos pertinentes para un entorno de criticidad media, alineando ambas normativas para una resolución más eficiente.

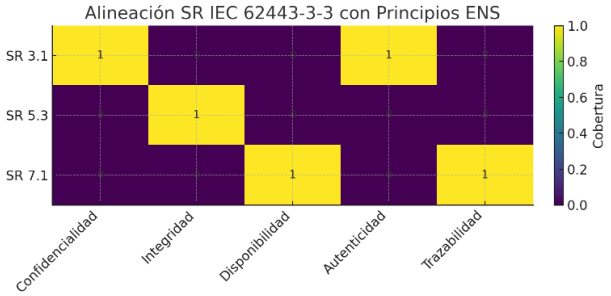


Fig. 6. Mapa de calor de cumplimiento normativo ENS e IEC-62443-3-3.

Si bien el ENS fue diseñado originalmente para entornos IT, su aplicación en entornos OT requiere ciertas consideraciones específicas. Por ejemplo, controles como MP.SI.1 (Seguridad en redes) o MP.SG.1 (Protección frente a código dañino) deben adaptarse a sistemas con alta sensibilidad a la latencia y baja tolerancia a cambios no planificados. En OT, aplicar medidas como el análisis de malware o inspección profunda de paquetes puede interferir en procesos industriales si no se ejecutan en zonas intermedias (como la DMZ) o en tiempos programados. Asimismo, controles como OP.ACC.1 (Gestión de identidades y control de accesos) implican integrar mecanismos de autenticación en dispositivos que a menudo carecen de capacidades avanzadas de gestión de usuarios, lo que obliga a soluciones compensatorias como segmentación lógica y control de acceso en red. Esta adaptación de controles ENS al contexto OT resulta clave para preservar la disponibilidad de los procesos y, al mismo tiempo, elevar el nivel de madurez en ciberseguridad industrial.

IV. Metodología de Validación y Resultados

La validación se llevó a cabo en el laboratorio virtual GNS3 descrito en la Sección III. Los ensayos se realizaron con cien instancias de sensores IIoT, cada una publicando una medición por segundo, lo que genera un caudal agregado de 100 mensajes por segundo hacia el *broker* MQTT, ritmo representativo de una planta de tamaño medio. La latencia extremo a extremo se capturó con la marca temporal que Node-RED agrega a cada mensaje y se midió de nuevo tras su escritura en InfluxDB. La carga de CPU de la pasarela se midió con *docker stats* y la eficacia de filtrado con los registros de OPNSense. Cada ensayo duró 30 min y se repitió tres veces para descartar valores atípicos.

Las Figuras 7 y 8 muestran la configuración de interfaces de OPNSense y las reglas de firewall, donde se aprecia la separación física de las VLAN 10, 20 y 30. Por su parte, la Figura 9 ilustra un sensor IIoT en funcionamiento y confirma la publicación periódica de datos mediante MQTT/TLS.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

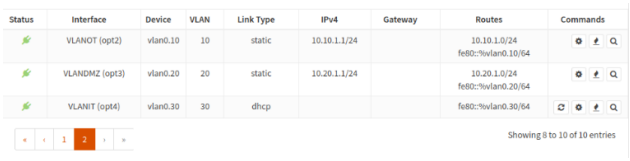


Fig. 7. Interfaces Firewall

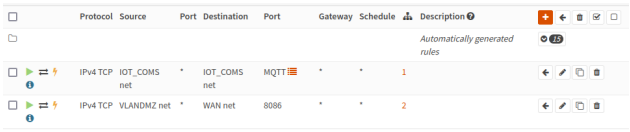
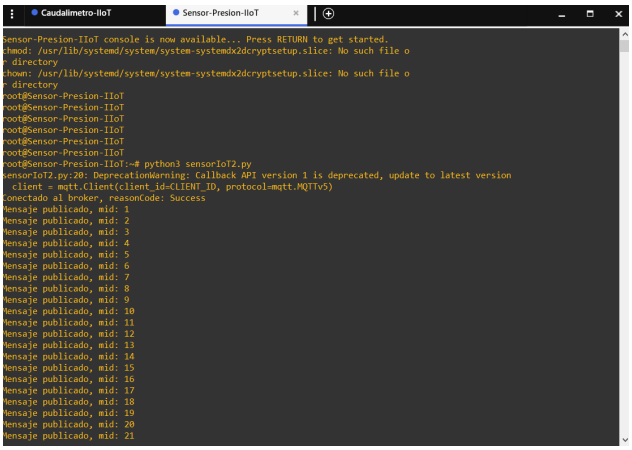
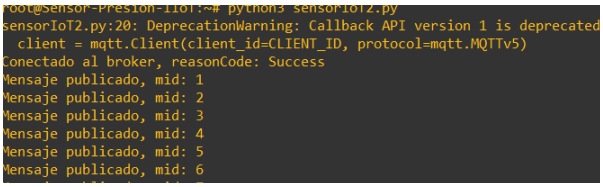


Fig. 8. Reglas / ACL del *firewall* perimetral OPNSENSE.



(a) Screenshot



(a) Zoom

Fig. 9. Dispositivo IIoT en funcionamiento.

Los valores medios obtenidos fueron: latencia sensor->DB 22,63 ± 7,49 ms, carga CPU en Node-RED 2,85 ± 0,81 %, y solo 3 reglas activas en OPNSense que admiten únicamente el tráfico estrictamente necesario, siendo el resto bloqueado por la regla *default deny*. Estos resultados confirman que la sobrecarga introducida por la pasarela es marginal y que la arquitectura cumple holgadamente los requisitos típicos de disponibilidad para gestión y supervisión remota de plantas industriales.

La Figura 10 recoge la gráfica en tiempo real generada en Grafana a partir de InfluxDB y evidencia la entrega continua de los datos sin pérdidas.

La aplicación práctica de los controles se ilustra en Figura 8, donde se listan las reglas de cortafuegos que materializan SR 3.1 de IEC-62443-3-3 (control de acceso de red) y el principio de prevención del ENS. Para asegurar integridad y autenticidad (SR 5.3, principios ENS de integridad y trazabilidad), se configuró InfluxDB con *bucket* "IoT-INTEGRATION" y *token* "IoT-aguas". Así, el proceso se documenta en la Figura 11, que agrupa la creación de *bucket*, *token* y nodo de salida en Node-RED.

Los ensayos confirman que la DMZ industrial y los conductos seguros impiden la comunicación OT -> IT directa, cumpliendo la separación de dominios. La latencia se mantiene por debajo de los 100 ms sin optimizaciones específicas, lo que sugiere que la arquitectura es escalable al menos hasta 500 mensajes/s (500 sensores enviando 1 mensaje por segundo) con hardware equivalente. Además, la gestión de identidades basada en tokens reduce el riesgo de credenciales expuestas y facilita la auditoría.

En este punto, es importante indicar las limitaciones presentes en este trabajo. La validación se efectuó en un entorno virtual sin tráfico malicioso deliberado. No se incluyó un sistema de detección de intrusos (*Intrusion Detection System*, IDS) ni se evaluó la resiliencia ante ataques de Denegación de Servicio (*Denial of Service*, DoS). El flujo se limitó a MQTT/TLS sobre TCP y, por tanto, protocolos como OPC UA o AMQP no se han testado.

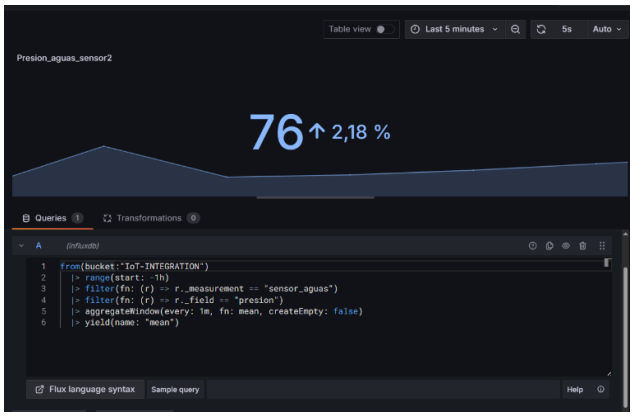


Fig. 10. Creación *dashboard* y visualización de datos IIoT.

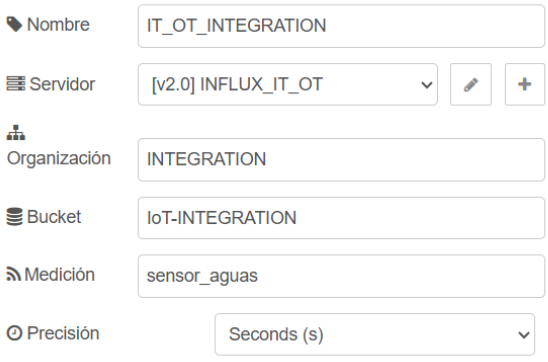


Fig. 11. Configuración del Nodo de Salida en InfluxDB.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

V. Conclusión

Se ha validado una arquitectura IT-OT de tres zonas (OT, DMZ, IT) que haciendo uso de software *open-source* y contenedores *docker*, cumple algunos de los requisitos de ENS e IEC-62443-3-3 aplicables. En un laboratorio GNS3 con 100 sensores IIoT (100 mensajes/s agregados) se midió una latencia extremo a extremo de 22,63 ± 7,49 ms y una carga CPU en la pasarela Node-RED del 2,85 ± 0,81 %, sin pérdidas de mensajes. El cortafuegos OPNSense, apoyado en VLAN y TLS1.2, aseguró la separación de dominios y la trazabilidad exigida. La propuesta ofrece un prototipo base para ayudar y simplificar la convergencia OT-IT bajo marcos normativos exigentes. Como trabajos futuros, se plantea integrar un IDS en la DMZ, automatizar el despliegue con *docker-compose* y pipelines CI/CD que verifiquen IEC-62443-4-2, evaluar la adopción de un modelo *Zero Trust* para acceso remoto a estaciones SCADA para su gestión mediante comunicaciones entre Firewalls y tokens de MFA y, por último, la validación de este tipo de entornos con tramas de tráfico malicioso y ataques DDoS para sistemas IIoT.

Agradecimientos

Este trabajo forma parte del proyecto Laboratorio de I+D+i en Ciberseguridad, Privacidad y Comunicaciones Seguras (TRUST Lab), financiado por la Unión Europea NextGeneration-EU, Plan de Recuperación, Transformación y Resiliencia, a través del INCIBE.

Referencias

[1] N. Laan, R. Gupta, A. W. Koehler, y W. Van Hill, «Implementing Cybersecurity for Industrial Connected Products: How to Use Standards to Protect Operational Technology», *IEEE Industry Applications Magazine*, pp. 2-11, 2025, doi: 10.1109/MIAS.2025.3559862.

[2] Ministerio de Asuntos Económicos y Transformación Digital, *Real Decreto 311/2022, de 3 de mayo, por el que se regula el Esquema Nacional de Seguridad*. Spain, 2022, pp. 61715-61804.

[3] International Electrotecnichal Commission, «IEC 62443-4-1:2018», IEC 62443 (all parts). Accedido: 11 de junio de 2025. [En línea]. Disponible en: https://webstore.iec.ch/en/publication/series/5029

[4] I. Cindrić, M. Jurčević, y T. Hadjina, «Mapping of Industrial IoT to IEC 62443 Standards», *Sensors*, vol. 25, n.º 3, p. 728, ene. 2025, doi: 10.3390/sms30728.

[5] G. Schäfer, H. Waclawek, S. Riedmann, C. Binder, C. Neureiter, y S. Huber, «IT/OT Integration by Design», *INCOSE International Symposium*, vol. 34, n.º 1, pp. 337-352, jul. 2024, doi: 10.1002/iis2.13150.

[6] M. Bhole, W. Kastner, y T. Sauter, «IT Security Solutions for IT/OT Integration: Identifying Gaps and Opportunities», en *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, sep. 2024, pp. 01-08. doi: 10.1109/ETFA61755.2024.10710968.

[7] C. Göttel, M. Kabir-Querrec, D. Kozhaya, T. Sivanthi, y O. Vuković, «Qualitative Analysis for Validating IEC 62443-4-2 Requirements in DevSecOps», en *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, sep. 2023, pp. 1-8. doi: 10.1109/ETFA54631.2023.10275637.

[8] H. Hmiddouch, A. Villafranca, R. Castro, V. Dubetskyy, y M.-D. Cano, «Enhancing Industrial Cybersecurity with Virtual Lab Simulations», *International Journal of Advanced Computer Science and Applications*, vol. 16, n.º 5, 2025, doi: 10.14569/IJACSA.2025.0160505.

[9] Centro Criptológico Nacional, «Guía de Seguridad de las TIC CCN-STIC 808 ENS. Verificación del cumplimiento Octubre 2023», oct. 2023. [En línea]. Disponible en: https://cpage.mpr.gob.es

[10] P. Viorel, «The MING Stack: What It Is and How It Works», julio de 2024. [En línea]. Disponible en: https://www.iiot-world.com/industrial-iot/connected-industry/the-ming-stack-what-it-is-and-how-it-works/

[11] Inductive Automation, «Design Like a Pro: Best Practices for IIoT», 13 de diciembre de 2018. [En línea]. Disponible en: https://inductiveautomation.com/resources/article/design-like-a-pro-best-practices-for-iiot

[12] Marcel van Rensburg, David Dolezilek, y Jason Dearien, «Case Study: Lessons Learned Using IEC 61850 Network Engineering Guideline Test Procedures to Troubleshoot Faulty Ethernet Network Installations», *Power and Energy Automation Conference Spokane, Washington*, mar. 2017.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

A Reinforcement Learning-Based Telematic Routing Protocol for the Internet of Underwater Things

Mohammadhossein Homaei*, Mehran Tarif†, Agustín Di Bartolo*, Víctor González Morales*, Mar Ávila Vegas*
*Department of Computer Systems Engineering and Telematics, University of Extremadura, Cáceres, 10003, Extremadura, Spain
Email: mhomaein@alumnos.unex.es, adibartolo@unex.es, victorgomo@unex.es, mmavila@unex.es
†Department of Computer Science, University of Verona, 37134 Verona, Italy
Email: mehran.tarifhokmabadi@univr.it

The Internet of Underwater Things (IoUT) has a lot of problems, like low bandwidth, high latency, mobility, and not enough energy. Routing protocols that were made for land-based networks, like RPL, don't work well in these underwater settings. This paper talks about RL-RPL-UA, a new routing protocol that uses reinforcement learning to make things work better in underwater situations. Each node has a small RL agent that picks the best parent node depending on local data such the link quality, buffer level, packet delivery ratio, and remaining energy. RL-RPL-UA works with all standard RPL messages and adds a dynamic objective function to help people make decisions in real time. Aqua-Sim simulations demonstrate that RL-RPL-UA boosts packet delivery by up to 9.2%, uses 14.8% less energy per packet, and adds 80 seconds to the network's lifetime compared to previous approaches. These results show that RL-RPL-UA is a potential and energy-efficient way to route data in underwater networks.

***Keywords*—Internet of Underwater Things, Reinforcement Learning, RPL, Adaptive Routing, Energy Efficiency.**

I. INTRODUCTION

The IoUT is becoming an important technology for things like monitoring the environment in the ocean, exploring underwater, and checking infrastructure on the ocean floor. These systems use acoustic sensor networks that work in tough conditions with high latency, unstable connections, and severe limits on energy and bandwidth [1], [2]. Underwater networks use acoustic signals instead of radio waves, which are slower and less dependable than radio waves. Because nodes are generally powered by batteries that are hard to replenish, it's very important that they use as little energy as possible. In addition, node movement can affect the network topology, which makes routing and data delivery uncertain. These problems call for communication methods that are flexible and light.

Many terrestrial IoT devices employ the IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL). It uses metrics like hop count or Expected Transmission Count (ETX) to create routing trees called Destination-Oriented Directed Acyclic Graphs (DODAGs). RPL works well in static and low-power situations, but it doesn't work well in dynamic and delay-prone underwater environments since it doesn't change much and doesn't react quickly enough. In our last paper [3], we changed RPL's objective function (OF) and communication logic to make it work better underwater.

However, our previous work [3] had some problems: (i) it used static objective function weights that couldn't change when the network conditions changed, (ii) it couldn't learn in real time to make better routing decisions based on past performance, and (iii) it needed to have its parameters manually tuned for each deployment scenario. These problems show how important it is to find a routing method that is both flexible and energy-efficient, and that works with current RPL systems. We want to use reinforcement learning to build a system like this so that underwater nodes can respond to changing conditions without having to pay a lot for processing or transmission.

This study builds on our prior work [3] to show RL-RPL-UA, a variant of RPL for IoUT that uses RL. A lightweight RL agent runs on each node and chooses the next hop depending on things like energy level, link quality, buffer condition, and delivery history. The protocol uses a composite, adjustable OF to help choose parents while still being fully compatible with RPL control messages (DIO/DAO). RL works well for adaptive routing

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

in networks that change over time because it lets nodes make better judgements based on what they learn. But most RL-based protocols either don't work with RPL or need a lot of resources. RL-RPL-UA solves this problem by providing a system that can grow and use resources efficiently, and it can adapt to underwater conditions without changing the structure of the RPL protocol.

The rest of the paper is organized as follows: Section II reviews relevant routing protocols for IoUT. Section III describes the RL-RPL-UA architecture and objective function. Section IV presents the simulation setup and performance evaluation. Finally, Section V summarizes conclusions and outlines future work.

II. RELATED WORKS

IoUT networks are very hard to utilise because of their specific physical and operational limits, such as high propagation latency, limited energy supplies, frequent disconnections, and dynamic topologies that happen as nodes move around. There have been several suggestions for routing systems that might fix these issues. There are five primary groups of these: (i) clustering and depth-based protocols; (ii) game-theoretic and opportunistic strategies; (iii) AI- and RL-driven approaches; (iv) bio-inspired and meta-heuristic algorithms; and (v) extensions based on RPL.

A. Depth-Based and Clustered Routing

Depth-based and clustered routing approaches try to reduce latency and energy use by putting nodes into clusters or leveraging their depth information. Early protocols like C-GCo-DRAR[4] and U-(ACH)^2 [5] use node depth and adaptive cluster building to lower latency and transmission overhead. FLCEER employs fuzzy logic [6] to choose the best choice for the cluster head and extend the life of the network. IDA-OEP employs smart data analytics to transmit data in a way that saves energy [1], whereas BES uses bald-eagle-search optimisation to make sure that data is sent in the most energy-efficient way [7]. These plans work well in situations that are mobile or very dynamic, but they are typically not adaptable and need to be calibrated to the environment accurately, even though they work well in static conditions.

B. Opportunistic and Game-Theoretic Methods

Protocols that are opportunistic and game-theoretic deal with void zones, make things more reliable, and consume less energy. For example, GTRP employs Nash equilibria to choose relays in 3-D acoustic networks [8]. PCR [9] changes the power of the gearbox in real time, but hybrid solutions like A-ANTD [10] and TARD [2] use autonomous underwater vehicles (AUVs) to collect data that can handle delays. These designs do make some deployments work better, but they usually need a lot of pre-configuration or centralised control, which stops them from being able to work on their own on a big scale.

C. AI and RL Approaches

Recent work has employed AI to make routing that adapts to itself. Li et al. utilise multi-agent RL for optical IoUT links [11], Khan et al. use Q-learning for void mitigation [8], and Nandyala et al. create topology-aware Q-routing [12], [13]. Tarif et al. use fuzzy inference to stabilise pathways while they are moving. Tarif et al. (2025) present UWF-RPL, a fuzzy-logic extension of RPL that weights ETX, depth, residual energy, and latency in a Mamdani controller [14], [15]. This results in a 17% PDR gain and 15% energy savings over baseline RPL [16]. But unlike the lightweight Q-learning agent we utilised in our RL-RPL-UA, its rule base is fixed and it can't alter weights while it's running.

D. Meta-Heuristic and Bio-Inspired Algorithms

Bio-inspired and meta-heuristic approaches like FFRP (Firefly) [17], EORO (enhanced PSO) [18], and BES [7] employ swarm intelligence to find paths that consume less energy. Pradeep et al. [19] provide a fuzzy region-based approach that works with sink mobility. Even if their simulation results are promising, they often can't be used in the real world since they need global optimisation and don't learn continuously.

E. RPL-Based Extensions

The RPL was first created for sensor networks on land, but it has now been changed to function underwater. Because RPL is the primary protocol for IoT on land, a lot of research has gone into making it work with underwater acoustic communication. In UW/MRPL [3], we made RPL work better in underwater settings by adding routing metrics that take depth into account and support for mobility. It was better than basic protocols like OF0 and MRHOF, but it couldn't adjust in real time and used fixed objective function (OF) weights. To solve the problems of compatibility and energy balancing, UWF-RPL [16] added a fuzzy logic-based OF to conventional RPL control messages (DIO/DAO). It didn't, meanwhile, use feedback systems to make routing decisions better, and membership functions still had to be changed by hand. Our proposed RL-RPL-UA improves RPL by adding an RL agent that automatically changes OF weights in real time. This keeps complete protocol compatibility and gets rid of the requirement for human configuration.

Table I
COMPARISON OF RECENT ROUTING PROTOCOLS WITH RL-RPL-UA

Protocol	RL	RPL	Adaptive OF	Mobility	Citation
C-GCo-DRAR	–	–	Static OF	Limited	[4]
FLCEER	–	–	Static OF	Moderate	[6]
IDA-OEP	–	–	Static OF	Limited	[1]
GTRP	–	–	Static OF	Moderate	[8]
RL Protocol	✓	–	Static OF	Moderate	[13]
Q-Learning	✓	–	Dynamic OF	Moderate	[12]
Multi-agent RL	✓	–	Static OF	High	[11]
UA-RPL	–	✓	Static OF	Moderate	[20]
URPL	–	✓	Dynamic OF	Moderate	[14]
Fuzzy-CR	–	✓	Decision Making	Moderate	[15]
UWF-RPL	–	✓	Static Fuzzy OF	Moderate	[16]
UW/MRPL (prev. work)	–	✓	Static OF	High	[3]
RL-RPL-UA (this work)	✓	✓	Dynamic	High	–

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

We briefly review them above, and use Tables I and II to illustrate why RL-RPL-UA is necessary and novel.

III. PROPOSED PROTOCOL: RL-RPL-UA

In this section, we introduce RL-RPL-UA, a novel routing protocol that enhances the conventional RPL protocol by embedding an RL agent into each node of the underwater IoT network. Unlike traditional RPL implementations that rely on static OFs, our model leverages an adaptive learning mechanism to select optimal routes under the harsh and dynamic conditions of underwater communication.

A. Protocol Architecture

The architecture of RL-RPL-UA integrates an RL agent within the standard RPL stack. Each node consists of the following modules:

- **Sensing Unit:** Gathers state information including residual energy, buffer occupancy, and signal strength.
- **Communication Module:** Interfaces with an acoustic modem or underwater simulation module (AquaSim, NS-2 with UAN).
- **RL Agent:** A local Q-learning or DQN model.
- **Extended RPL Stack:** Supports modified DIO/DAO messages carrying dynamic state and learned metrics.

B. RL Model

The routing process is modeled as a Markov Decision Process (MDP), where each node learns an optimal routing policy by interacting with its environment.

1) State Space: The state s_t at time t includes:

$$s_t = [E_t, \text{LQI}_t, Q_t, \text{PDR}_t, T_t] \tag{1}$$

where E_t is residual energy, LQI_t is link quality indicator, Q_t is current queue size, PDR_t is historical packet delivery ratio, and T_t is time since last successful transmission.

2) Action Space: The action a_t is the selection of a next-hop parent from among n neighbors:

$$a_t \in \{\text{Parent}_1, \text{Parent}_2, \ldots, \text{Parent}_n\} \tag{2}$$

As shown in Equation (2), the action space consists of the set of all neighboring nodes that can serve as the next hop in the routing process.

3) Reward Function: The reward signal r_t is defined to balance reliability, delay, and energy consumption:

$$r_t = \alpha \cdot \text{PDR}_t - \beta \cdot \text{Delay}_t - \gamma \cdot \text{EnergyCost}_t \tag{3}$$

As shown in Equation (3), this formulation enables the agent to optimize routing decisions by weighing the positive effect of packet delivery against the negative impact of delay and energy consumption. Here, α, β, and γ are tunable hyperparameters that control the trade-offs between these objectives. This reward is used to update the RL agent's policy.

Algorithm 1 RL-enhanced RPL Routing

Require: Initialization of Q-table or DQN weights, neighbor table, default Rank
Ensure: Energy-efficient and adaptive routing in underwater IoT
Initialize RL agent (Q-table or DQN), default Rank
$s \leftarrow$ OBSERVE_STATE
Broadcast DIO with $OF_{RL}(n_i)$ and node state
while Node is active **do**
  Receive DIOs from neighbors
  for all neighbor n_i in NeighborTable **do**
    Extract state features: $s_{n_i} = [E, LQI, Q, PDR, T]$
    Compute $OF_{RL}(n_i)$ using Equation 5
    Estimate $Q(s, a = n_i)$ using RL model (Q-table or DQN)
  end for
  Select Parent:
  $a^* \leftarrow \arg\max_{n_i} Q(s, a = n_i)$ ▷ Best next-hop based on RL
  Update RPL Rank based on selected parent and $OF_{RL}(a^*)$
  Forward data packets to a^*
  Wait for Acknowledgement or Timeout
  Observe outcome:
  Measure PDR_t, Delay_t, EnergyCost_t
  Compute reward r_t using Equation 3
  $s_{t+1} \leftarrow$ OBSERVE_STATE
  RL Update:
  if Using Q-learning **then**
    Update Q-table using Equation 4
  else if Using DQN **then**
    Store (s_t, a^*, r_t, s_{t+1}) in ReplayBuffer
    Train DQN via minibatch sampling
  end if
  $s_t \leftarrow s_{t+1}$
  Periodically broadcast updated DIO with new Rank and OF_{RL}
end while
function OBSERVE_STATE
  Measure local energy E, link quality LQI, buffer queue Q, packet success rate PDR, time since last ACK T
  return $[E, LQI, Q, PDR, T]$
end function

4) Policy Learning: The RL agent seeks to learn a policy $\pi(a|s)$ that maximizes the expected cumulative reward:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \eta \Big[r_t + \gamma \cdot \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \Big] \tag{4}$$

Equation 4 is the standard Q-learning update rule, where η is the learning rate and γ is the discount factor for future rewards.

C. Adaptive OF

To replace the static OFs in RPL, we define a composite and dynamic OF:

$$OF_{RL}(n_i) = w_1 \cdot E(n_i) + w_2 \cdot R(n_i) + w_3 \cdot Q(n_i) + w_4 \cdot PDR(n_i) \tag{5}$$

Where:

- $E(n_i)$: Normalized remaining energy of neighbor n_i,
- $R(n_i)$: Link reliability (e.g., inverse of ETX),
- $Q(n_i)$: Queue length or buffer utilization,
- $PDR(n_i)$: Historical delivery ratio,
- w_1 to w_4: Adaptive weights tuned by the RL agent.

This OF is broadcast in DIO messages, allowing each node to evaluate its neighbors and update its rank dynamically.

D. Routing Decision Process

The routing process in RL-RPL-UA involves the following steps:

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Table II
KEY DIFFERENCES BETWEEN UWF-RPL, UWMRPL (OUR PREVIOUS WORKS), AND RL-RPL-UA

Feature	UWF-RPL [16]	UWMRPL [3] (Previous work)	RL-RPL-UA (This Work)
Main Concept	Fuzzy-logic RPL for optimized routing	Mobility-aware RPL with static tunable OF	RL-based dynamic routing with local agents
Routing Adaptability	Semi-adaptive via static fuzzy logic rules	Semi-adaptive via predefined logic	Fully adaptive via real-time RL updates
OF	Static fuzzy logic (depth, energy, latency, ETX)	Static/custom (ETX, depth)	Dynamic composite (energy, LQI, queue, PDR)
RPL Compatibility	Extended DIO/DAO with fuzzy logic metrics	Standard-compliant	Extended DIO/DAO with RL metrics
Learning Agent	None (static fuzzy logic)	None	Q-learning or DQN per node
Reward Mechanism	None	None	$\alpha \cdot \text{PDR} - \beta \cdot \text{Delay} - \gamma \cdot \text{Cost}$
Overhead	Moderate (fuzzy logic computations)	Low (no learning updates)	Low (optimized RL updates)
Mobility Handling	Reactive via fuzzy logic evaluation	Reactive DAG repair	Proactive via learned feedback
Queue Management	Included (congestion control)	Not included	Included (adaptive queue management)
Energy Efficiency	Good (static optimized)	Moderate (no dynamic optimization)	High (real-time optimization)
Key Contribution	Improved stability and efficiency via fuzzy logic	Mobility and energy-aware extension of RPL	Online adaptive parent selection via RL

1) DIO Exchange: Each node broadcasts its current state and OF_{RL} value using an extended DIO message.
2) Neighbor Table Update: On receiving DIOs, nodes update their neighbor tables and estimate Q-values.
3) Parent Selection: The parent with the highest Q-value is selected as the preferred next-hop.
4) Data Forwarding: Data packets are forwarded along the learned optimal path.
5) Learning Update: After each transmission, the node observes outcomes and updates its Q-table using Equation 4.

E. Underwater-Specific Enhancements

The following improvements are tailored to the underwater environment:

- Delay Estimation: Nodes estimate propagation delay based on distance and water temperature to better model the reward function.
- Energy-Aware Slot Scheduling: TDMA (Time Division Multiple Access) is used as a MAC protocol to assign non-overlapping time slots to nodes, reducing collisions and idle listening in underwater acoustic networks.
- Clustered Learning: In large networks, cluster heads can aggregate policies and periodically disseminate updates.

F. Compatibility and Overhead

RL-RPL-UA remains compatible with legacy RPL nodes by embedding new fields in the optional sections of RPL messages. In terms of complexity:

- The Q-learning implementation requires minimal computational resources and is suitable for constrained devices.
- Communication overhead is slightly increased due to additional state sharing, but overall packet retransmissions are reduced.

G. Security Considerations

RL-RPL-UA can be extended to support trust-aware routing by integrating reputation scores into the reward function, allowing the network to avoid compromised nodes.

H. Resource and Energy Cost Estimation

To evaluate the feasibility of deploying RL-RPL-UA in real-world IoUT scenarios, we estimate the energy and processing cost based on standard underwater sensor node specifications. We consider nodes equipped with low-power microcontrollers (e.g., MSP430, ARM Cortex-M4) and acoustic modems such as the WHOI Micromodem or EvoLogics S2C.

1) Energy Cost per Transmission: Assuming a transmission power of 0.5 W and transmission time of 1.5 seconds per packet, the energy cost per transmission is calculated as:

$$E_{tx} = P_{tx} \times t = 0.5 \times 1.5 = 0.75 \text{ J} \tag{6}$$

As shown in Equation (6), each data transmission consumes approximately 0.75 joules.

2) Energy Cost per RL Update: The Q-table update process requires approximately 500–1000 CPU cycles. For a 16 MHz processor operating at 1.8 V and 3 mA, the energy cost is given by:

$$E_{cpu} = V \times I \times \frac{\text{cycles}}{f} = 1.8 \times 0.003 \times \frac{1000}{16 \times 10^6} \approx 0.34 \mu\text{J} \tag{7}$$

According to Equation (7), the energy consumption for a single RL update is approximately 0.34 microjoules.

3) Memory and Storage: The Q-table of size $n \times a$ with 8-bit values, for example with 10 neighbors and 5 actions, requires approximately 50 bytes. This is feasible for microcontrollers with at least 32 KB of SRAM.

4) Discussion: Compared to traditional RPL, RL-RPL-UA introduces minimal computational overhead due to the small Q-table and low RL update cost. However, it improves energy efficiency by reducing retransmissions and adapting paths proactively.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

IV. SIMULATION RESULTS

A. SIMULATION PARAMETERS

To assess the performance of the proposed RL-RPL-UA protocol, we conducted simulations using Aqua-Sim, an extension of the NS-2 (Network Simulator 2) framework specifically designed for underwater acoustic network environments [22]. Aqua-Sim is available at https://github.com/rmartin5/aqua-sim-ng. The simulated network consists of both static and mobile sensor nodes deployed within a 3D underwater space using acoustic communication links. Each node independently executes the RL-RPL-UA algorithm and exchanges routing information via modified DIO/DAO messages.

We compare RL-RPL-UA against several baseline protocols, including standard RPL (OF0), Q-learning-only approaches, and cluster-based routing. We specifically selected Co-DRAR [4] as it represents state-of-the-art depth-based clustered routing, UA-RPL [20] as a recent RPL adaptation for underwater networks, and UWF-RPL [16] as the most recent fuzzy-logic enhancement of RPL, along with our prior work UW/MRPL [3]. These protocols were chosen because they cover different routing paradigms (clustered, RPL-based, fuzzy-enhanced, and mobility-aware) and allow comprehensive evaluation of RL-RPL-UA's improvements in adaptability, energy efficiency, and delivery reliability. Evaluation metrics include Packet Delivery Ratio (PDR), End-to-End Delay, Energy Consumption, Routing Overhead, and Network Lifetime.

The main simulation parameters are listed in Table III.

Table III
SIMULATION PARAMETERS FOR RL-RPL-UA EVALUATION

Parameter	Value
Simulation Area	$300 \times 300 \times 300$ m^3
Number of Nodes	50
Initial Energy per Node	5 J
Transmission Power	0.5 W
Acoustic Bandwidth	10 kHz
Propagation Speed	1500 m/s
MAC Protocol	TDMA
Routing Protocols	RL-RPL-UA, RPL (OF0), Q-Routing
RL Algorithm	Q-learning (tabular)
Learning Rate (η)	0.1
Discount Factor (γ)	0.9
Reward Weights (α, β, γ)	(1.0, 0.6, 0.4)
Simulation Time	1000 s
Traffic Model	CBR, 1 packet/10 s
Packet Size	64 Bytes
Mobility Model	Random Waypoint (0.1–0.3 m/s)

B. Packet Delivery Ratio

PDR is calculated over K simulation trials as [21]:

$$\text{PDR}_{\text{mean}} = \frac{1}{K} \sum_{k=1}^{K} \left(\frac{R_k}{S_k} \right) \times 100 \tag{8}$$

$$\sigma_{\text{PDR}} = \sqrt{\frac{1}{K-1} \sum_{k=1}^{K} \left(\frac{R_k}{S_k} - \text{PDR}_{\text{mean}} \right)^2} \tag{9}$$

In the static scenario, RL-RPL-UA achieves a mean Packet Delivery Ratio (PDR) of 94.3% with a standard deviation of 1.7, outperforming UWF-RPL (89.2%, σ=2.2), UWRPL (85.1%, σ=3.0), UA-RPL (83.5%, σ=3.0), and Co-DRAR (81.2%, σ=2.8). The results show that while UWF-RPL enhances PDR over traditional RPL variants by using adaptive metrics, RL-RPL-UA delivers a further 5.1% improvement over UWF-RPL and 9.2% over UWRPL, confirming the impact of RL in static deployments.

In the mobile scenario, RL-RPL-UA maintains the highest delivery performance with a PDR of 92.8% (σ=1.9), surpassing UWF-RPL (90.5%, σ=2.0), UWMRPL (88.2%, σ=2.3), UA-RPL (80.2%, σ=3.1), and Co-DRAR (78.4%, σ=3.2). Although UWF-RPL narrows the performance gap in mobile conditions through fuzzy logic and energy-aware decisions, RL-RPL-UA outperforms all baselines, confirming that its real-time learning approach significantly improves delivery reliability under dynamic underwater environments.

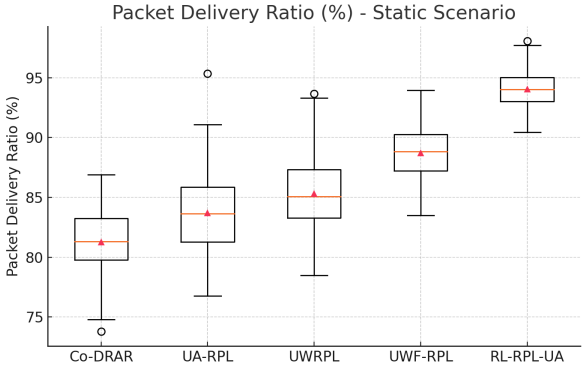


Fig. 1. PDR in the static scenario.

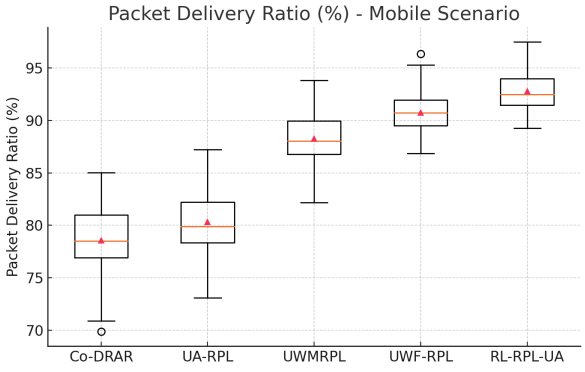


Fig. 2. PDR in the mobile scenario.

C. End-to-End Delay

The average packet delay in trial k is:

$$\text{Delay}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} \left(t_j^{\text{recv}} - t_j^{\text{send}} \right) \tag{10}$$

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

The overall mean and deviation:

$$\text{Delay}_{\text{mean}} = \frac{1}{K}\sum_{k=1}^{K}\text{Delay}_k \tag{11}$$

$$\sigma_{\text{Delay}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(\text{Delay}_k - \text{Delay}_{\text{mean}})^2} \tag{12}$$

In the static scenario, RL-RPL-UA achieves an average end-to-end delay of 1.8 s (σ=0.2), outperforming UWF-RPL (2.0 s, σ=0.25), UWRPL (2.4 s, σ=0.3), UA-RPL (2.7 s, σ=0.4), and Co-DRAR (2.9 s, σ=0.4). The introduction of UWF-RPL demonstrates improvement over conventional RPL extensions, yet RL-RPL-UA further reduces delay by 10% compared to UWF-RPL and by 25% relative to UWRPL.

In the mobile scenario, RL-RPL-UA sustains low latency with an average delay of 1.9 s (σ=0.2), outperforming UWF-RPL (1.95 s, σ=0.25), UWMRPL (2.1 s, σ=0.3), UA-RPL (2.8 s, σ=0.4), and Co-DRAR (3.1 s, σ=0.4). These results highlight the RL agent's effectiveness in minimizing transmission delay under mobile and dynamically changing underwater network conditions (Figures 3, 4).

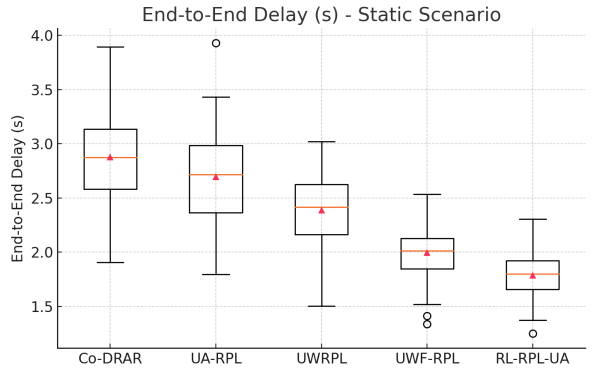


Fig. 3. End-to-End Delay in the static scenario.

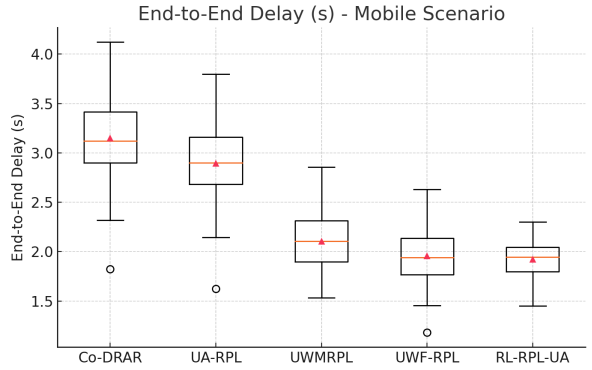


Fig. 4. End-to-End Delay in the mobile scenario.

D. Energy per Delivered Packet

Per trial, the energy cost per packet is:

$$\text{E}_k = \frac{E_{\text{total},k}}{R_k} \tag{13}$$

Mean and deviation:

$$\text{E}_{\text{mean}} = \frac{1}{K}\sum_{k=1}^{K}\text{E}_k \tag{14}$$

$$\sigma_{\text{E}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(\text{E}_k - \text{E}_{\text{mean}})^2} \tag{15}$$

In the static scenario, RL-RPL-UA achieves an average energy cost of 0.75 J per delivered packet (σ=0.05), significantly lower than UWF-RPL (0.78 J, σ=0.06), UWRPL (0.88 J, σ=0.07), UA-RPL (0.89 J, σ=0.07), and Co-DRAR (0.91 J, σ=0.08). While UWF-RPL improves energy efficiency by integrating adaptive cost metrics, RL-RPL-UA achieves an additional 3.8% energy saving over UWF-RPL and 14.8% over UWRPL, confirming its superior resource-awareness (Figure 5).

In the mobile scenario, RL-RPL-UA continues to deliver the most energy-efficient performance with an energy cost of 0.74 J (σ=0.05), followed by UWF-RPL (0.76 J, σ=0.055), UWMRPL (0.82 J, σ=0.06), UA-RPL (0.91 J, σ=0.08), and Co-DRAR (0.94 J, σ=0.09). These improvements reflect the effectiveness of the RL-based adaptive routing strategy in minimizing retransmissions and avoiding energy-intensive paths, even under dynamic network conditions (Figure 6).

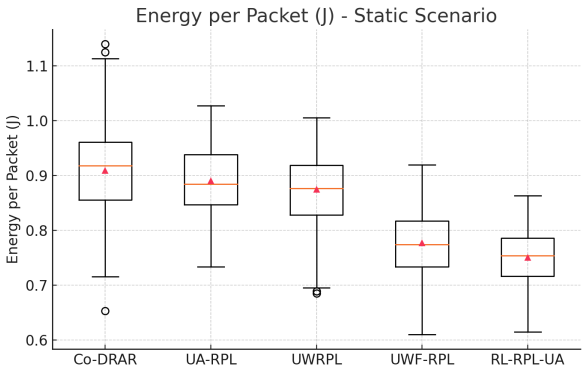


Fig. 5. Energy cost per delivered packet in the static scenario.

E. Routing Overhead Ratio

Overhead is computed as:

$$\text{OH}_k = \frac{C_k}{R_k} \tag{16}$$

With:

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

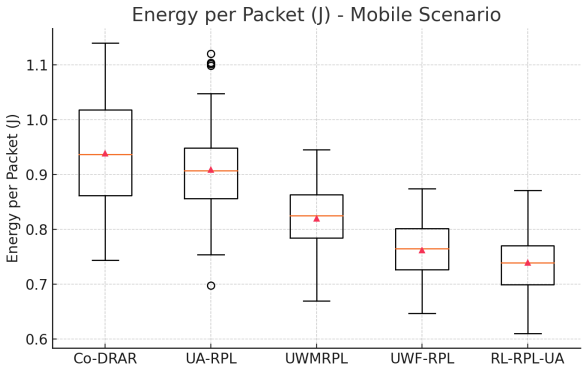


Fig. 6. Energy cost per delivered packet in the mobile scenario.

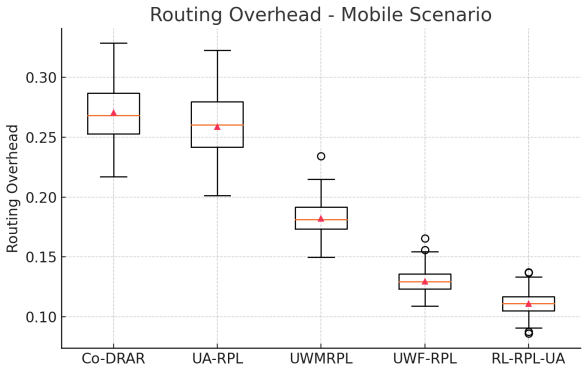


Fig. 8. Routing overhead ratio in the mobile scenario.

$$\mathrm{OH}_{\mathrm{mean}} = \frac{1}{K}\sum_{k=1}^{K}\mathrm{OH}_k \tag{17}$$

$$\sigma_{\mathrm{OH}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(\mathrm{OH}_k - \mathrm{OH}_{\mathrm{mean}})^2} \tag{18}$$

In the static scenario, RL-RPL-UA introduces the lowest control overhead with a mean routing overhead ratio of 0.12 (σ=0.01), outperforming UWF-RPL (0.14, σ=0.015), UWRPL (0.22, σ=0.02), UA-RPL (0.24, σ=0.023), and Co-DRAR (0.25, σ=0.025). Although UWF-RPL reduces overhead compared to UWRPL and other classical protocols, RL-RPL-UA further reduces control traffic by 14.3% over UWF-RPL and 45% over UWRPL (Figure 7).

In the mobile scenario, RL-RPL-UA maintains minimal overhead at 0.11 (σ=0.01), followed by UWF-RPL (0.13, σ=0.012), UWMRPL (0.18, σ=0.015), UA-RPL (0.26, σ=0.028), and Co-DRAR (0.27, σ=0.03). Presented in consistent protocol order, these results confirm the effectiveness of RL-RPL-UA in suppressing control overhead even in dynamic, mobile environments (Figure 8)..

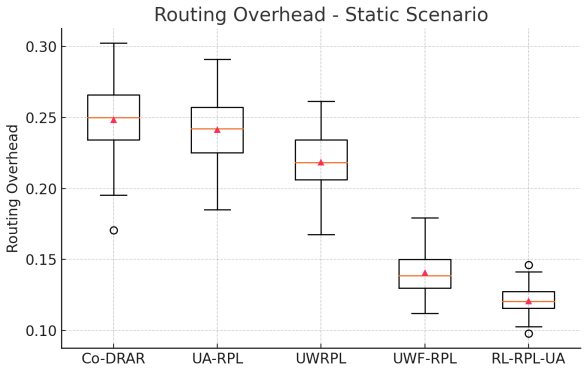


Fig. 7. Routing overhead ratio in the static scenario.

F. Network Lifetime

Lifetime is defined as the time until the first node in the network depletes its energy:

$$T_{\mathrm{mean}}^{\mathrm{death}} = \frac{1}{K}\sum_{k=1}^{K}T_{\mathrm{death}}^{(k)} \tag{19}$$

$$\sigma_T = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(T_{\mathrm{death}}^{(k)} - T_{\mathrm{mean}}^{\mathrm{death}}\right)^2} \tag{20}$$

In the static scenario, RL-RPL-UA achieves the longest network lifetime of 720 seconds (σ=15), followed by UWF-RPL (690 s, σ=18), UWRPL (640 s, σ=22), UA-RPL (610 s, σ=24), and Co-DRAR (600 s, σ=25). The integration of fuzzy optimization in UWF-RPL enhances node longevity, but RL-RPL-UA further extends the lifetime by 30 seconds over UWF-RPL and 80 seconds over UWRPL, confirming the benefit of RL in energy-aware route planning (Figure 9).

In the mobile scenario, RL-RPL-UA sustains the longest network operation at 710 seconds (σ=16), ahead of UWF-RPL (700 s, σ=17), UWMRPL (680 s, σ=20), UA-RPL (590 s, σ=26), and Co-DRAR (580 s, σ=28). The improvement stems from RL-RPL-UA's ability to distribute energy consumption more evenly across nodes by dynamically selecting optimal, energy-efficient paths under varying underwater mobility conditions (Figure 10).

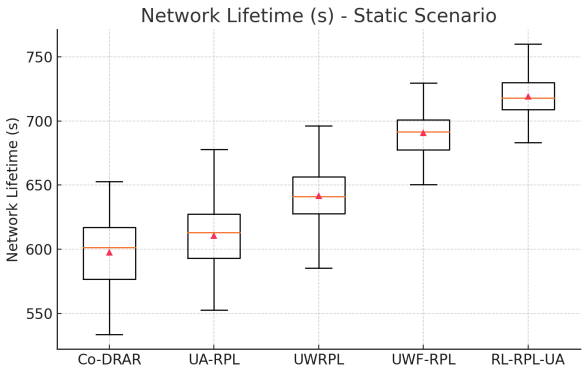


Fig. 9. Network lifetime (time until first node dies) in the static scenario.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

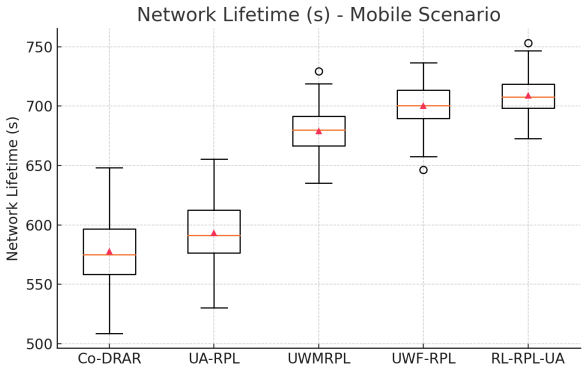


Fig. 10. Network lifetime (time until first node dies) in the mobile scenario.

V. CONCLUSION

This work presented RL-RPL-UA, a reinforcement learning-based extension of the RPL protocol designed for the challenges of the Internet of Underwater Things (IoUT). By incorporating Q-learning agents, the protocol adapts dynamically to changing network conditions and selects routing paths based on multiple performance criteria, including energy efficiency, link quality, queue length, and delivery reliability. The evaluation included a comprehensive comparison with recent baseline protocols under both static and mobile scenarios. Simulation results show that RL-RPL-UA offers consistent improvements in reliability, delay, energy consumption, control overhead, and network lifetime, suggesting that reinforcement learning can effectively enhance the adaptability and overall performance of routing protocols in underwater acoustic environments. Future research will focus on: (i) applying deep reinforcement learning to reduce training complexity and enable distributed decision-making among multiple agents in highly dynamic underwater networks; (ii) evaluating scalability with networks exceeding 100 nodes and node velocities above 1 m/s to assess convergence behavior under extreme mobility; and (iii) conducting preliminary hardware validation using acoustic modems such as the WHOI Micromodem or EvoLogics S2C to verify real-world applicability.

ACKNOWLEDGMENT

This initiative is carried out within the framework of the funds of the Recovery, Transformation and Resilience Plan, financed by the European Union (Next Generation) - National Cybersecurity Institute (INCIBE) in the project C107/23 “Artificial Intelligence Applied to Cybersecurity in Critical Water and Sanitation Infrastructures”.

REFERENCIAS

[1] Z. Wang, X. Gu, W. Xie, and D. Wu, “Qinghai mutton sales mode analysis and optimization strategy research,” *Procedia Comput. Sci.*, vol. 242, pp. 1370–1377, 2024, doi: 10.1016/j.procs.2024.08.123.

[2] K. Saleem, L. Wang, A. Almogren, E. Ntizikira, A. U. Rehman, S. Bharany, and S. Hussen, “Cognitive intelligence routing protocol for disaster management and underwater communication system in underwater acoustic network,” *Sci. Rep.*, vol. 15, no. 1, Mar. 2025.

[3] M. H. Homaei, A. J. Di Bártolo, R. Molano Gómez, P. G. Rodríguez, and A. Caro, “Enabling RPL on the Internet of Underwater Things,” *J. Netw. Syst. Manag.*, vol. 33, no. 3, May 2025.

[4] Y. Guo, J. Jiang, Q. Yan, and G. Han, “An opportunity routing protocol based on density peaks clustering in the Internet of Underwater Things,” in *Proc. Int. Conf. Intell. Commun. Netw. (ICN)*, Nov. 2023, pp. 175–179.

[5] M. Ismail, H. Qadir, F. A. Khan, S. Jan, Z. Wadud, and A. K. Bashir, “A novel routing protocol for underwater wireless sensor networks based on shifted energy efficiency and priority,” *Comput. Commun.*, vol. 210, pp. 147–162, Oct. 2023.

[6] S. Natesan and R. Krishnan, “FLCEER: Fuzzy logic cluster-based energy efficient routing protocol for underwater acoustic sensor network,” *Int. J. Inf. Technol. Web Eng.*, vol. 15, no. 3, pp. 76–101, Jul. 2020.

[7] N. Usman, O. Alfandi, S. Usman, A. M. Khattak, M. Awais, B. Hayat, and A. Sajid, “An energy efficient routing approach for IoT enabled underwater WSNs in smart cities,” *Sensors*, vol. 20, no. 15, p. 4116, Jul. 2020.

[8] Z. A. Khan, O. A. Karim, S. Abbas, N. Javaid, Y. B. Zikria, and U. Tariq, “Q-learning based energy-efficient and void avoidance routing protocol for underwater acoustic sensor networks,” *Comput. Netw.*, vol. 197, p. 108309, Oct. 2021.

[9] R. W. L. Coutinho, A. Boukerche, and A. A. F. Loureiro, “PCR: A power control-based opportunistic routing for underwater sensor networks,” in *Proc. 21st ACM Int. Conf. Modeling, Anal. Simul. Wireless Mobile Syst.*, Oct. 2018, pp. 173–180.

[10] Y. H. Robinson, S. Vimal, E. G. Julie, M. Khari, C. Expósito-Izquierdo, and J. Martínez, “Hybrid optimization routing management for autonomous underwater vehicle in the Internet of Underwater Things,” *Earth Sci. Inform.*, vol. 14, no. 1, pp. 441–456, Oct. 2020.

[11] X. Li, X. Hu, R. Zhang, and L. Yang, “Routing protocol design for underwater optical wireless sensor networks: A multiagent reinforcement learning approach,” *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9805–9818, Oct. 2020.

[12] C. S. Nandyala, H.-W. Kim, and H.-S. Cho, “QTAR: A Q-learning-based topology-aware routing protocol for underwater wireless sensor networks,” *Comput. Netw.*, vol. 222, p. 109562, Feb. 2023.

[13] İ. Eriş, Ö. M. Gül, and P. S. Bölük, “A novel reinforcement learning based routing algorithm for energy management in networks,” *J. Ind. Manag. Optim.*, vol. 20, no. 12, pp. 3678–3696, 2024.

[14] M. Tarif and B. N. Moghadam, “Proposing a dynamic decision-making routing method in underwater Internet of Things,” in *Proc. 10th Int. Conf. Artif. Intell. Robot. (QICAR)*, Feb. 2024, pp. 186–193.

[15] M. Tarif, M. Effatparvar, and B. N. Moghadam, “Enhancing energy efficiency of underwater sensor network routing aiming to achieve reliability,” in *Proc. 3rd Int. Conf. Distrib. Comput. High Perform. Comput. (DCHPC)*, May 2024, pp. 1–7.

[16] M. Tarif, M. H. Homaei, and A. Mosavi, “An enhanced fuzzy routing protocol for energy optimization in the underwater wireless sensor networks,” *Comput. Mater. Continua*, vol. 83, no. 2, pp. 1791–1820, 2025.

[17] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, “A stateless opportunistic routing protocol for underwater sensor networks,” *Wirel. Commun. Mob. Comput.*, vol. 2018, pp. 1–12, Jan. 2018.

[18] S. Ghoreyshi, A. Shahrabi, and T. Boutaleb, “A novel cooperative opportunistic routing scheme for underwater sensor networks,” *Sensors*, vol. 16, no. 3, p. 297, Feb. 2016.

[19] S. Pradeep, T. B. B. R. Bapu, R. Rajendran, and R. Anitha, “Energy efficient region-based source distributed routing algorithm for sink mobility in underwater sensor network,” *Expert Syst. Appl.*, vol. 233, p. 120941, Dec. 2023.

[20] Z. Liu, X. Jin, Y. Yang, K. Ma, and X. Guan, “Energy-efficient guiding-network-based routing for underwater wireless sensor networks,” *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21702–21711, Nov. 2022.

[21] M. H. Homaei, S. S. Band, A. Pescapè, and A. Mosavi, “DDSLA-RPL: Dynamic decision system based on learning automata in the RPL protocol for achieving QoS,” *IEEE Access*, vol. 9, pp. 63131–63148, 2021.

[22] Xie, P., Zhou, Z., Nicolaou, N., See, A., Cui, J.H., Shi, Z.: Aqua-Sim: an NS-2 based simulator for underwater sensor networks. In: Proceedings of MTS/IEEE OCEANS 2009, pp. 1-7 (2009)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

INCA-LLM: Innovative Network Slicing in 5G/6G Core Auto-Configuration with LLM

Felix Delgado-Ferro*, Ramón García-Verjaga†, Juan J. Ramos-Munoz*,
Jorge Navarro-Ortiz*, Juan M. Lopez-Soler*
* Department of Signal Theory, Telematics and Communications, University of Granada, Granada.
† Department of Computer Science and Artificial Intelligence, University of Granada, Granada.
felixdelgado@ugr.es, rgarver@correo.ugr.es, {jjramos, jorgenavarro, juanma}@ugr.es

Abstract

This work presents an initial approach to automation using Large Language Models, with the main objective of automatically configuring the 5G core for proper network slicing deployment. First, it explains the key concepts needed to understand 5G-LLM integration. Subsequently, it presents our proposal, called INCA-LLM, which uses In-Context Learning to automatically generate the configuration files for the 5G core. INCA-LLM applies few-shot prompting strategy for the first proof-of-concept tests. Preliminary results show that the generated files are suitable for deployment, with a 79.54% similarity to the expert-created configuration.

***Keywords*—Large Language Models, In-Context Learning, Few-shot prompting, Auto-Configuration, Next-Generation Networks, 5G/6G Core, Network Slicing.**

I. INTRODUCTION

Large Language Models (LLMs) have demonstrated strong capabilities in domains such as finance, law, and medicine, yet their adoption in highly technical areas like telecommunications remains limited. The configuration of Next-Generation Networks (NGNs), particularly in 5G and beyond, demands precise domain expertise and involves processes that are often manual, time-consuming, and prone to error. Even small mistakes can lead to service disruptions and significant costs, while the scarcity of qualified experts further increases the burden of network management. These challenges highlight the urgency of developing automation mechanisms capable of reducing human dependency, improving accuracy, and ensuring operational efficiency. Properly adapted, LLMs can help address these limitations by supporting configuration-intensive tasks with greater reliability.

At the same time, the complexity introduced by emerging paradigms such as Software-Defined Networking (SDN), Network Function Virtualization (NFV), Multi-access Edge Computing (MEC), and Network Slicing (NS) has rendered traditional management approaches insufficient. In this context, LLMs are positioned as enablers of self-managing networks, contributing to faster deployment, lower operational costs, and reduced human error, in line with the vision of Zero-touch network and Service Management (ZSM). Building on this perspective, this work introduces Innovative Network Slicing in 5G/6G Core Auto-Configuration with LLM (INCA-LLM), a framework for managing network slicing in industrial contexts. The system automates the generation of 5G core network configurations using data from the Amarisoft core, translating high-level user inputs into precise outputs while mitigating risks such as hallucinations and misconfigurations. By addressing one of the most resource-intensive aspects of network management, INCA-LLM contributes to more reliable, scalable, and efficient next-generation networks.

This paper is structured as follows: Section II reviews related work; Section III outlines key concepts for 5G-LLM integration; Section IV details the INCA-LLM framework; Section V presents validation and results; and Section VI concludes with insights and future directions.

II. RELATED WORKS

Large Language Models are playing an increasingly relevant role in automating NGNs. Their effectiveness depends on high-quality, domain-specific datasets, such as those based on 3rd Generation Partnership Project (3GPP) Technical Specifications (TS) [4]. Tools like 5G Instruct Forge [1] help to generate such datasets, enabling fine-tuning of models like LLaMA and Mistral for 5G tasks.

Managing NGNs, especially for network slicing, is complex due to the wide range of functionalities and the shortage of experts [5] [6]. This has led to the rise of Intent-Based Networking (IBN), where user-defined goals guide automated network management. LLMs are being explored to interpret these intentions and generate configurations, as seen in works like SliceGPT [7] and

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

LLM-xAPP [8], which apply LLMs to slice orchestration and radio resource management.

Our work builds on these ideas by focusing on automating the generation of 5G Core Network (5G-CN) configurations, specifically for the Amarisoft core. This is a key component for enabling IBN and ZSM. We emphasize accurate data collection to ensure reliable configuration generation, as errors could affect critical services. The work also presents integration methods with the 5G network and introduces our approach using In-Context Learning (ICL).

III. SYSTEM IMPLEMENTATION

A. 5G-LLM Integration Keys

The integration of LLM in a 5G network requires experience in both areas. Therefore, two key factors are broken down below: (1) LLM inference and utilization, and (2) 5G system and LLM integration.

LLM inference and utilization

Focusing on inference and utilization, note that LLMs specialize in the well-known prompt engineering. Prompt engineering is the process to generate desired outputs, specifically prompt creation requires no additional training and produces instant results based on user input. This represents a straightforward approach to using LLMs. Prompt engineering features several techniques, but in this work we focus on In-Context Learning (ICL). ICL uses formatted natural language prompts and integrates task descriptions and examples to guide LLMs in executing tasks. This approach allows LLMs to recognize and perform new tasks using contextual information. In this regard, four strategies are shown as follows:

1) **Zero-shot prompting:** Provide a clear instruction describing the task, e.g. "Generate a configuration file based on the following requirements...".
2) **Few-shot prompting:** Include 2-3 examples from the several triads directly in the prompt to show the LLM the desired input (natural language description) and output (configuration file) format. This approach is crucial for the task as it allows the model to learn the desired pattern.
3) **Role-playing:** Assign a specific role to the model, e.g. "You are an expert engineer in 5G networks specializing in core configurations. Your task is to generate a configuration file..."
4) **Output format instructions:** Explicitly request that the output conform to the configuration file structure. If possible, provide a template or schema description within the prompt to guide the model.

5G system and LLM integration

From a theoretical-practical framework, the standard 5G Service Based Architecture (SBA) is made up of two parts (RAN and CN). Additionally, there are multiple parts that can be considered within the architecture, such as Near-Real-Time RAN Intelligent Controller (Near-RT RIC) and IP Multimedia Subsystem (IMS) Core. From this point on, all of them will be considered as a 5G system. Fig. 1 shows the 5G System and also the relevant interfaces (green arrows) and IMS connections (purple elements).

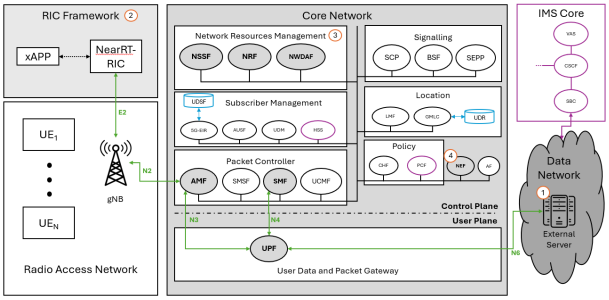


Figure 1. 5G-LLM Integration on 5G System with NearRT-RIC, 5G Core Service Based Architecture and IMS Core

In addition, 5G network functions related to network slicing management need to be well understood (highlighted and in bold). These are defined below:

- **Network Slice Selection Function (NSSF):** It selects the appropriate slice network for a User Equipment (UE) based on service requirements and policies.
- **Network Repository Function (NRF):** It maintains a repository of available network functions and their capabilities, allowing other network functions to discover and communicate with the required functions. It is essential for the management and discovery of network functions that can be used in different network slices.
- **Network Data Analytics Function (NWDAF):** It provides analysis of network data to improve efficiency and performance in real time, e.g. resources for optimisation of resource allocation of different slices.
- **Access and Mobility Management Function (AMF):** It manages the mobility and access of the UE, interacting with the NSSF to ensure that the UE connects to the appropriate network slice according to its service needs.
- **Session Management Function (SMF):** It manages UE data sessions, including session configuration within network slices, ensuring that they meet specific Quality of Service (QoS) requirements.
- **User Plane Function (UPF):** It manages user data traffic for different network slices, including routing, packet forwarding, policy enforcement and QoS management of each slice.
- **Network Exposure Function (NEF):** It exposes network capabilities and services to external and secure applications, enabling the integration of third party services.

We highlight four integration methods (marked numerically on Fig.1):

1) **External server:** It is the most versatile method due to the fact that can encompass both RAN and CN. It consists of deploying the LLM service on a server external to the 5G system. Once requested, it communicates with the 5G system to auto-configure

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

or modify parameters to optimize network slices according to the user requirements. These servers can be deployed at the Edge or Cloud.

2) **xApp:** This approach involves creating a framework that integrates an LLM as an xApp running on the NearRT-RIC of O-RAN, which is part of the RAN. Therefore, it allows for the monitoring and adaptive management of radio resources, translating high-level UE intentions into RAN configurations, such as optimizing PRBs based on UE needs. It is important to note that this method focuses solely on the RAN.
3) **Network Resource Management:** This method, focused solely on the CN, involves the implementation or enhancement within NWDAF. Essentially, the LLM receives data collected from other NFs, analyzes it, and optimizes network and resource management internally within the CN.
4) **Network Exposure:** It is similar to the external server. The main difference is that the requirements or prompt must be provided to the system via NEF.

B. INCA-LLM

Testbed

INCA-LLM is deployed on an edge server as shown in Fig. 2, as this setup offers an efficient and practical way to generate initial configurations. It also allows for future extensions to both RAN and CN. INCA-LLM focuses on Amarisoft core due to its compact configuration format, which uses fewer files compared to alternatives like open5GS or OAI. As shown in Fig. 2, the system integrates INCA-LLM with the 5G network and demonstrates a configuration with multiple slices, being the main objective of the auto-generated configuration files. In this case, the network would support three slices. The first slice is for typical internet traffic and data calls over the eMBB slice with a low priority 5QI (e.g., 5QI 9). The second slice, called URLLC, is for critical industrial traffic and will use a high priority with 5QI 90. Finally, the third slice, called mMTC, will contain the monitoring traffic, generated by multiple industrial devices, with a medium priority and 5QI 86.

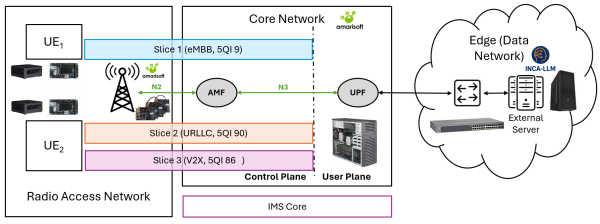


Figure 2. Integration 5G System with INCA-LLM

Fig. 2 also illustrates the components of the testbed where the system is deployed and evaluated. The system has been deployed on a testbed to evaluate its functionality and results. Primarily, the INCA-LLM service is deployed on an external server, specifically a DELL 1445 server. The 5G-Amarisoft system is implemented on a general-purpose computer equipped with two PCIe SDR50. Finally, to verify connectivity, we use two UEs consisting of an Intel NUC BXNUC10I7FNH2 paired with a Quectel RM500Q-GL board.

INCA-LLM Implementation

INCA-LLM is designed and implemented following the principles of ICL. It implements two modes. Both are based on the Few-shot prompting strategy, using the default configuration of Amarisoft and a multiple slicing configuration files that contains the necessary settings for the correct deployment of the slices. Our implementation, see Fig. 3, is based on the LLM Ollama and the model deepseek-r1:14b have been tested. These mode are:

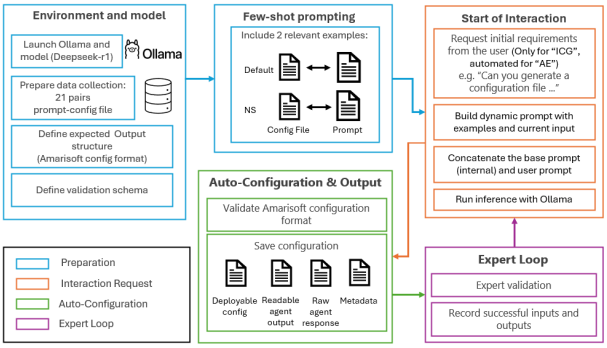


Figure 3. INCA-LLM implementation: Network slicing auto-configuration for 5G/6G CN.

1) **Interactive Configuration Generation (ICG):** This mode allows for the structured generation of CN configurations, using the two cited configurations as references while maintaining the Amarisoft core configuration format (following [9])). The Amarisoft format has a high level of complexity due to its key-value structure (similar to JSON) and C directives. Through user-model interaction, the LLM service generates a configuration based on the requirements provided in natural language.
2) **Automated Evaluation of Prompts (AEoP):** This mode consists of a systematic process designed to evaluate multiple iterations based on different prompt designs on the model, comparing their outputs with the expected configurations previously defined by an expert in 5G core configuration of Amarisoft. It works from a set of test cases, where each prompt is evaluated in parallel, and their outputs are automatically validated through structural comparison with reference files.

IV. VALIDATION & RESULTS

The INCA-LLM system is evaluated through four validation steps:

1) **Format validation (V_1):** It checks the configuration file generated by the model complies with the expected format for the Amarisoft core configuration, ensuring that the syntactic structure is valid.
2) **Expert validation (V_2):** It implies that an expert manually reviews the configurations generated by

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

the model to ensure that the values, parameters, and structure are appropriate, secure, and consistent, strictly adhering to the user's requirements.

3) **Deployment (V_3):** involves deploying the configuration in the testbed and observing if the 5G system exhibits any deployment failures. This validation ensures that the configuration is not only theoretically correct but also operationally functional.
4) **Structural similarity (V_4):** This metric quantitatively evaluates the degree of alignment between the model-generated configuration and the reference configuration provided by experts. It considers the hierarchical organization, presence of mandatory fields, and consistency of internal object relationships. Formally expressed as Eq. 1.

$$S(A,B) = \frac{2M}{|A| + |B|} * 100 \tag{1}$$

where A and B represent the normalized reference and generated configurations, M denotes the number of matching characters identified in the longest common subsequences, and $|A|$, $|B|$ correspond to their respective lengths. The resulting value $S(A,B) \in [0,100]\%$ provides a robust indicator of structural similarity, with values closer to 100 % denoting stronger alignment between the two configurations.

The results obtained are condensed in Table I, showing the validations passed and the percentage of similarity. It should be noted that the percentage of similarity is never 100% because it is a file generated by the model and the inference obtained, so there may be some differences from the reference configurations, such as simple comments within the configuration.

Table I
PRELIMINARY RESULTS OF INCA-LLM SYSTEM.

Test	V_1	V_2	V_3	V_4
T_1	✓	✓	✗	74.51%
T_2	✓	✓	✓	80.23%
T_3	✓	✓	✓	73.75%
T_4	✓	✓	✗	87.86%
T_5	✓	✓	✗	71.08%
T_6	✓	✓	✓	89.82%

The results show that 100% of the generated configurations correctly follow the Amarisoft structure. At the same time, the expert states that the parameters referred to in the prompts are correctly modified. But as we can see, there are still some parameters not specified that are configured wrongly, making 50% of the files generated for testing invalid for deployment.

V. CONCLUSION

In this paper, we introduce the integration of a 5G/6G system with an LLM service, called INCA-LLM. The proposed integration consists of deploying the service at the edge with the aim of automatically generating configuration files for the network core, focusing mainly on network slice deployments. INCA-LLM is based on ICL, specifically on the few-shot prompting strategy. Our deployment and evaluations show that INCA-LLM, without the need for training (fine-tuning) and specialized in Amarisoft, correctly generates the files according to user specifications, demonstrating significant structural similarities with the files provided by the 5G network configuration expert (approximately 79.54%). This work advances the use of LLMs in telecommunications and provides a system that can be extended to other software such as OAI or open5GS. Looking ahead, our research has important implications for the development of future networks like 6G, which will rely on technologies such as zero-touch management systems.

The research lines to be followed involve the extension of LLMs for the generation of Amarisoft files corresponding to the rest of the 5G system (gNB, DRB data base, IMS Core and UE data base). Subsequently, the addition of auto-configuration without the need for expert review will be implemented, and finally, monitoring and new LLM agents will be deployed as specified in the integration methods to optimize the network.

ACKNOWLEDGMENT

This work has been financially supported in part by the Spanish Ministry of Science and Innovation through the PID2022-137329OB-C43 project funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU; and in part by the Spanish Ministry of Universities (Grant FPU20/02621).

REFERENCES

[1] A. Idir Ait Said, A. Mekrache, K. Boutiba, K. Ramantas, A. Ksentini and M. Rahmani, "5G INSTRUCT Forge: An Advanced Data Engineering Pipeline for Making LLMs Learn 5G," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 2, pp. 974-986, April 2025.

[2] Y. Xia, J. Zhang, N. Jazdi and M. Weyrich, "Incorporating Large Language Models into Production Systems for Enhanced Task Automation and Flexibility," *Preprint Arxiv*, July 2024.

[3] C. Benzaid and T. Taleb, "AI-Driven Zero Touch Network and Service Management in 5G and Beyond: Challenges and Research Directions," *IEEE Network*, vol. 34, no. 2, pp. 186-194, March/April 2020.

[4] 3GPP, "The Mobile Broadband Standard," 3GPP, 2025. [Online]. Available: https://www.3gpp.org/. [Accessed: 26- May- 2025].

[5] A. Mekrache and A. Ksentini, "LLM-enabled Intent-driven Service Configuration for Next Generation Networks," in *2024 IEEE 10th International Conference on Network Softwarization (NetSoft)*, Saint Louis, MO, USA, pp. 253-257, 2024.

[6] D. M. Manias, A. Chouman and A. Shami, "Towards Intent-Based Network Management: Large Language Models for Intent Extraction in 5G Core Networks," in *2024 20th International Conference on the Design of Reliable Communication Networks (DRCN)*, Montreal, QC, Canada, pp. 1-6, 2024.

[7] E. Bandara *et al.*, "SliceGPT – OpenAI GPT-3.5 LLM, Blockchain and Non-Fungible Token Enabled Intelligent 5G/6G Network Slice Broker and Marketplace," in *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2024, pp. 439-445, 2024.

[8] X. Wu, J. Farooq, Y. Wang and J.Chen, "LLM-xApp: A Large Language Model Empowered Radio Resource Management xApp for 5G O-RAN," in *Workshop on Security and Privacy of Next-Generation Networks (FutureG)*, 2025.

[9] Amarisoft, "Overall Architecture of Amari Callbox & Configuration Guide," Amarisoft, 2025. [Online]. Available: https://tech-academy.amarisoft.com/ConfigurationGuide.html [Accessed: 08-Sep - 2025]

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Vulnerabilidades en Redes Intra-Vehiculares Modernas: Un Análisis de Ataque de Denegación de Servicio sobre Automotive Ethernet

Isabel Bravo Sánchez, Adrián Ortega Escalona, Gines Domenech-Asensi, Juan Carlos Aarnoutse-Sanchez, Maria-Dolores Cano
Departamento de Tecnologías de la Información y las Comunicaciones (TIC),

Universidad Politécnica de Cartagena

TRUST Lab, Edif. ELDI O3-10, 30202 Cartagena

isabel.bravo@upct.es, adrian.ortega@upct.es, gines.domenech@upct.es, juanc.sanchez@upct.es, mdolores.cano@upct.es

El avance hacia vehículos definidos por software y la conducción autónoma ha impulsado la adopción de Automotive Ethernet como la red troncal de comunicación intra-vehicular debido a su alto ancho de banda y flexibilidad. Sin embargo, esta conectividad incrementada introduce nuevas superficies de ataque, siendo los ataques de Denegación de Servicio (DoS) una amenaza significativa para la disponibilidad y seguridad de las funciones vehiculares. Este trabajo tipo *work in progress* presenta un análisis de las vulnerabilidades de DoS en redes Automotive Ethernet, discutiendo arquitecturas de red, vectores de ataque comunes y una metodología para su simulación y evaluación utilizando herramientas de simulación como Vector CANoe. Se exploran las métricas clave para cuantificar el impacto de dichos ataques, sentando las bases para el desarrollo futuro de contramedidas robustas.

Palabras Clave: Ciberseguridad Automotriz, Automotive Ethernet, Denegación de Servicio (DoS), Redes Intra-Vehiculares, Vector CANoe, Simulación de Ataques.

I. Introducción

La arquitectura electrónica de los vehículos modernos ha experimentado una transformación radical. De sistemas aislados y funcionalmente limitados, hemos pasado a complejos ecosistemas interconectados que gestionan desde el info-entretenimiento hasta funciones críticas para la seguridad como los Sistemas Avanzados de Asistencia al Conductor (*Advanced Driver Assistance Service*, ADAS) y la conducción autónoma [1]. Esta evolución ha sido posible gracias a la diversificación y especialización de las redes intra-vehiculares.

Históricamente, redes como *Controller Area Network* (CAN), *Local Interconnect Network* (LIN), FlexRay y *Media Oriented Systems Transport* (MOST) han sido el pilar para el control del tren motriz y chasis (véase la Tabla I). LIN ha gestionado sensores y actuadores de bajo coste. FlexRay ofreció mayor ancho de banda y determinismo para sistemas de seguridad activa. Por último, MOST se dedicó a aplicaciones multimedia [1]. Si bien estas redes persisten, la creciente demanda de ancho de banda, baja latencia y conectividad IP ha impulsado la adopción de Automotive Ethernet [2-3]. A diferencia de la Ethernet tradicional, diseñada para entornos informáticos con cableado más voluminoso y sin restricciones de tiempo real, Automative Ethernet se adapta a las exigencias del vehículo moderno mediante enlaces punto a punto, cableado reducido y transmisión determinista. Estándares como IEEE 802.3bw (100BASE-T1) e IEEE 802.3bp (1000BASE-T1), junto con extensiones para *Time Sensitive Networking* (TSN), lo posicionan como la red troncal para dominios como ADAS, info-entretenimiento avanzado y actualizaciones *Over-The-Air* (OTA).

A pesar de sus ventajas, la transición a Automotive Ethernet y su interconexión con *legacy networks* a través de *gateways* exponen al vehículo a ciber amenazas bien conocidas en el mundo de Tecnologías de la Información (TI), entre las que destacan los ataques de Denegación de Servicio (*Denial of Service*, DoS). Un ataque DoS busca agotar los recursos de un sistema o red, impidiendo que los usuarios legítimos accedan a servicios esenciales [4-5]. En un vehículo, un DoS efectivo podría degradar o deshabilitar funciones críticas, comprometiendo la seguridad.

Este *work-in-progress* presenta una metodología de simulación reproducible en el software Vector CANoe [6] para cuantificar el impacto de un DoS de inundación sobre una red Ethernet a 1 Gbps intra-vehicular. El escenario incluye ocho Unidades de Control Electrónico (*Electronic Control Units*, ECUs) y un nodo atacante que, mediante

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Communication Access Programming Language (CAPL), genera ráfagas de 1500 bytes alcanzando un 99 % de carga de bus y 8 k pkt/s, lo que demuestra la degradación extrema de la disponibilidad. Esperamos que el marco experimental (métricas, scripts y configuraciones) sirva como base para investigaciones futuras en contramedidas y que permita, en siguientes etapas, validar la resiliencia de arquitecturas reales.

El resto del artículo se organiza de la siguiente forma. La Sección II describe la metodología, poniendo en contexto los distintos vectores de ataque de DoS, incluyendo la configuración para simulación del ataque en Vector CANoe y definiendo las métricas de impacto. La Sección III presenta y analiza los resultados, comparando el estado de red antes y después del ataque. El artículo finaliza destacando las conclusiones principales y el trabajo futuro.

Tabla I
TABLA DE TECNOLOGÍAS DE COMUNICACIÓN VEHICULAR

Red	Característica Principal
CAN	Robusta, económica, control distribuido
LIN	Bajo costo, simplicidad (Maestro/Esclavo)
FlexRay	Determinista (TDMA), tolerante a fallos
MOST	*Streaming* A/V síncrono, QoS para multimedia
Automotive Ethernet	Alto ancho de banda, IP nativo, flexible y escalable

II. METODOLOGÍA

Para analizar el impacto de los ataques DoS sobre Automotive Ethernet, se propone una metodología basada en simulación permitiendo la creación de escenarios controlados y evaluación sistemática de vulnerabilidades.

A. Arquitectura de Red Intra-Vehicular y Automotive Ethernet

Las arquitecturas vehiculares modernas son heterogéneas, combinando múltiples tecnologías de red (Fig. 1). Automotive Ethernet suele actuar como una red troncal (*backbone*) de alta velocidad, interconectando diferentes dominios funcionales (ej. ADAS, *infotainment*, telemática) y *gateways* que, a su vez, comunican con redes legadas como CAN o LIN. Dentro del dominio de Automotive Ethernet, se emplean switches para segmentar el tráfico y conectar múltiples ECUs.

Los estándares IEEE para Automotive Ethernet (ej. 100BASE-T1, 1000BASE-T1) definen la capa física, mientras que protocolos como TCP/IP y UDP/IP son comúnmente utilizados para la comunicación de datos. *Scalable service-Oriented MiddlewarE over IP* (SOME/IP) es frecuentemente usado para la comunicación orientada a servicios [2].

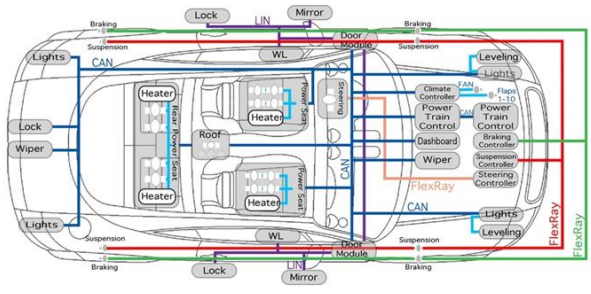


Fig. 1. Ejemplo de cooperación entre las distintas redes vehiculares (antes de Ethernet).

Las características de seguridad inherentes a Automotive Ethernet, como MACsec (IEEE 802.1AE) para la seguridad en la capa de enlace, y el uso de firewalls y Sistemas de Detección de Intrusos (*Intrusion Detection Systems*, IDS) en *gateways* y ECUs centrales, son cruciales, pero no infalibles contra todos los tipos de ataques DoS [4].

B. Vectores de Ataque de Denegación de Servicio

Los ataques DoS buscan saturar la red o los recursos de las ECUs. En el contexto de Automotive Ethernet, algunos vectores relevantes incluyen: ataques de inundación (*Flood Attacks*), ataques de amplificación/reflexión y ataques a protocolos específicos.

Dentro de los flood attacks, tenemos los ataques por inundación SYN (SYN *Flood*). El atacante envía un gran número de paquetes TCP SYN, a menudo con direcciones IP de origen falsificadas, para agotar los recursos de conexión del servidor (ECU) [3]. También están los ataques de inundación UDP (UDP *Flood*), donde se envían grandes cantidades de paquetes UDP a puertos aleatorios o específicos de la ECU víctima, consumiendo ancho de banda de red y recursos de procesamiento de la ECU [3]. Por último, los de inundación ICMP (ICMP *Flood*) son similares a la inundación UDP, pero utilizando paquetes ICMP.

Respecto a los ataques de amplificación/reflexión, el atacante envía solicitudes a servicios de red de terceros con la dirección IP de origen falsificada como la de la víctima. Las respuestas, a menudo más grandes que las solicitudes, se dirigen a la víctima, amplificando el ataque.

Por su parte, los ataques a protocolos específicos incluyen la explotación de vulnerabilidades en SOME/IP o su mecanismo de descubrimiento de servicios (SOME/IP-Service Discovery) para generar tráfico excesivo o interrumpir la comunicación legítima [5]. También en esta categoría estaría la manipulación de tráfico TSN, como la alteración de la frecuencia de mensajes o la manipulación de los mecanismos de control de puertas (*gate control lists*), para degradar la calidad de servicio de flujos críticos [5].

C. Herramienta de Simulación: Vector CANoe

Vector CANoe es una herramienta de desarrollo, simulación y prueba ampliamente utilizada en la industria automotriz para ECUs y redes completas. Utilizamos Vector CANoe Desktop Edition con un modelo de licencia por suscripción basado en dispositivo, de carácter académico, conforme al End User License Agreement.

Su opción Ethernet proporciona capacidades extensivas para:

- Simular redes y ECUs Ethernet, incluyendo stacks TCP/IP individuales.
- Soportar diversos protocolos de Automotive Ethernet como SOME/IP, DoIP y los protocolos de transporte TCP/UDP.
- Interactuar con hardware de interfaz Ethernet de Vector (familia VN5000) que soporta múltiples capas físicas (10BASE-T1S, 100BASE-T1, 1000BASE-T1).
- Generar y manipular tráfico Ethernet utilizando el lenguaje de programación CAPL (Communication Access Programing Language) o .NET, permitiendo la creación de scripts de ataque DoS.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- Monitorizar y analizar el tráfico de red en detalle (*Trace Window*, análisis de protocolos).

D. Configuración de la Simulación del Ataque DoS

Para el análisis propuesto en este trabajo, se configura un escenario de simulación en CANoe que incluye:

- Red Simulada: Una red Automotive Ethernet con ECUs críticas simuladas (ej. una ECU de ADAS que procesa datos de sensores, una ECU de gateway y otras 5 ECUs adicionales).
- Nodo Atacante: Una ECU simulada o un script CAPL configurado para generar tráfico DoS (ej. inundación de paquetes Ethernet) dirigido a las otras ECUs.

Como se aprecia en la Figura 2, el escenario de simulación en Vector CANoe representa una red Ethernet vehicular compuesta por múltiples ECUs, así como un nodo atacante y un switch central que actúa como *backbone*. En la parte superior se visualizan las capas de configuración del sistema, mientras que en la inferior se representan los módulos funcionales y sus conexiones. Este entorno permite modelar la interacción entre componentes reales y simulados. De la misma forma, en el Algoritmo 1 se muestra un procedimiento simplificado de ataque DoS mediante envío masivo de tramas Ethernet. A partir de parámetros como las direcciones MAC de origen y destino, tamaño de paquete, número de paquetes por ráfaga e intervalo de envío, se genera un bucle que transmite ráfagas sucesivas, reproduciendo así un escenario de sobrecarga que compromete la disponibilidad de la red.

E. Métricas de Evaluación del Impacto

Para cuantificar el impacto del ataque DoS, se observará en las simulaciones las siguientes métricas de rendimiento de la red y de la ECU:

- Carga del Bus de Transmisión (*Tx Bus Load*) (%): Porcentaje de la capacidad del enlace (1 Gpbs) utilizado por el tráfico de transmisión generado por la fuente de ataque.
- Tasa de Transmisión de Paquetes (*Tx Packet Rate*)(pkt/s): Número de paquetes enviados por segundo desde la interfaz que origina el paquete.
- Tasa de Transmisión de Bits (*Tx Bit Rate*) (Mbit/s): La cantidad efectiva de datos (bits) transmitidos por segundo desde la interfaz que origina el paquete.
- Volumen Total de Datos Transmitidos (*Tx Bytes Total*) (byte): Cantidad acumulada de bytes enviados desde la interfaz que origina el ataque durante la simulación.

Fig. 2. Escenario de Simulación en Vector CANoe.

Algoritmo 1: Ataque DoS en CAPL (versión compacta).

```
1.  Procedimiento DoS (srcMAC, dstMAC, pkgSize, burstCount,
    interval)
2.      // Ajuste de pkgSize al rango [64, 1500]
3.      pkgSize ← max(64, min (pkgSize, 1500))
4.      Mientras ataqueActivo hacer
5.          for i ← 1 to burstCount hacer
6.              Enviar (Ethernet (srcMAC, dstMAC, pkg-Size));
                Pausa(interval)
7.              Fin for
8.      Fin Mientras
9.  Fin Procedimiento
```

III. Resultados y Discusión

Los resultados obtenidos mediante simulaciones en Vector CANoe evidencian el impacto significativo de un ataque de DoS sobre una red Automotive Ethernet. Para evaluar ese impacto, se llevó a cabo un análisis comparativo de la interfaz de la red antes y durante la materialización del ataque, utilizando un conjunto de métricas clave que nos proporciona el propio simulador, las cuales hemos mencionado en la sección anterior.

La Tabla II presenta un resumen comparativo de las métricas de rendimiento clave observadas en la interfaz física Ethernet Eth1, contrastando el funcionamiento normal de la red con un escenario bajo un ataque DoS. Para la obtención de estos datos, se utilizó la herramienta Vector CANoe en modo de acceso basado en canal (*Channel-Based Access*). Este enfoque nos permite una monitorización directa del tráfico y las estadísticas a nivel del canal físico, reflejando así el estado real de 'la red en sí' a través de Eth1. Aunque las métricas presentadas son específicas del canal, la integridad de Eth1 es crucial para ECUs criticas como ADAS u otras funciones de asistencia (por ejemplo, Cámara Frontal (CAMF) y Radar de Largo Alcance (LRR)); por consiguiente, la degradación de este canal, detallada en la tabla, afecta directamente la operatividad de dichos sistemas.

Tabla II
Comparativa de Métricas en Estado Normal y Bajo Ataque

Métrica	Estado Normal (Eth 1)	Bajo Ataque DoS (Eth 1)
Tx Bus Load %	0,00	99,12
Tx Packet Rate (pkt/s)	0	8120
Tx Bite Rate (Mbit/s)	0	99,129
Tx Bytes Total (bytes)	0	4.959.500

Después de alrededor de unos 30 segundos de duración del ataque, tenemos que los resultados muestran que, durante el escenario normal, la interfaz Eth1 muestra una actividad de transmisión prácticamente nula, en la que no tenemos actividad. La carga del bus de transmisión se mantiene en 0,00 % y consecuentemente, tanto la tasa de transmisión de paquetes (0 pkt/s) como la tasa de transmisión bits (0,00 Mbit/s) son cero. Esto nos indica que, en ausencia del ataque, Eth1, no está generando tráfico de salida o este es insignificante. Por el contrario, durante el escenario bajo ataque DoS, la interfaz Eth1 se convierte en una fuente masiva de tráfico como podemos ver en la Figura 3, la cual observamos como por nuestro canal se está enviando multitud de paquetes cada milisegundo, todos enviados por nuestra

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

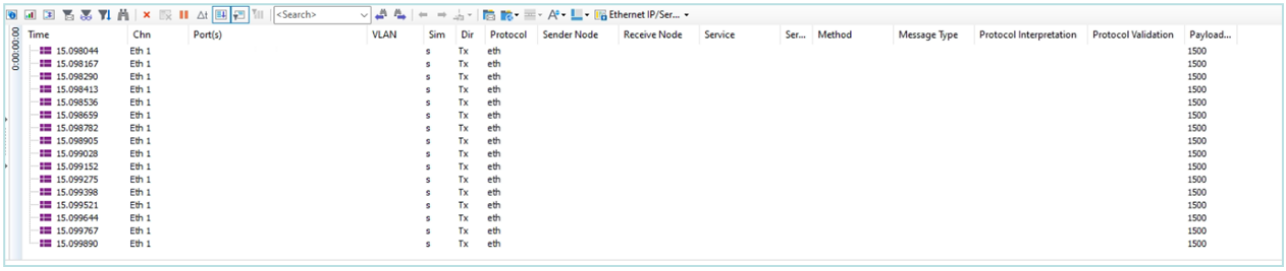


Fig. 3. Ataque DoS observado en ventana *Trace*.

ECU atacante, la cual envía como se observa en la parte derecha de la imagen un tamaño de paquetes de 1500 bytes. Además de observar la ventana de *Trace*, CANoe incluye una consola donde ofrece una visión interna del mecanismo de generación del ataque DoS, en la figura 4, presentando un extracto del registro de la consola emitido por el programa CAPL. Cada entrada en la columna 'Message' indica el total de tramas enviadas hasta el momento por el script de ataque.

- La Carga del Bus de Transmisión (*Tx Bus Load*) asciende a 99,12 %, lo que significa que el ataque está utilizando casi la totalidad de la capacidad de transmisión del enlace de 1 Gbps Ethernet.
- La Tasa de Transmisión de Paquetes (*Tx Packet Rate*) se dispara a 8120 paquetes por segundo. Esta alta frecuencia de envío de paquetes es característica la característica fundamental de un ataque de inundación (*flood attack*), diseñado para sobrecargar tanto la red como los nodos receptores, los cuales se quedarían sin sen apenas funcionales.
- Consistentemente, la Tasa de Transmisión de Bits (*Tx Bit Rate*) alcanza los 99,129 Mbit/s, corroborando que la alta carga del bus se debe a un flujo continuo y denso de datos maliciosos.
- En cuanto al volumen total, durante el periodo de observación del ataque, Eth1 transmitió 3250 paquetes, sumando un total de 4.959.500 bytes (aproximadamente 4,96 MB) de datos. Estas cifras cuantifican la magnitud del tráfico inyectado por el ataque.

En conjunto, estas métricas demuestran de manera inequívoca que el ataque DoS simulado transforma la interfaz Eth1 en un potente generador de tráfico de inundación. Este comportamiento satura el canal de comunicación desde su origen, lo que previsiblemente tendría consecuencias severas para la disponibilidad y el rendimiento de cualquier ECU servicio que depende de la comunicación a través de este segmento de la red Automotive Ethernet. La capacidad de la red para transmitir información crítica, como la de los sistemas ADAS que dependen de Eth1, se vería comprometida directamente por esta actividad maliciosa.

IV. Conclusiones

La rápida adopción de Automotive Ethernet como columna vertebral de los vehículos definidos por software convierte la disponibilidad de su *backbone* en un requisito de seguridad funcional tan decisivo como la integridad o la con-

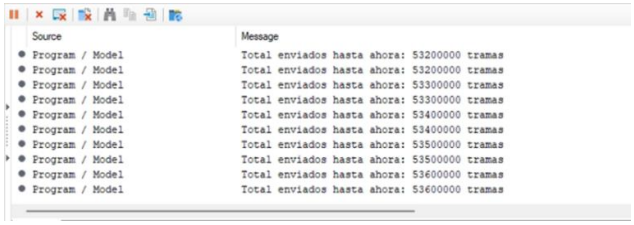


Fig. 4. Registro de la consola de Vector CANoe.

fidencialidad. En este contexto, la simulación realizada en Vector CANoe demuestra que una ráfaga de tramas de 1500 bytes puede elevar la red intra-vehicular de 1 Gbps al 99 % de carga, degradando la disponibilidad de las ECU esenciales y evidenciando la severidad del vector DoS. Este banco de ensayo constituye el primer paso de una línea de investigación destinada a caracterizar y mitigar este y otros tipos de ataque en Automotive Ethernet. Los trabajos futuros profundizarán en la caracterización estadística de los fenómenos observados y en el diseño de mecanismos de defensa que refuercen la resiliencia de las redes vehiculares de próxima generación.

Agradecimientos

Este trabajo forma parte del proyecto Laboratorio de I+D+i en Ciberseguridad, Privacidad y Comunicaciones Seguras (TRUST Lab), financiado por la Unión Europea NextGeneration-EU, Plan de Recuperación, Transformación y Resiliencia, a través del INCIBE.

Referencias

[1] W. Zeng, M. A. S. Khalid and S. Chowdhury, "In-Vehicle Networks Outlook: Achievements and Challenges," in IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1552-1571, thirdquarter 2016, doi: 10.1109/COMST.2016.2521642.

[2] K. Matheus and T. Königseder, "Automotive Ethernet", 3rd ed., Cambridge, U.K.: Cambridge Univ. Press, 2021.

[3] National Instruments. "Automotive Ethernet Overview." National Instruments. [Online]. Available: https://www.ni.com/docs/en-US/bundle/ni-xnet/page/summary-of-the-ethernet-standard.html. [Accessed: Jun. 5, 2025].

[4] H. Goswami, "A framework for detecting and preventing DoS attacks in automotive Ethernet switches," M.Sc. thesis, University of Twente, Enschede, The Netherlands, 2023.

[5] A. Martínez-Cruz, K. A. Ramírez Gutiérrez, C. Feregrino-Uribe, and A. Morales-Reyes, "Security on in-vehicle communication protocols: Issues, challenges, and future research directions," *Computer Communications*, vol.180, pp. 1-20, 2021. Doi: 10.1016/j.comcom.2021.08.027

[6] Vector. "CANoe.Ethernet." Vector. [Online]. Available: https://www.vector.com/int/en/products/products-a-z/software/canoe/option-ethernet/. [Accessed: Jun. 5, 2025]

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Propuesta de implementación de catálogos federados para espacios de datos sobre DataHub

Carlos Aparicio de Santiago, Pablo Viñuales Esquinas, Irene Plaza Ortiz,
Andres Munoz-Arcentales, Gabriel Huecas, Joaquín Salvachúa, Enrique Barra
Information Processing and Telecommunications Center, Universidad Politécnica de Madrid.
ETSI Telecomunicación. Avda. Complutense 30, Madrid 28040, Spain.
ce.aparicio@upm.es, p.vinualese@upm.es, irene.plaza.ortiz@alumnos.upm.es, joseandres.munoz@upm.es,
gabriel.huecas@upm.es, joaquin.salvachua@upm.es, enrique.barra@upm.es

En la era digital, los espacios de datos emergen como ecosistemas clave para el intercambio seguro y controlado de información entre participantes. Para ello, son esenciales componentes como los catálogos de metadatos y los conectores de espacios de datos. Este documento propone una implementación y solución de integración para ambas piezas, considerando las directrices de estandarización en formatos de datos, metadatos y protocolos, lo que garantiza la interoperabilidad.

Se presenta una solución híbrida: DataHub se utiliza como catálogo federado para una gestión robusta de metadatos, aprovechando sus capacidades avanzadas de ingesta, gobernanza y linaje. Por otro lado, la implementación propia, Rainbow Catalog, gestiona las políticas ODRL de acceso y uso. Esta integración permite consultar *datasets* de DataHub y asociarles políticas ODRL, facilitando los flujos de negociación y transferencia definidos por el Dataspace Protocol. El resultado es un sistema que combina la potencia de DataHub para la catalogación a gran escala con la gestión de políticas del conector, clave para la soberanía y confianza en los espacios de datos.

***Palabras Clave*—espacios de datos, Catálogos federados, DataHub, Conector de espacios de datos.**

I. INTRODUCCIÓN

La digitalización acelerada de las últimas décadas ha transformado profundamente todos los sectores de la sociedad, incluida la economía. Esto ha impulsado la consolidación de los datos como recurso esencial, tanto en contextos profesionales como en la vida cotidiana.

En respuesta a esta transformación, a principios de los años 2000 surgió el concepto de Open Data, promovido por organismos como la Unión Europea, con el objetivo de facilitar el acceso público a la información y reforzar la transparencia institucional [1]. Desde entonces, se han creado numerosos portales de datos abiertos [2], que actúan como repositorios con funciones de búsqueda y filtrado [3]. Sin embargo, la adopción de estos portales ha estado frenada por obstáculos como la publicación de datos incompletos, formatos no interoperables, metadatos deficientes y la escasa participación del sector privado [4], [5].

Los espacios de datos surgen como una alternativa a estos entornos. Dichos espacios de datos pretenden la generación de valor alrededor del dato mediante su compartición en un entorno de soberanía, confianza digital y seguridad [6], [7]. Estos espacios de datos están construidos sobre principios de descentralización, interoperabilidad, gobernanza y transparencia [8].

Para cumplir estos requisitos, se propone el uso de tecnologías emergentes como la identidad autosoberana (SSI) [9][10], apoyada en estándares del W3C como los *Decentralized Identifiers* (DID) y las *Verifiable Credentials* (VC). Asimismo, se emplean protocolos comunes como el *Dataspace Protocol* [11]. Para definir políticas de acceso y uso, se recurre al modelo ODRL 2.2 [12], y para la descripción de recursos y gestión de catálogos, al vocabulario DCAT3 [13][14]; ambos son recomendaciones oficiales del W3C. Además, los espacios de datos requieren infraestructuras avanzadas que habiliten no solo el intercambio de datos, sino también servicios de análisis, inferencia y compartición segura de datos personales, así como aplicación de políticas de control de acceso y uso que estén vinculadas a los activos digitales.

A diferencia de los entornos de datos abiertos, los espacios de datos no implican acceso libre e ilimitado a la información. En este nuevo paradigma, el acceso está sujeto a acuerdos explícitos entre proveedores y consumidores de datos, lo que requiere mecanismos de negociación de condiciones de uso. Una vez alcanzado un acuerdo, se formaliza mediante un contrato que define los términos de la transferencia. Este proceso es gestionado por un componente técnico denominado conector de espacios de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

datos (*Dataspace Connector*) [15].

Adicionalmente, el ecosistema requiere de un catálogo de metadatos, encargado de almacenar, publicar y facilitar el acceso a los conjuntos de datos disponibles. También requiere, a su vez, la simplificación y automatización de ingesta de metadatos y conexión con sistemas finales de datos. Este catálogo actúa como punto de descubrimiento, permitiendo a los proveedores registrar sus recursos y a los consumidores localizarlos, explorarlos o acceder a ellos, mediante funcionalidades avanzadas de búsqueda y agregación. Además, permite definir las políticas de uso y acceso asociadas a dichos servicios de datos para tener un entorno de gobernanza. [16]

Este artículo propone una arquitectura técnica para un catálogo federado en espacios de datos, integrando la plataforma DataHub con el conector interoperable Rainbow. El modelo híbrido resultante aprovecha DataHub para la gestión avanzada de metadatos y Rainbow para el control de acceso mediante políticas ODRL. Se analiza cómo el uso de estándares como DCAT3 garantiza la interoperabilidad del ecosistema, abordando así los principales retos de la federación y gobernanza de datos.

II. CATÁLOGO FUENTE Y CATÁLOGO FEDERADO

Los catálogos constituyen un componente fundamental en la arquitectura de los espacios de datos, al encargarse de la gestión y exposición de conjuntos de datos enriquecidos con metadatos. Actúan como punto de encuentro entre proveedores, que publican sus recursos, y consumidores, que acceden a ellos, los consultan y los tratan bajo determinadas condiciones.

Con todo, es posible distinguir dentro de los espacios de datos dos tipos complementarios de catálogos: catálogos fuente y catálogos federados [17].

En primer lugar, los catálogos fuente, también llamados catálogos de origen, son gestionados por los propios proveedores de datos. En ellos se publican *datasets*, acompañados de metadatos que describen características clave (por ejemplo, el formato, la frecuencia de actualización o las condiciones de acceso). Esta información es esencial, ya que consigue enriquecer los conjuntos de datos de modo que los consumidores puedan evaluar la relevancia y adecuación de los recursos antes de acceder a los mismos.

A medida que crece el número de proveedores y catálogos individuales, se hace necesaria la existencia de catálogos federados. Estas plataformas permiten integrar y agrupar los recursos publicados en múltiples catálogos fuente, ofreciendo a los usuarios un punto de entrada unificado al ecosistema. Esto no solo facilita la búsqueda y el descubrimiento de datos, sino que también amplía el alcance de los proveedores y fomenta un entorno más accesible y eficiente. [18]

III. CONECTOR DE ESPACIOS DE DATOS

Aunque todavía no existe una definición plenamente consensuada a nivel de arquitectura de lo que constituye un Espacio de Datos, hay esfuerzos por parte de organismos de estandarización y grupos de trabajo de UNE, CEN-CENELEC para definirla y se están haciendo contribuciones a los estándares W3C, IETF y ETSI en la misma línea. No obstante, la mayoría de enfoques coinciden en señalar al conector como uno de sus componentes esenciales. Esta pieza cumple un papel central en la arquitectura del ecosistema, ya que actúa como el mecanismo que habilita la interoperabilidad entre participantes. En concreto, el conector es responsable de gestionar los procesos de negociación de políticas de acceso y uso, así como de transferencia de datos, garantizando que estas operaciones se realicen bajo marcos comunes y estandarizados. [15]

Además de garantizar la interoperabilidad entre actores, los conectores deben incorporar mecanismos de autenticación y autorización que se alineen con los principios de identidad soberana (SSI). Para ello, se pueden emplear sistemas basados en *Verifiable Credentials* y *Verifiable Presentations*, implementados a través de protocolos como OpenID4VCI [19], OpenID4VP [20], DCP [21] o GNAP4VP [10]. Estos protocolos permiten verificar la identidad de los participantes controlando las credenciales presentadas, garantizando un marco de confianza digital.

Los conectores de espacios de datos manejan los cuatro tipos de actores ilustrados en la Figura 1:

- *Dataspace Provider*: componente software responsable de exponer el catálogo de metadatos, gestionar el acceso a los datos, y ejecutar los procesos de negociación y transferencia según estándares de interoperabilidad. Además, se conecta con los sistemas finales que alojan los datos reales.
- *Dataspace Consumer*: contraparte del proveedor, implementa los protocolos necesarios para negociar y consumir los datos conforme a las políticas establecidas.
- *Sistemas finales de datos*: fuentes donde residen los datos, accesibles únicamente por el *Dataspace Provider*.
- *Clientes de datos*: procesos o aplicaciones que acceden a los datos a través del Dataspace Consumer.

Además de los actores funcionales, el ecosistema también contempla entidades de gobernanza distribuida y autoridades de confianza digital, que no se abordan en este trabajo. En términos de arquitectura, el cliente de datos y el *Dataspace Consumer* operan dentro de una misma infraestructura y bajo un marco de gobernanza común, al igual que el *Dataspace Provider* y el sistema final de datos. Estos dos dominios —aislados entre sí— se comunican mediante el conector, que orquesta los flujos de autenticación, negociación de políticas y transferencia de datos. En conjunto, el *Dataspace Provider*, el *Dataspace Consumer* y sus mecanismos de interacción constituyen lo que se conoce como un conector de espacio de datos.

IV. PAPEL DEL CATÁLOGO EN EL CONECTOR DE ESPACIO DE DATOS

Como se explicó en el epígrafe anterior, uno de los requisitos del actor *Dataspace Provider* es exponer un

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

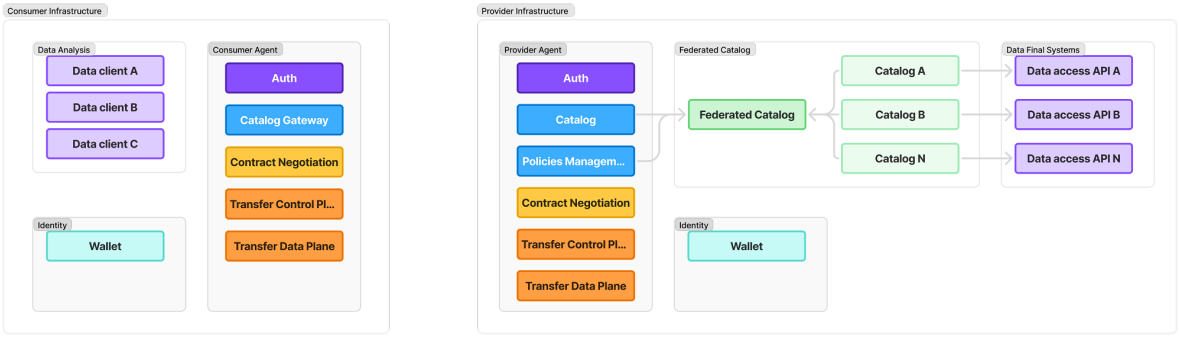


Fig. 1. Diagrama de contexto de un conector de espacios de datos.

catálogo de metadatos. Esto resulta esencial, ya que habilita los flujos clave del ecosistema: la negociación de contratos y la negociación de transferencia de datos. Ambos procesos están definidos en el *Dataspace Protocol* y requieren que los recursos y sus políticas asociadas estén previamente publicados en el catálogo.

A. *Flujo de negociación de contrato*

El proceso de negociación de contrato se modela como una máquina de estados basada en el intercambio de mensajes entre el *Dataspace Consumer* y el *Dataspace Provider* (como se muestra en la Figura 2). Este flujo puede iniciarse por cualquiera de las partes, mediante mensajes como *ContractRequestMessage* o *ContractOfferMessage*. A través de sucesivos intercambios, se alcanza un acuerdo formalizado o contrato (también denominado *Agreement*), que definirá las condiciones de acceso y uso del recurso.

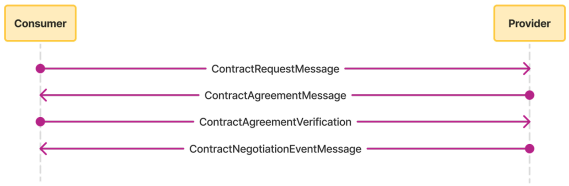


Fig. 2. Flujo de negociación de contrato en relación a ODRL

Estos mensajes intercambiados durante la negociación del contrato contienen referencias a políticas expresadas en ODRL. ODRL (*Open Digital Rights Language*) es un lenguaje de expresión de políticas que permite describir condiciones de acceso y uso de datos con un alto nivel de granularidad. Sin profundizar en el modelo, es importante señalar que cada política hace referencia explícita al recurso al que se aplica, el cual debe estar previamente registrado en el catálogo de metadatos.

De este modo, la existencia del catálogo no solo es necesaria para la identificación del recurso, sino también para validar la coherencia entre lo negociado y lo publicado. El contrato resultante, por tanto, queda vinculado al recurso definido en el catálogo, así como a los participantes del acuerdo: el proveedor (*Dataspace Provider*) y el consumidor (*Dataspace Consumer*). Actualmente, existen esfuerzos dedicados a mejorar y extender el lenguaje ODRL 2.2 para que se adapte a las necesidades específicas de los espacios de datos [22] y se está contribuyendo en la especificación de ODRL 3.0.

B. *Flujo de negociación de transferencia*

El *Transfer Process Protocol*, representado en la Figura 3, describe la negociación para la transferencia efectiva del recurso. Este flujo también se estructura como una máquina de estados. En este caso, el *Dataspace Consumer* inicia el proceso con un *TransferRequestMessage*, que hace referencia al identificador del contrato previamente alcanzado, y al tipo de distribución deseada. Una vez

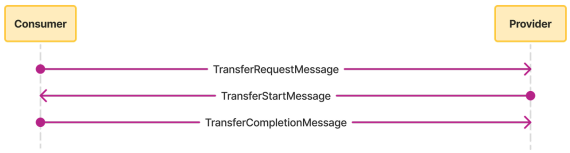


Fig. 3. Flujo de negociación de transferencia en relación a ODRL.

verificado, el sistema genera un punto de acceso a los datos, utilizando la información registrada en el catálogo.

V. CAPTURA DE REQUISITOS

La definición de requisitos para un catálogo federado en el contexto de los espacios de datos ha sido abordada ampliamente en la literatura reciente [23]. Existen, además, referencias relevantes en materia de estandarización de metadatos, gestión del linaje de datos y otros aspectos técnicos derivados de experiencias prácticas de implementación e integración. Dichos requisitos se definen en la Tabla I.

VI. IMPLEMENTACIÓN PROPUESTA - RAINBOW CATALOG

Como primera aproximación técnica a los requisitos identificados, se ha generado un servicio propio de catálogo de metadatos integrado en la implementación del conector de espacios de datos. Este desarrollo, denominado Rainbow, se ha centrado especialmente en garantizar

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
TABLA DE REQUISITOS.

R1	Capacidad de generar metadatos desde sistemas federados con esquemas unificados. [24]
R2	Capacidad de resolver información sobre los metadatos en los flujos de negociación de transferencia y de políticas. [24]
R3	Capacidad del sistema de metadatos de poder almacenar y gestionar políticas de acceso y uso, así como la capacidad de indexarlas.
R4	Definición de metadatos de manera interoperable con estándares internacionales reconocidos. [25]
R5	Definición de políticas con estándares internacionales reconocidos. [26]
R6	Sistemas automatizados para ingestión de metadatos desde las fuentes de datos o desde sistemas federados. Facilidad de integración con sistemas propios de ingeniería de datos.
R7	Capacidad de generar información sobre el linaje de los datos y la calidad del dato. [27]
R8	Capacidad de gestión de búsqueda y descubrimiento de *datasets*.

la interoperabilidad, adoptando como base el vocabulario estandarizado DCAT3 para la descripción de metadatos, y permitiendo la asociación de políticas de acceso mediante el modelo ODRL 2.2, conforme a sus especificaciones. Cabe señalar que este enfoque es coherente con las directrices del propio *Dataspace Protocol*, que adopta DCAT3 en su definición del *Catalog Protocol*. El desarrollo de Rainbow es público y el repositorio se encuentra en el siguiente enlace: https://github.com/EunomiaUPM/rainbow

Una de las principales ventajas de esta solución es el cumplimiento de los requisitos de interoperabilidad, gracias a la adopción de estándares. Asimismo, al tratarse de una implementación desarrollada como parte integral del conector, se consigue una integración nativa con los flujos de negociación de contrato y de transferencia de datos.

Otro aspecto destacable es la capacidad de definir y vincular políticas ODRL directamente a los metadatos del catálogo, lo que permite una gestión coherente del control de acceso y uso desde la propia estructura del sistema.

No obstante, esta implementación presenta algunas limitaciones. En particular, la funcionalidad de descubrimiento de datos no está suficientemente cubierta, ya que actualmente no incorpora mecanismos avanzados de búsqueda, filtrado o enriquecimiento semántico de metadatos. Tampoco contempla la automatización del proceso de ingestión mediante sistemas de extracción de metadatos, lo que implica una alta carga operativa para los administradores encargados de mantener actualizado el catálogo.

Por último, en contextos donde se requiere gobernanza distribuida o se manejan grandes volúmenes de datos —como en organizaciones complejas o administraciones públicas—, es imprescindible contar con catálogos federados. En su estado actual, Rainbow no da soporte a este tipo de despliegue, lo que limita su aplicabilidad en escenarios más exigentes.

VII. USO DE DATAHUB PARA INTEGRACIÓN DE FUENTES DE DATOS

Dada la importancia de sistemas federados y la limitación de Rainbow exploramos Datahub. DataHub es una plataforma moderna de metadatos.

Una de las principales ventajas de la plataforma es la alta capacidad que tiene para integrarse con otras fuentes, desde las cuales se ingestan los datos. Estas fuentes pueden ser plataformas para la gestión de bases de datos relacionales o fuentes especializadas en procesos de extracción, transformación y carga. También admite la integración con sistemas de procesamiento de datos en tiempo real. Otra ventaja relevante que es que actualmente tiene una amplia adopción en sistemas en producción.

Internamente, DataHub organiza los metadatos como un grafo de entidades interconectadas. Esta estructura no solo potencia su motor de búsqueda y descubrimiento de recursos, sino que también permite visualizar el linaje de los datos de forma robusta, una capacidad clave para la gobernanza. También permite la inclusión de metadatos, sistemas de búsquedas basadas en términos clave, gestión de dominios, ect.

DataHub, no obstante, presenta ciertas limitaciones para su integración en espacios de datos, especialmente en términos de interoperabilidad y gestión de políticas.

Una primera limitación se enfoca en que las propiedades de los *datasets* se estructuran de forma no estandarizada, siendo meros atributos organizados como pares de clave-valor. Esto implica que para que se adhiera a estándares como DCAT sea preciso un control en en los sistemas ETL, lo que implica un paso innecesario de complejidad.

La segunda, y más importante para este trabajo, es que DataHub carece de soporte nativo para la gestión de políticas de acceso y uso. La plataforma no permite definir ni asociar políticas expresadas en lenguajes estándar como ODRL, un requisito esencial para la soberanía y la confianza en los espacios de datos. Estas carencias motivan la necesidad de una solución híbrida.

VIII. COMPARATIVA ENTRE DATAHUB Y RAINBOW CATALOG

Como se puede ver en la Tabla II, Rainbow Catalog está fuertemente dirigido a adherirse a estándares específicos como DCAT3 y ODRL 2.2, enfocándose en entornos que requieran interoperabilidad basada en estos estándares como es el caso de los espacios de datos. Por otro lado, DataHub es una solución más generalista y madura, destacando por su interfaz gráfica avanzada, sus amplias capacidades de integración para la ingesta automática de metadatos desde diversas fuentes, y su flexibilidad para la gobernanza de datos mediante un modelo propio y funcionalidades como el etiquetado y el linaje. Es una opción muy popular en la industria para construir un catálogo de datos federado y moderno.

A la luz de esta comparativa, concluimos que una buena opción puede ser una integración por capas, donde se haga una gestión de metadatos en DataHub aprovechando la capacidad de gestión de los mismos y las integraciones

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla II
TABLA DE COMPARATIVA DE SOLUCIONES.

Característica	**Rainbow Catalog**	**DataHub**
Función principal	Sistema de catálogo de metadatos.	Plataforma avanzada de gestión y descubrimiento de metadatos, enfocada en gobernanza y catálogo de datos moderno.
Modelo de Metadatos	Capacidad de resolver información sobre los metadatos en los flujos de negociación de transferencia y de políticas.	Modelo interno propio: DataHub Entities Model.
API	Capacidad del sistema de metadatos de poder almacenar y gestionar políticas de acceso y uso, así como la capacidad de indexarlas.	API estable (usualmente GraphQL), bien documentada y ampliamente utilizada.
Interfaz Gráfica (GUI)	Actualmente en desarrollo.	GUI muy completa y probada, uno de sus puntos fuertes para la exploración y gestión de metadatos.
Integraciones	Definición de políticas con estándares internacionales reconocidos.	Integración nativa con una vasta gama de sistemas (Bases de datos, Data Lakes, Warehouses, BI, ETL) para extracción automática de metadatos (ingesta).
Políticas	Soporte nativo para políticas ODRL 2.2 (a nivel de *Catalog*, *dataset*, Distribution, DataService). Roadmap para gestión de herencia.	Foco en gobernanza mediante etiquetado (tagging) flexible, dominios, glosarios y propietarios. No soporta ODRL nativamente.
Versionado/Linaje	Roadmap para versionado (DCAT3) y linaje (Relations).	Capacidades robustas de linaje de datos (visualizable en la GUI) y versionado implícito de metadatos.
Notificaciones	Sistema Pub/Sub para notificar cambios de estado (altas, modificaciones, bajas).	Emite eventos de cambio de metadatos, usualmente vía Kafka, para integración con otros sistemas.
Estándares/Compliance	Enfocado en DCAT3. Roadmap para DPROD, VCs (Verifiable Credentials) y Gaia-X.	Ampliamente adoptado en la industria, pero no específicamente diseñado alrededor de estándares como DCAT3 o Gaia-X de forma nativa.
Desarrollo/Comunidad	Parte del proyecto Rainbow. Roadmap activo	Proyecto Open Source muy activo y popular, con una gran comunidad y desarrollo constante. Ampliamente utilizado en la industria

que tiene, y por otro lado trabajar en la gestión de políticas ODRL en Rainbow Catalog.

IX. INTEGRACIÓN DEL CATÁLOGO Y EL CONECTOR DE ESPACIOS DE DATOS

En este apartado se describe la arquitectura y el flujo de integración del catálogo federado de DataHub con el conector de espacios de datos Rainbow.

Para aportar un caso de uso significativo y validar la implementación, se ha construido un escenario práctico que federa dos catálogos a partir de fuentes de datos abiertas utilizando DataHub, Apache Airflow y Apache NiFi. Adicionalmente, se ha habilitado la gestión de políticas ODRL sobre los *datasets* de ese catálogo desde Rainbow, de modo que se alcance gobernanza descentralizada, requisito esencial en un espacio de datos.

A. Flujo de integración

El flujo extremo a extremo que implementamos se organiza en cinco etapas:

1. Ingesta de metadatos desde catálogos individuales: La extracción y normalización de metadatos de cada catálogo de origen se realiza con Apache NiFi. NiFi permite definir grafos de procesamiento que cubren tanto la extracción desde las fuentes como las transformaciones necesarias. En este paso es crítico conformar los metadatos a DCAT. Una vez procesados, se cargan en DataHub.

2. Ingesta y federación: Para consolidar varios catálogos, aprovechamos las capacidades nativas de ingesta de DataHub, incluida la integración entre diferentes catálogos. La orquestación de procesos, frecuencias y monitorización se resuelve con Apache Airflow. El resultado es un catálogo federado que agrega y expone de forma unificada los metadatos provenientes de múltiples dominios.

3. Exposición de *datasets* y catálogos a los participantes del espacio de datos: DataHub ofrece un *endpoint* GraphQL para aplicaciones de terceros sobre su modelo interno de entidades. Sin embargo, en entornos de espacios de datos se exige una interfaz neutra e interoperable para consulta de metadatos. El *Dataspace Protocol*, a través de su *Catalog Protocol* y el uso de DCAT, define precisamente ese punto de acceso estandarizado expuesto por el *Dataspace Provider*.

Por tanto, incorporamos una instancia puente que actúa de *proxy* entre el catálogo interno de DataHub y el *endpoint* DCAT requerido en el proveedor del espacio de datos. En la práctica, Rainbow dispone de un servicio de gestión de metadatos y una fachada que consulta DataHub y publica el punto de entrada DCAT conforme al *Dataspace Protocol*, de modo que consumidores y entidades federadoras externas no necesitan conocer detalles internos de DataHub.

4. Gestión de políticas ODRL aplicadas a *datasets* del catálogo federado: Dado que Rainbow está diseñado de manera nativa para manejar políticas ODRL, lo usamos como sistema de gestión. El ciclo de políticas incluiría:

- Descubrimiento: desde Rainbow se listan y seleccionan los recursos del catálogo federado (vía fachada a DataHub).
- Asociación: se crean políticas ODRL y se asocian al *target* que identifica de forma única al *dataset*

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

federado.
- Persistencia: las políticas se guardan en Rainbow, incluyendo metadatos de auditoría y versionado lógico cuando aplica.
- Resolución: Rainbow expone resolución, búsqueda y aplicación de políticas.

5. Uso de políticas en negociación de contrato y transferencia: Con el catálogo DCAT y las políticas ODRL disponibles, los participantes ya pueden iniciar las negociaciones de contrato, y ya se puede definir un sistema de control de políticas de acceso y uso asociado.

B. Arquitectura de integración

La arquitectura adopta una topología en capas con acoplamiento débil, definidas en Figuras 4 y 5:

- Capa de metadatos (DataHub) Persistenacia de metadatos: ingesta, catalogación, linaje y descubrimiento.
- Capa de políticas y contratos (Rainbow) Implementa la gestión de políticas ODRL, la exposición DCAT y los flujos de negociación del *Dataspace Protocol*. Esta capa es la que habilita la soberanía y el control de uso.
- Capa de integración (Fachada en Rainbow) Componente clave que desacopla DataHub y Rainbow. Dicha capa actúa como: a) Cliente de la API de DataHub, y b) Mapeo semántico entre el modelo interno de DataHub y representaciones DCAT para publicación en el *endpoint* del *Dataspace Provider*.

Esta separación permite evolucionar cada pieza de forma independiente: DataHub puede actualizar la información y metadatos; Rainbow puede evolucionar su modelo de políticas, herencias y mecanismos de evaluación.

De aquí se derivan dos flujos de trabajo: Un primer flujo de consulta de metadatos, y un segundo de negociación, persistencia y ejecución de políticas.

C. Flujo de exploración y selección de target

Este flujo arranca en Rainbow mediante una fachada que actúa como pasarela contra DataHub, como se aprecia en la Figura 4. Primero, la fachada se autentica y obtiene un token. Con ese contexto, emite consultas al servidor GraphQL habilitado en DataHub para listar *domains* que se proyectan como catálogos a nivel de DCAT y, después, recuperar el detalle de un *domain* junto a los *datasets* que contiene.

Para cada *dataset* seleccionado se solicita su detalle operativo, extrayendo la identidad única (el URN canónico de DataHub) y la información de distribución necesaria para los flujos del *Dataspace Protocol*: tipo de distribución, puntos de acceso (*endpoints*) y posibles configuraciones del servicio de acceso/consumo (p. ej., autenticación), así como más metadatos definidos. El URN canónico de DataHub se emplea como atributo *target* de la política ODRL.

Con esta información, Rainbow mapea el modelo interno de DataHub a DCAT y publica una vista DCAT neutra (*endpoint* del *Dataspace Provider*), de modo que los participantes del espacio de datos pueden consultar catálogos y *datasets* sin conocer el modelo de entidades de DataHub.

D. Flujo de creación y aplicación de políticas

Con el *dataset* identificado en el flujo anterior, Rainbow Catalog gestiona políticas ODRL de forma nativa como se ilustra en la Figura 5. El ciclo es:

1) Consulta/validación del *dataset*: antes de asociar una política, la Rainbow ODRL API pide a la fachada *dataset* para verificar que el *dataset* existe y que los metadatos operativos (distribución, *endpoint*) son consistentes.
2) Creación y persistencia de la política: se conforma la política ODRL del *dataset* en DataHub y se persiste en Rainbow con su propio identificador.
3) Negociación de contrato (CN): cuando el *Dataspace Consumer* inicia la negociación, Rainbow realiza una doble comprobación: vuelve a consultar el *dataset* vía fachada y recupera la política. De ser aceptada por ambas partes se genera un *Agreement* con el contenido ODRL de dicha política.
4) Negociación de transferencia (TP): usando el *target* de la política almacenado en el *Agreement*, Rainbow resuelve de nuevo, a través de la fachada, el tipo de distribución, el punto de acceso y las configuraciones pertinentes del sistema final. Con ello autoriza y entrega al consumidor el acceso conforme a las condiciones pactadas.

Con este esquema en dos fases, DataHub aporta gobernanza, linaje, ingesta y federación a gran escala, mientras que Rainbow aporta exposición DCAT interoperable y gestión de políticas ODRL estrechamente acoplada a los flujos de negociación de contrato y negociación de transferencia del *Dataspace Protocol*. El resultado es una integración coherente: descubrimos y resolvemos metadatos operativos (fase 1) y negociamos/transferimos con control de uso (fase 2), siempre anclados al URN del *dataset* federado.

E. Aplicación al caso de uso

Sobre el caso de uso, se ha aplicado todo el sistema de ETL desde los catálogos de datos para extraer los metadatos con Apache NiFi, la orquestación de con Apache Airflow de manera exitosa. Se puede, por tanto, generar una instancia federadora y dos instancias federadas con DataHub. El repositorio se puede encontrar en este enlace: https://github.com/pvinualese/datahub_dockerized.

Por otro lado, se lanzó una instancia de Rainbow, donde se generó un módulo ad-hoc para operar como *proxy* contra DataHub. Dicho flujo se encuentra en el notebook bajo este enlace: https://github.com/ging/ds-deployment/blob/main/src/pull_flow_rpc_datahub.ipynb. Como se puede observar en el flujo, se usa el *proxy* para acceder a los datos de los *datasets* de DataHub, y posteriormente se pueden persistir y usar políticas ODRL asociadas al *dataset* y persistirlas en Rainbow.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

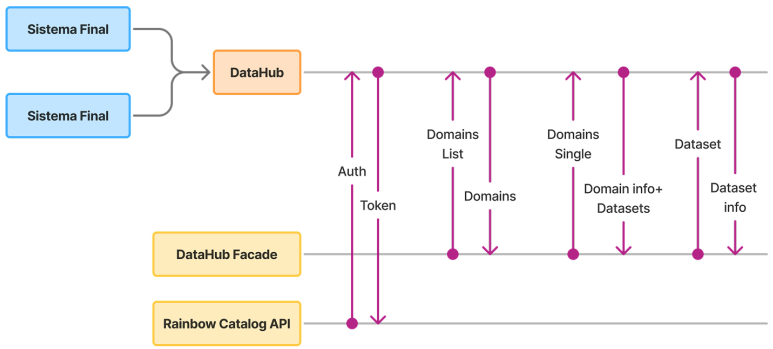


Fig. 4. Flujo de consulta de metadatos.

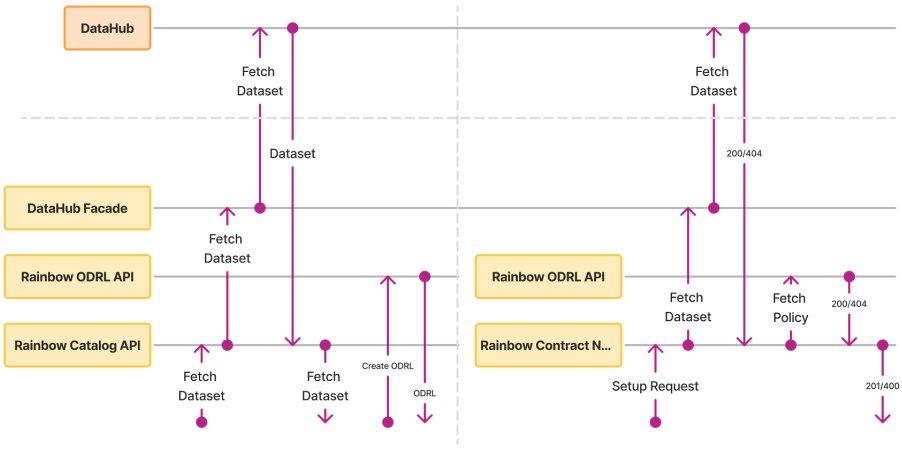


Fig. 5. Flujo de creación y aplicación de políticas

F. Desafíos y trabajos a futuro

Durante la implementación, han aparecido retos técnicos que se tratará de abordar en futuros trabajos, tales como:

- Coherencia y sincronización de metadatos: En caso de haber cambios en DataHub como modificaciones o bajas en los mismos, las políticas quedarían invalidadas al no poder resolver *targets*. Una opción posible sería valorar sistemas de suscripción a cambios y definir una interfaz para ello.
- Rendimiento en consultas federadas: En caso de haber un número alto de *datasets* puede que el sistema de consultas basado en GraphQL actúe como cuello de botella. Una posible solución sea la de consultas estrechas a DataHub (sólo con datos necesarios), paginaciones o preindexaciones y sistemas de caché en la fachada de Rainbow.
- Mapeo semántico de DataHub a DCAT. DataHub tiene un sistema de entidades propio y el mapeo a DCAT no es trivial. Tanto como para la gestión de campos complejos como ODRL, como para campos relacionales como sistemas de versiones, ect.
- Gestión de herencia y ámbito de políticas. Actualmente las políticas se asocian con *datasets*, pero se debería investigar en asociar las políticas a nivel de Catálogo, u otras entidades definidas en DCAT. Para ello haría falta una redefinición del uso de políticas a nivel de *Dataspace Protocol* que pretendemos abordar en trabajos futuros.

X. CONCLUSIONES

Este documento se ha centrado en hacer una propuesta de implementación sobre una cuestión crucial dentro de los ecosistemas de espacios de datos y su posible adopción masiva por parte de actores proveedores de servicios de datos que es la integración de catálogos federados que puedan funcionar a gran escala y que estén basados en soluciones existentes y probadas en la industria, y al mismo tiempo que puedan abordar los requisitos de interoperabilidad inherentes a los espacios de datos.

En dicho documento se ha hecho una revisión del papel que juega el catálogo centrándonos en los flujos de negociación y transferencia. Por otro lado se ha hecho una comparativa sobre una implementación existente y desarrollada por el grupo de investigación que propone este artículo, que parte de los principios de interoperabilidad y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

uso de estándares, muy cercana al caso de uso relacionado con los flujos de un espacio de datos; y por otro lado de una solución como DataHub con gran adopción a nivel global desde el punto de vista de gestión y gobernanza del dato.

En la investigación hemos detectado la necesidad de hacer una implementación híbrida, definiendo las capas de ingesta, gestión y persistencia de metadatos bajo DataHub, y definiendo las capas de creación, adjudicación, gestión y persistencia de las políticas ODRL en Rainbow. Entendemos que es una solución que desde el punto de vista de arquitectura es coherente, y además en la implementación práctica está resultando exitosa.

Esta integración es total y abre la puerta al desarrollo de una interfaz de usuario, desde la cual se podrán hacer ofertas para acceder a un recurso, negociar las condiciones de acceso. De esta manera, consumidores y proveedores podrán llegar a acuerdos de cara permitir la transferencia de los datos. Al mismo tiempo se recalca que implementaciones de catálogos federados permiten escalar el impacto de los espacios de datos a casos de uso masivos en gestión de metadatos y de gestión de la gobernanza desde un punto de vista integral.

AGRADECIMIENTOS

Los autores agradecen el apoyo del proyecto EUNOMIA, "Soluciones para la soberanía, confianza y Seguridad en los espacios de datos" (ETD202300196), financiado por el INCIBE (Instituto de Ciberseguridad).https://eunomia.dit.upm.es, y de la Agencia Estatal de Investigación (AEI), 10.13039/501100011033., a través del proyecto FuN4Date con la subvención PID2022-136684OB-C22.

REFERENCIAS

[1] Attard, Judie and Orlandi, Fabrizio and Scerri, Simon and Auer, "A systematic review of open government data initiatives", Government information quarterly, vol. 32, no 4, pp. 399-418, 2015.

[2] Schmidt, Sophie C and Thiery, Florian and Trognitz, Martina,"Practices of linked open data in archaeology and their realisation in Wikidata", Digital, vol. 2, no 3, pp. 333-364, 2022.

[3] Wahyo Setiawan Wibowo, Dana Indra Sensuse, Sofian Lusa, Prasetyo Adi, Wibowo Putro, Wibowo and Alivia Yulfitri, "A systematic literature review on open government data: Challenges and mapped solutions" Journal of Theoretical and Applied Information Technology, vol. 101, no. 5, pp. 1806–1818, 2023.

[4]] J. Umbrich, S. Neumaier, and A. Polleres, "Quality assessment andevolution of open data portals," 3rd international conference on future internet of things and cloud. IEEE, pp. 404–411, 2015

[5] M. Beno, K. Figl, J. Umbrich, and A. Polleres, "Open Data Hopes and Fears: Determining the Barriers of Open Data," 2017 Conference for E-Democracy and Open Government (CeDEM), pp. 69–81, 2017.

[6] Simon Scerri, Tuomo Simon, Tuomo Tuikka, Irene Lopez de Vallejo, and Edward Curry, "Common european data spaces: Challenges and opportunities", Data Spaces: Design, Deployment and Future Directions, 2022, p. 337-357.

[7] Manlio Bacco, Alexander Kocian, Stefano Chessa, Antonino Crivello, Antonino and Paolo Barsocchi, "What are data spaces? Systematic survey and future outlook", Data in Brief, vol. 57, p. 110969, 2024.

[8] Angelo Martella, Angelo, Cristian Martella, and Antonella Longo, "Designing Data Spaces: Navigating the European Initiatives Along Technical Specifications", arXiv preprint arXiv:2503.15993, 2025.

[9] World Wide Web Consortium. "ODRL Information Model 2.2," W3C Recommendation, Feb. 15, 2018. [Online]. Available: https://www.w3.org/TR/odrl-model.

[10] Rodrigo Menéndez, Andres Munoz-Arcentales, Joaquín Salvachúa, Carlos Aparicio, Irene Plaza, Gabriel Huecas, "Next Generation Authentication for Data Spaces: An Authentication Flow Based On Grant Negotiation And Authorization Protocol For Verifiable Presentations (GNAP4VP)", arXiv preprint arXiv:2505.24698, 2025.

[11] International Data Spaces Association, "Dataspace Protocol 2024-1 – Release Candidate," Eclipse Dataspace Working Group, Jul. 2024. [Online]. Available: https://internationaldataspaces.org/publications.

[12] Tobias Dam, Andreas Krimbacher and Sebastian Neumaier, "Policy Patterns for Usage Control in Data Spaces", arXiv preprint arXiv:2309.11289, 2023.

[13] World Wide Web Consortium. "Data Catalog Vocabulary (DCAT) - Version 3," W3C Recommendation, Aug. 22, 2024. [Online]. Available: https://www.w3.org/TR/vocab-dcat-3

[14] Marco Hauff, Lina Molinas Comet, Paul Moosmann, Christoph Lange, Ioannis Chrysakis, Ioannis and Johannes Theissen-Lipp, "FAIRness in Dataspaces: The Role of Semantics for Data Management", Proc. of SDS, 2024.

[15] Tobias Dam, Lukas Daniel Klausner, Sebastian Neumaier and Torsten Prieb, "A Survey of Dataspace Connector Implementations", arXiv preprint arXiv:2309.11282, 2023.

[16] Adriana Morejón, Alberto Berenguer, Lucía Espona, David Tomás, Jose-Norberto Mazón, "Exploring Content-Based Catalogs for Enhanced Discovery Services in Data Spaces", 2025.

[17] Theissen-Lipp, Johannes and Hemid, Ahmad and Lange, Christoph and Gillmann, Christina "owards a Federated Data and Service Catalogue", 2024.

[18] Jahnke, Nils and Fraunhofer, ISST and Möller, Frederik, "Federated Data Catalogs for Data Sharing - Towards Design Principles, 2024.

[19] OpenID Foundation. "OpenID for Verifiable Credential Issuance," OpenID Connect Working Group Draft, Dec. 2024. [Online]. Available: https://openid.github.io/OpenID4VCI/openid-4-verifiable-credential-issuance-wg-draft.html

[20] OpenID Foundation. "OpenID for Verifiable Presentations," OpenID Digital Credentials Protocols Workgroup Draft, Jun. 10, 2025. [Online]. Available: https://openid.github.io/OpenID4VP/openid-4-verifiable-presentations-wg-draft.html

[21] Eclipse Dataspace Working Group. "Eclipse Decentralized Claims Protocol," v1.0-RC3, Mar. 31, 2025. [Online]. Available: https://eclipse-dataspace-dcp.github.io/decentralized-claims-protocol/v1.0-RC3/

[22] Irene Plaza Ortiz, Andres Munoz-Arcentales, Joaquín Salvachúa, Carlos Aparicio, Gabriel Huecas and Enrique Barra, "Authentication and authorization in Data Spaces: A relationship-based access control approach for policy specification based on ODRL", arXiv preprint arXiv:2505.24742, 2025.

[23] Jahnke, Nils, and Boris Otto, "Data catalogs in the enterprise: applications and integration.", Datenbank-Spektrum, vol. 23, no 2, p. 89-96, 2023.

[24] Li, Yanhuang. "Interoperability and Negotiation of Security Policies." Diss. Ecole Nationale Supérieure des Télécommunications de Bretagne-ENSTB, 2016.

[25] Brechtel, M., D. Petrik, K. Hölzle, "From Challenges to Solution Pathways for IndustrialData Ecosystems - A Socio-Technical Perspective" Wirtschaftsinformatik 2023 Proceedings, 2023.

[26] Noardo, Francesca, et al. "Standards for Data Space Building Blocks." Remote Sensing 16.20 (2024): 3824.

[27] Altendeitering, M., S. Dübler, T. Guggenberger. "Data Quality in Data Ecosystems: Towardsa Design Theory" Twenty-eighth Americas Conference on Information Systems, 2022.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Towards a quantum-safe 5G: Quantum Key Distribution in core networks

Asier Atutxa, Ane Sanz, Jorge Sasiain, Jasone Astorga, Eduardo Jacob
Department of Communications Engineering, Faculty of Engineering of Bilbao
University of the Basque Country (UPV/EHU)
Plaza Ingeniero Torres Quevedo n. 1, Bilbao, 48013, Spain
asier.atutxa@ehu.eus, ane.sanz@ehu.eus, jorge.sasiain@ehu.eus, jasone.astorga@ehu.eus, eduardo.jacob@ehu.eus

This paper presents a TLS authentication and encryption proposal in 5G multi-site environments leveraging entanglement-based Quantum Key Distribution (QKD), securing communications between NFs in the 5G core network with QKD keys. Additionally, this paper presents a QKD key repository using the Network Repository Function (NRF), which stores and relays TLS session keys to authenticated consumer NFs to access producer NFs in a fast but quantum-safe process.

***Palabras Clave*—5G, QKD, Service-Based Architecture (SBA), Network security**

I. PROPOSED SECURITY SYSTEM

This paper presents a novel system [1] depicted in Fig.1 to implement quantum-based security among NFs in 5G multi-site core networks, encompassing direct QKD links between the 5G core and several MEC sites. In this context, this proposal includes two different handshake approaches, depending on the process in which the NRF or the NFs take part: the **TLS handshake with QKD** and the **TLS handshake with Trusted-Relay QKD (TR-QKD)**.

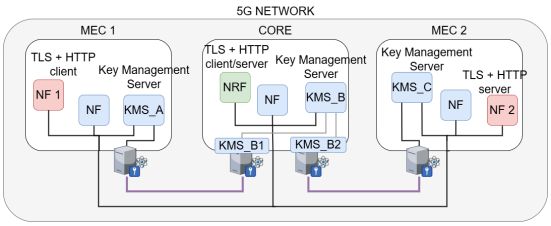


Fig. 1: Proposal for QKD implementation in 5G networks.

On the one hand, a **TLS handshake with QKD**, must be performed between each NF and the NRF before any Service Registration or Service Discovery. This handshake establishes a QKD key between them that can be used directly to secure their communications and make them quantum-safe. Additionally, after the Service Discovery request by the consumer NF, the NRF securely relays the QKD key acquired in the previous handshake with the producer NF, enabling both consumer and producer NFs to share the same QKD key, regardless of whether they have a direct QKD connection. On the other hand, before any Service Request process, assuming that consumer and producer NFs do lack a direct QKD link, the **TR-QKD** takes part. In this case, the consumer NF makes use of the QKD key previously relayed by the NRF, ensuring quantum-resistant communication between NFs.

II. RESULTS AND DISCUSSION

The results show that the fastest type of TLS handshake is the out-of-bound PSK system, based on previously distributed keys. It requires on average 21.08 milliseconds, 7.36 milliseconds, and 9.52 milliseconds in the three phases (registration, discovery, and request). However, the system proposed in this paper also achieves equally fast results after the Service Registration handshake, once the key has been obtained from the QKD equipment. The first handshake requires on average 95.62 milliseconds due to the API calls to the KMS, but the time measured in the next two handshakes are 7.92 and 9.62 respectively.

The out-of-bound system requires a method to distribute the keys, a challenge that our proposed QKD system already addresses. Therefore, a better comparison is done against the SECP521 and certificate-based TLS version of the test, which uses DH to obtain an encryption key and digital certificates for authentication. The obtained results are marginally faster in the first handshake (67.96 milliseconds), but significantly slower in the following two handshakes (46.9 and 51.2 respectively), which make the proposed QKD-based system faster on average.

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Science and Innovation in the project 6G-INSPIRE (PID2022-137329OB-C44), the SareQuant project from IKUR (IKUR-DIPCPRTR- 23/02) and the QFirst project (KK-2022/00062).

REFERENCIAS

[1] Asier Atutxa, Ane Sanz, Jorge Sasiain, Jasone Astorga, Eduardo Jacob, Towards a quantum-safe 5G: Quantum Key Distribution in core networks, Computer Communications, Volume 224, 2024, Pages 145-158, ISSN 0140-3664, https://doi.org/10.1016/j.comcom.2024.06.005.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Interpretación de vídeos mediante Inteligencia Artificial. Estudio preliminar para la detección de errores en cuidados médicos o de enfermería

Eva Gil[1], César Fernández[2], María Asunción Vicente[2], Irene Carrillo[2], José Joaquín Mira[1]

[1]Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO)
Hospital General Universitario, Ed. Anexo II, 3a planta, Camí de L´Almassera 11, Elche (03203), Alicante.

[2]Departamento de Psicología de la Salud. Universidad Miguel Hernández (UMH).
Avenida de la Universidad s/n Elche (03202) Alicante.

eva.gilh@umh.es, c.fernandez@umh.es, suni@umh.es, icarrillo@umh.es, jose.mira@umh.es

Los recientes modelos de Inteligencia Artificial (IA) son capaces de interpretar escenas complejas en imágenes y vídeos. Hemos probado su fiabilidad para detectar errores cometidos por cuidadores a cargo de personas dependientes. Partimos de tres videos que muestran tareas de cuidado rutinarias en las que, deliberadamente, se introducen diversos errores y evaluamos si tres diferentes modelos de IA (OpenAI, xAI y Gemini), son capaces de detectar esos errores. El análisis comparativo de los tres modelos ofreció unas tasas de acierto del 92.86% (xAI), 78.58% (Gemini) y 42.86% (OpenAI). La medida de repetibilidad en 10 ejecuciones realizada con el modelo Gemini ofreció un promedio del 90% de respuestas coincidentes. Sin embargo, las fiabilidades que las IAs estiman para sus propios resultados son poco realistas. Dado el escaso número de pruebas realizadas, se trata de resultados previos que requerirán comprobaciones más exhaustivas.

***Palabras Clave*- inteligencia artificial, visión por computador, salud, cuidadores.**

I. Introducción

La evolución del software de visión artificial ha ido acompañada del aumento del nivel en el que se programan los algoritmos; desde la programación basada en código de bajo nivel, con ayuda de librerías como OpenCV [1]; avanzando hacia entornos de programación por bloques, que limitaron la necesidad de escribir código [2]. Los modelos de IA, basados en las redes neuronales convolucionales (CNNs [3]), la arquitectura *transformer* y el concepto de atención [4] permiten un salto nunca imaginado: sólo indicamos a la IA el resultado que queremos obtener, sin especificar qué preprocesamiento, qué tipo de detección de contornos, qué métodos de segmentación, qué procedimiento de detección de movimiento, etc. Es el nivel de programación más alto posible: la IA diseña todo el algoritmo de visión por computador.

En paralelo, se ha producido un avance continuo en cuanto a la capacidad de reconocimiento de escenas complejas. En los comienzos de la visión por computador, se requerían fondos uniformes para poder detectar objetos (por ejemplo, piezas en una cinta transportadora o vehículos en una carretera vistos desde una imagen cenital) [5]. Los procesamientos se limitaban a extraer los contornos (mediante segmentación o detección de bordes) [6,7] e intentar reconocer el objeto por similitud a los almacenados en una base de datos, muy limitada en cuanto a número de objetos diferentes. Más adelante, se pudo detectar objetos sobre fondos complejos, a través de descriptores invariantes como SIFT o SURF [8] y aparecieron técnicas específicas para la detección de caras humanas [9]. Posteriormente, se consiguieron identificaciones de personas a partir de su rostro [10]. De nuevo, la IA y el entrenamiento de los algoritmos de visión por computador con grandes bases de datos de imágenes y vídeos suponen un salto respecto al ritmo de avance previo: actualmente es posible reconocer cualquier escena compleja, identificando los objetos que contiene y la relación entre ellos.

En este contexto, nos planteamos si las herramientas de IA serían aplicables para tareas de supervisión, en concreto, para la detección de los errores que los cuidadores (principalmente los no profesionales) pueden cometer al realizar los cuidados rutinarios a sus familiares o personas de las que cuidan. Es una aplicación relevante en la sociedad actual, con un porcentaje de personas dependientes muy elevado que, normalmente, reciben cuidados de familiares sin formación médica. La relevancia del problema queda patente al comprobar las diferentes perspectivas desde las que se ha abordado [11,12].

El objetivo del estudio es evaluar la capacidad de los modelos de IA más recientes para interpretar vídeos grabados mientras se realizan cuidados a personas dependientes y detectar en ellos errores comunes, si se producen.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

II. Metodología

A. Videos a analizar.

Se partió de una base de datos de vídeos de cuidadores realizando tareas cotidianas de cuidado de personas dependientes generada en el marco del proyecto Reality Care (https://realitycare.es). En tales vídeos, los cuidadores cometen a propósito errores comunes (por ejemplo, no lavarse las manos al iniciar el procedimiento, no desinfectar la zona antes de administrar un medicamento con jeringuilla, etc.). La lista completa de vídeos es accesible en esta lista de reproducción de YouTube: http://bit.ly/3ZrKMsa. Se seleccionaron 3 vídeos y para cada uno de ellos se generó un listado de posibles errores de cuidados, algunos cometidos realmente en el vídeo y otros no, con el objeto de estudiar si la IA es capaz de discernir cuáles de los errores se producen en realidad.

B. Modelos de IA utilizados.

Cada vídeo se analizó con tres modelos diferentes de IA: el modelo o3 de OpenAI (ChatGPT), el modelo DeeperResearch de xAI (Grok) y el modelo DeepSearch de Gemini (Google). En todos los casos, se trata de las opciones con mayor capacidad de razonamiento de cada proveedor hasta la fecha de realización de los experimentos. Otros proveedores se descartaron por no permitir procesar videos, ni como ficheros adjuntos ni mediante enlaces.

C. Prompt para la detección de errores.

En todos los casos se utilizó el mismo prompt, únicamente adaptado a cada vídeo (el listado de errores a detectar varía) y a la forma de proporcionar el vídeo al modelo (como enlace o como fichero adjunto):

{Adjunto un vídeo | Te indico la URL de un vídeo de Youtube} en el que se administra insulina a un paciente.

Quiero que analices tanto el audio (que está en castellano y puedes transcribir) como el vídeo y compruebes varios aspectos del procedimiento: 1) si la persona cuidadora se limpia las manos antes de comenzar; 2) si desinfecta la zona antes del pinchazo; 3) si ejerce presión hacia adelante en el dedo para que salga la sangre en la medición de glucosa; 4) si calienta la aguja de insulina; 5) si aprieta y profundiza en la zona del pinchazo para administrar la insulina; 6) si encapucha la aguja al terminar.

Responde exclusivamente con el resultado, en formato array json, con los siguientes campos: CAMPO 1, con nombre "pregunta" y que debe contener el nombre de cada una de las preguntas 1 a 6 que te realizo; CAMPO 2, con nombre "respuesta" y valores "sí" o "no"; CAMPO 3, con nombre "fiabilidad" y un valor en el rango 0-100 indicando la fiabilidad de la respuesta anterior; CAMPO 4, con nombre "explicación" y que contendrá una explicación textual tan larga como necesites indicando los procesamientos de imagen realizados y justificando el resultado. Tómate todo el tiempo que necesites para los procesamientos.

D. Comparación entre modelos.

El experimento de comparación entre modelos consistió en ejecutar el prompt para cada uno de los tres vídeos y para cada uno de los tres modelos. El principal resultado que se analizó fue el porcentaje de aciertos en las respuestas (si la IA acierta detectando qué errores se cometen en el vídeo y qué errores no se cometen). También se analizó la fiabilidad que la propia IA estima para sus respuestas. Los datos de esa fiabilidad estimada fueron comparados con la fiabilidad real (acierto o fallo), con el objeto de determinar si la fiabilidad estimada sería útil en un entorno real.

E. Estudio de repetibilidad.

Dado el carácter no determinista del comportamiento de las IAs, se realizó un estudio adicional de repetibilidad. Este estudio se llevó a cabo únicamente con el modelo Gemini por ser el único que a la fecha de realización de los experimentos permitía acceder a la interpretación de vídeos a través de su API. Para el experimento se lanzaron 10 repeticiones del prompt para cada uno de los 3 vídeos y se analizaron las variaciones obtenidas en los resultados.

F. Disponibilidad de resultados en abierto.

Se ha posibilitado el acceso a los resultados del experimento, así como al código fuente utilizado, a través de la página web https://lcsi.umh.es/realitycare.

III. Resultados

A. Videos y posibles errores seleccionados.

Los tres vídeos seleccionados para el análisis, junto con sus detalles (procedimiento, preguntas a realizar y respuesta correcta) se muestran en la Tabla I.

Tabla I
Vídeos para analizar mediante IA

Procedimiento	Posibles errores del cuidador	S/N
Administración de heparina	¿Se lava las manos antes de empezar?	N
	¿Desinfecta la zona antes del pinchazo?	N
	¿Quita el aire de la aguja?	S
	¿Aplica presión después del pinchazo?	S
	¿Encapucha la aguja antes de tirarla?	S
Medida de glucosa y administración de insulina	¿Se lava las manos antes de empezar?	N
	¿Desinfecta la zona antes del pinchazo?	N
	¿Ejerce presión hacia adelante en el dedo para que salga la sangre?	N
	¿Calienta la aguja de insulina?	N
	¿Aprieta y profundiza en la zona del pinchazo para administrar la insulina?	S
	¿Encapucha la aguja antes de tirarla?	S
Administración de morfina	¿Hace bien la disolución?	N
	¿Purga el catéter?	N
	¿Se lava las manos antes de empezar?	N

B. Ejemplo de respuestas obtenidas.

Como ejemplo de las respuestas obtenidas, se muestra el resultado devuelto por el modelo o3 de OpenAI para la primera pregunta del primero de los videos (administración de heparina):

{
"pregunta": "¿La persona se limpia las manos antes de comenzar?",

"respuesta": "sí",

"fiabilidad": 70,

"explicación": "Se extrajeron fotogramas cada 0,5 s tras descartar los primeros 5 s de introducción identificados

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

mediante baja varianza de histograma de brillo. En el intervalo 00:05–00:15 se detectó movimiento de manos bajo un grifo: flujo de píxeles brillantes con componente cromática azul-verdosa y patrón ondulante típico de agua corriente, junto a la detección de espuma (segmentación por valores altos en el canal V y baja saturación). La sucesión 'manos-agua-toalla' se clasificó como lavado de manos. El algoritmo de clasificación de acciones basado en diferencias de contorno y optical flow dio una confianza media del 0,70, por lo que se asigna fiabilidad 70."
}

La totalidad de las respuestas obtenidas se puede consultar en la web https://lcsi.umh.es/realitycare.

C. *Comparación entre modelos.*

En cuanto al nivel de aciertos en las respuestas de las IAs, la Tabla II muestra los resultados promediados entre todas las imágenes, según el modelo de IA empleado. Puede apreciarse cómo los mejores resultados fueron obtenidos por el modelo xAI (92.86% de aciertos en promedio para todas las preguntas de todos los vídeos); y los peores resultados fueron obtenidos por el modelo OpenAI (42.86% de respuestas acertadas). Curiosamente, este último modelo, a pesar de obtener peores resultados, es más optimista con la fiabilidad que indica, llegando a indicar fiabilidades del 100% para respuestas que realmente son erróneas.

A este respecto, se realizó un análisis para determinar cuál de los modelos de IA ofrecía fiabilidades más realistas. Los resultados se muestran en la Fig. 1 (OpenAI), en la Fig. 2 (xAI) y en la Fig. 3 (Gemini). En los tres casos, se representa, para cada valor de fiabilidad indicado por el modelo (eje x), el porcentaje de respuestas correctas acumulado (o porcentaje de aciertos entre todas las respuestas para las que el modelo de IA ha estimado una fiabilidad igual o inferior a ese límite). Idealmente, el porcentaje de aciertos debería ser mayor en los casos en los que las IAs indican que su respuesta tiene una alta fiabilidad, lo que supondría unos diagramas de barras de altura creciente. Este resultado sólo se aprecia en el modelo Gemini, con lo que en general no es posible confiar en el parámetro de fiabilidad estimada por la IA.

Tabla II
RESULTADOS OBTENIDOS POR CADA MODELO DE IA

Medida	OpenAI	xAI	Gemini
Porcentaje de acierto promedio	42.86	92.86	78.58
Fiabilidad indicada promedio	62.50	57.86	80.29

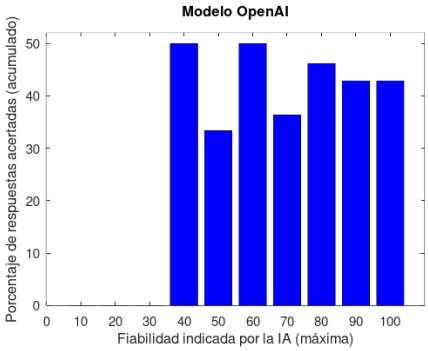


Fig. 1. Fiabilidad indicada vs precisión real para el modelo OpenAI.

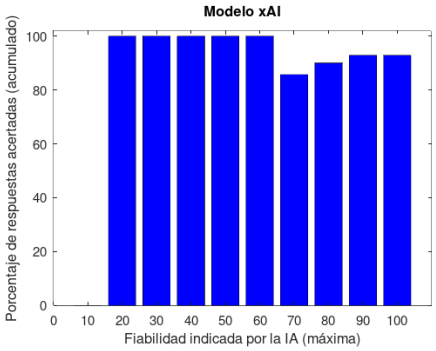


Fig. 2. Fiabilidad indicada vs precisión real para el modelo xAI.

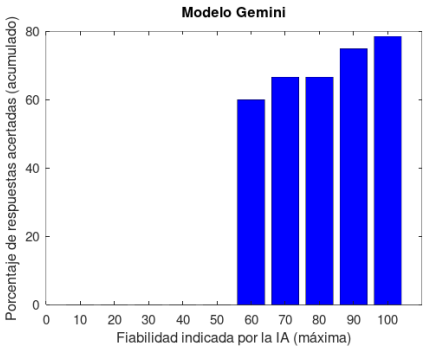


Fig. 3. Fiabilidad indicada vs precisión real para el modelo Gemini.

D. *Estudio de repetibilidad.*

Los resultados del estudio de repetibilidad se muestran en la Tabla III. Las columnas A hasta J representan cada una de las 10 repeticiones de los experimentos. Las filas P1 hasta P14 representan las preguntas realizadas en el conjunto de los tres vídeos (P1 a P5 para el vídeo de administración de heparina, P6 a P11 para el vídeo de administración de insulina y P12 a P14 para el vídeo de administración de morfina). Cada celda indica S/N como la respuesta de la IA (sí/no) a cada pregunta. La última fila muestra el porcentaje de respuestas atípicas frente a la respuesta mayoritaria (sea o no la respuesta correcta).

Tabla III
REPETIBILIDAD (REPETICIONES A HASTA J, PREGUNTAS 1 A 14)

	A	B	C	D	E	F	G	H	I	J	Medidas atípicas
P1	N	N	N	N	N	N	N	N	N	N	0%
P2	N	N	N	N	N	N	N	N	N	N	0%
P3	S	S	S	S	S	S	S	S	S	S	0%
P4	S	S	S	S	S	S	S	S	S	S	0%
P5	S	S	S	S	S	S	S	S	S	S	0%
P6	N	N	N	N	N	N	N	N	N	N	0%
P7	N	N	N	N	N	S	N	N	N	N	10%
P8	N	S	N	N	S	S	S	S	N	S	40%
P9	N	N	N	N	N	N	N	N	N	N	0%
P10	N	S	N	N	S	N	S	N	S	S	50%
P11	S	S	N	N	N	S	S	S	S	S	30%
P12	N	N	N	N	N	N	N	N	N	N	0%
P13	S	S	S	S	S	S	S	S	S	S	0%
P14	N	N	S	N	N	N	N	N	N	N	10%
Promedio											10%

En general, los resultados son completamente repetitivos (0% de medidas atípicas), pero en ciertas preguntas hay gran variabilidad, como la pregunta P10 (Vídeo insulina: ¿Aprieta

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

y profundiza en la zona del pinchazo para administrar la insulina?) o la pregunta P8 (Vídeo insulina: ¿Ejerce presión hacia adelante en el dedo para que salga la sangre en la medición de glucosa?). En ambos casos las preguntas se refieren a fragmentos de los vídeos cuya interpretación deja lugar a dudas incluso para un revisor humano, con lo cual es justificable que la variabilidad en las respuestas sea elevada en tales casos. El promedio, teniendo en cuenta todas las preguntas de todos los vídeos, es de un 10% de medidas atípicas.

IV. Discusión y limitaciones

Dado que el prompt dejaba abierto deliberadamente el tipo de procesamiento a realizar para obtener el resultado pedido, cada modelo optó por una estrategia diferente. En concreto, el modelo de OpenAI utilizó procesamiento de imágenes, reconocimiento de sonidos (por ejemplo, agua de un grifo) y, en algunos casos, transcripciones de las conversaciones. El modelo de Gemini se centró mayoritariamente en las imágenes; por último, el modelo de Grok se basó exclusivamente en las transcripciones de las conversaciones, pero, a pesar de ello, obtuvo los mejores resultados en cuanto a tasa de aciertos.

En https://lcsi.umh.es/realitycare se pueden descargar las respuestas completas de todos los modelos en el experimento comparativo. Aunque el detalle de los procedimientos utilizados es variable, es posible comprobar algunas estrategias utilizadas, como en el caso del modelo OpenAI a la pregunta de si la persona cuidadora se lava las manos. En este caso se detalla que en el procesamiento se han combinado tres fuentes de información diferentes: 1) las imágenes, en las que intenta detectar tanto el patrón de movimiento de las manos como el parpadeo brillante del agua; 2) el audio de fondo, en el que intenta detectar el sonido característico del flujo de agua en un grifo; y 3) la transcripción de la voz de la persona cuidadora, en la que busca alguna frase indicativa de lavado de manos. Adicionalmente, tiene en cuenta la concordancia temporal entre todas estas detecciones.

En general, los resultados de todos los modelos de IA podrían mejorar notablemente si el prompt detallara más el procedimiento a seguir para las detecciones, pero el objetivo del estudio era comprobar la capacidad de las diferentes IAs para realizar la tarea sin ninguna indicación adicional, salvo el resultado que se desea obtener. También se habrían obtenido mejores resultados con interacciones posteriores, como por ejemplo "realiza otra vez la transcripción del audio y conserva sólo los fragmentos reconocidos con mayor fiabilidad" o "repite los procesamientos de imágenes para comprobar los resultados anteriores", pero tampoco es el objetivo del estudio, que por ahora busca evaluar la capacidad de los modelos de IA para actuar en modo automático, no semiautomático.

Una limitación importante del estudio es el número de vídeos utilizado, y la falta de repeticiones en el experimento comparativo. Se trata de un estudio preliminar que requerirá un mayor número de pruebas para poder extraer conclusiones fiables, pero que muestra la potencialidad de los modelos actuales de IA para interpretar vídeos y detectar las acciones realizadas en ellos con gran detalle.

V. Conclusiones y trabajos futuros

Este estudio pone de manifiesto las capacidades y limitaciones de un sistema de supervisión automática de tareas de cuidado de personas dependientes basado en IA. Aunque los resultados obtenidos son preliminares, muestran la potencialidad de la solución propuesta. El trabajo futuro se centrará en realizar pruebas más exhaustivas y en la búsqueda de estrategias que permitan mejorar la capacidad de acierto y también la corrección de las medidas de fiabilidad devueltas por los modelos de IA. En paralelo, la aparición constante de modelos con mayores capacidades hará necesaria la realización de nuevas evaluaciones.

El estudio se ha planteado en base a un entorno en el que la IA analiza cada uno de los vídeos independientemente, y devuelve un resultado indicando la lista de errores detectados. Un planteamiento diferente consistiría en utilizar la IA para generar código fuente genérico con el que construir un sistema de interpretación de vídeos para su utilización posterior. Es una alternativa sobre la que aún no hemos trabajado, pero que también consideraremos en futuros estudios.

Agradecimientos

Esta investigación ha sido financiada por el Ministerio de Ciencia e Innovación (MCIN)/AEI (Agencia Estatal de Investigación) / DOI 10.13039/501100011033 (Digital Object Identifier) (TED 2021-130383B-I00), y por la Unión Europea "NextGenerationEU"/PRTR (Plan de Recuperación, Transformación y Resiliencia).

Referencias

[1] Bradski, G., Kaehler, A. "OpenCV". *Dr. Dobb's journal of software tools,* 3(2), 2000.

[2] Bendale, A., Chiu, K., Marwah, K., & Raskar, R. "Visionblocks: A social computer vision framework". *Proc. of 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pp. 521-526, 2011.

[3] Bharati, P., & Pramanik, A. "Deep learning techniques—R-CNN to mask R-CNN: a survey". *Computational Intelligence in Pattern Recognition: Proceedings of CIPR*, pp. 657-668, 2020.

[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. "Attention is all you need". *Advances in neural information processing systems*, 30, 2017.

[5] Forsyth, D. A., & Ponce, J. (2002). *Computer vision: a modern approach*. Prentice Hall professional technical reference.

[6] Yu, Y., Wang, C., Fu, Q., Kou, R., Huang, F., Yang, B., ... & Gao, M. "Techniques and challenges of image segmentation: A review". *Electronics*, 12(5), 2023.

[7] Gong, X. Y., Su, H., Xu, D., Zhang, Z. T., Shen, F., & Yang, H. B. "An overview of contour detection approaches". *International Journal of Automation and Computing*, 15, pp. 656-672, 2018.

[8] Juan, L., & Gwun, O. "A comparison of sift, pca-sift and surf". Inter*national Journal of Image Processing (IJIP),* 3(4), pp. 143-152, 2009.

[9] Wang, Y. Q. "An analysis of the Viola-Jones face detection algorithm". *Image Processing On Line*, 4, pp. 128-148, 2014.

[10] Gururaj, H. L., Soundarya, B. C., Priya, S., Shreyas, J., & Flammini, F. "A Comprehensive Review of Face Recognition Techniques, Trends and Challenges". *IEEE Access*, 20214.

[11] Pérez-Esteve C, Guilabert M, Matarredona V, Srulovici E, Tella S, Strametz R, Mira JJ. "AI in Home Care—Evaluation of Large Language Models for Future Training of Informal Caregivers: Observational Comparative Case Study". *JMIR*, 27:e70703, 2025.

[12] Ballester P, Pérez-Esteve C, Sánchez-García A, Gil-Hernández E, Guilabert M, Mira JJ. "Understanding the Roles of Male and Female Carers: A Comparative Analysis of Informal Caregiving over 65-year-old". *Geriatrics*, 10(3), 2025.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Fomentando el interés científico en las aulas mediante el Aprendizaje Basado en Juegos en áreas STEM: Matemáticas, Pensamiento Computacional y Ciberseguridad

Iván Marsá-Maestre[1], Juan R. Velasco[1], Pablo Garaizar[2], David Orden[3], Andrea de la Fuente[4], Rocío Garrido[4], Belén Palop[5]
[1]Departamento de Automática, Universidad de Alcalá
[2]Facultad de Ingeniería
[3]Departamento de Física y Matemáticas, Universidad de Alcalá
[4]Departamento de Didácticas Específicas, Universidad Autónoma de Madrid
[5]Departamento de Didáctica de las Ciencias Experimentales, Sociales y Matemáticas, Universidad Complutense de Madrid

[1,3]Escuela Politécnica Superior, Crtra. A2, Km. 32,600, 28805, Alcalá de Henares
[2]Facultad de Ingeniería, Avda. de las Universidades, 24, 48007, Bilbao
[4]Facultad de Formación de Profesorado y Educación, Francisco Tomás y Valiente, 3, 28049, Madrid
[5]Edificio La Almudena, C/ Rector Royo Villanova, 1, Ciudad Universitaria, 28040, MADRID

ivan.marsa@uah.es, juanramon.velasco@uah.es, garaizar@deusto.es, david.orden@uah.es, andrea.delafuente@uam.es, rocio.garrido@uam.es, b.palop@ucm.es

Entre julio de 2023 y junio de 2024, el Proyecto JUEGA (JUegos Educativos para Generar interés científico en las Aulas) ha permitido producir un total de 1400 copias de dos juegos diferentes, que se han distribuido en centros educativos y entre profesionales de la educación. Estos juegos persiguen atraer vocaciones a la formación universitaria STEM, en general, y a las ingenierías y las matemáticas, en particular. Tras los buenos resultados de este proyecto, arranca en julio de 2025 su continuación, con el desarrollo de dos nuevos juegos, esta vez centrados en la ciberseguridad y las matemáticas, respectivamente.

***Palabras Clave*- Aprendizaje Basado en Juegos, Programación, Internet, Matemáticas, Docencia, Divulgación**

I. Introducción

Los proyectos JUEGA (JUegos Educativos para Generar interés científico en las Aulas), financiados por la FECYT (FCT-22-18656 y FCT-24-20718), tienen como objetivos principales la utilización de juegos de mesa para contribuir al desarrollo de habilidades relacionadas con las Tecnologías de la Información y las Comunicaciones (TIC) entre estudiantes de educación secundaria, así como contribuir a despertar vocaciones para el estudio de estas disciplinas en la universidad, especialmente en mujeres. En el marco del primer proyecto (realizado entre julio de 2023 y junio de 2024) se han diseñado, desarrollado y producido dos juegos de mesa: H3L4D0S, centrado en el desarrollo del pensamiento algorítmico y la programación y Net'wor Forest, que permite comprender la esencia de los mecanismos de enrutamiento en redes de comunicación. Ambos juegos están orientados a estudiantes de último ciclo de primaria y de toda la secundaria. Asimismo, se han repartido más de 600 ejemplares de cada juego en diferentes colegios e institutos para su utilización en las aulas. Finalmente, se han entregado copias a profesionales del ámbito de educación que trabajan con niños y niñas utilizando habitualmente juegos de mesa.

El segundo proyecto JUEGA mantiene los mismos objetivos y se extiende entre julio de 2025 y junio de 2026, desarrollando y distribuyendo dos nuevos juegos: Alerta Zorro, orientado al aprendizaje de conceptos matemáticos en el ámbito de la probabilidad y Ctrl+Z, con un trasfondo que permite ilustrar los conceptos básicos de utilización de cortafuegos en sistemas telemáticos.

Para la difusión y seguimiento de estos proyectos se ha desarrollado una web (https://juegafecyt.web.uah.es), en la que es posible conocer de primera mano los juegos desarrollados, obtener materiales educativos asociados e incluso descargar versiones *print and play* de los juegos.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

II. Estado del arte y justificación

El número de estudiantes STEM que tenemos en las universidades españolas es insuficiente para las necesidades de nuestra industria y nuestra sociedad. En el curso 2022/2023, el número de estudiantes STEM suponía el 24,3%[1] del total, lo que supone un claro descenso desde 1998, cuando se situaba en un 33%. Por otra parte, si analizamos el sesgo de género en el ámbito de las TIC y las Matemáticas, en el curso 2021/2022 el 36,9% de las personas matriculadas en Matemáticas eran mujeres [1]. En el caso de las ingenierías (en general), la cifra desciende al 26,5%, situándose en el 14,9% en el caso específico de la Informática. En el ámbito de la ciberseguridad nos encontramos con un 24% de mujeres [2]. Por todo ello, es urgente realizar acciones que motiven a estudiantes de todas las etapas a descubrir el mundo STEM, en general, y de las TIC y las matemáticas, en particular.

El uso de juegos en el aula es una actividad cada vez más extendida y forma parte de la metodología ABJ (Aprendizaje Basado en Juegos). La literatura en este ámbito es extensa y hoy en día pocas personas del ámbito de la educación dudan de su efectividad cuando se usan correctamente [3], [4] y [5]. Estos proyectos pretenden colaborar con el aumento del conocimiento y la comprensión de la ciencia y la tecnología, a través del aprendizaje basado en juegos y con un enfoque de género neutro. Todo ello con el objetivo de estimular la curiosidad de las estudiantes por comprender más profundamente la ciencia que se esconde en ella para prevenir el actual rechazo al estudio de la tecnología.

Tras la primera edición del proyecto JUEGA, donde se han producido y entregado los dos primeros juegos, para el segundo proyecto se prevén cuatro acciones íntimamente relacionadas entre sí. Inicialmente, diseñaremos y publicaremos dos nuevos juegos de mesa que introduzcan conceptos de las matemáticas y de las TICs. A continuación, elaboraremos guías didácticas y realizaremos talleres de formación para el profesorado que imparte las asignaturas de Matemáticas y Tecnología, de manera que puedan utilizar estos juegos en el aula, y que hagan ver a sus estudiantes de manera sencilla la importancia de las temáticas abordadas en nuestras vidas. En tercer lugar, difundiremos las diferentes actividades del proyecto, tanto a través de la web del proyecto (https://juegafecyt.web.uah.es), como de otras acciones específicas dentro del plan de comunicación del proyecto. Por último, realizaremos un concurso de diseño de juegos educativos STEM, para explorar nuevas posibilidades para futuras ediciones del proyecto.

III. Juegos desarrollados

Los juegos desarrollados en el proyecto están pensados para jugarse de manera colaborativa, de manera que se fomenten, al mismo tiempo, otras competencias, como el trabajo en equipo o la gestión del liderazgo. En el primer proyecto JUEGA se han desarrollado los juegos H3L4D0S y Net'wor Forest. En el segundo proyecto JUEGA se están desarrollando los juegos Alerta Zorro y Ctrl+Z. Debe tenerse en cuenta que, en el momento de escribir el presente trabajo, estos dos últimos juegos se encuentran en fase de prototipo, por lo que su descripción puede variar respecto al juego que finalmente se lleve a producción.

En el diseño de los juegos se ha tenido en cuenta la inclusión en cuanto al género, evitando los tradicionales sesgos [6], trabajando desde la cooperación e incluyendo personajes femeninos no estereotipados. Consideramos muy importante la visibilización, tanto de los referentes femeninos en los sectores de las matemáticas, la computación y las redes, como de los desafíos y cambios que se demandan para el papel de la mujer en los mismos, con el fin de promocionar vocaciones y oportunidades para las nuevas generaciones que redunden en una igualdad real. En esta misma línea de promover una mayor afectividad de las niñas por estos campos, se ha optado por la versión cooperativa en los juegos. Numerosos estudios, como [7], establecen diferencias entre géneros que inclinan la balanza en contra de las niñas en cuanto a la menor asertividad y menor tendencia a interrumpir las conversaciones. En estas diferencias destaca especialmente la mayor ansiedad matemática [8], identificada como uno de los orígenes de la brecha en las carreras STEM y un problema de mayor gravedad en el contexto español particularmente. Los juegos cooperativos, con una adecuada gestión de la colaboración por parte del docente, favorecen la participación de todo el estudiantado al favorecer espacios seguros de debate desde un clima de tolerancia al error y de reducción de la ansiedad [9].

Como novedad del segundo proyecto respecto al primer proyecto JUEGA, se ha abierto una línea de cofinanciación del proyecto por medio de una actividad de micromecenazgo, por lo que será posible adquirir los juegos para uso personal. De esta manera, los beneficios obtenidos mediante crowdfunding servirán para incrementar el presupuesto del proyecto JUEGA-II y llegar a más centros educativos.

A. H3L4D0S

H3L4D0S es un juego solitario ambientado en una heladería controlada por el robot BIT, que ha perdido sus programas en memoria. Por este motivo, quien juega tendrá que ayudar a preparar los pedidos de los clientes, programando el funcionamiento del robot mediante unas cartas que permiten realizar bucles y acciones. Gracias a este juego, es posible desarrollar competencias relacionadas con el pensamiento computacional, tales como el pensamiento algorítmico, el reconocimiento de patrones o la abstracción [10].

El juego tiene dos modos: uno más sencillo en el que hay que ejecutar el programa para fabricar el helado, y otro más complejo, en el que debemos imaginar el programa necesario para hacer el helado que nos muestran. Un ejemplo de correspondencia entre programa y helado, formado con elementos del juego, puede verse en la Fig. 1.

B. Net'Wor Forest

Net'wor Forest es un juego ambientado en un bosque en el que han desaparecido las señales en los cruces de caminos, por lo que las personas que habitan en los pueblos que rodean el bosque no pueden atravesarlo. A lo largo de la partida las hadas (quienes están jugando) irán reponiendo las flechas en los cruces. Al mismo tiempo, las brujas queman las flechas de

[1] Elaboración propia a partir de "La universidad española en Cifras 2021/2022", disponible en https://www.crue.org/wp-content/uploads/2024/06/UEC-2021-2022.pdf (accedido el 10 de octubre de 2024) y de la "Estadística de Estudiantes Universitarios", disponible en https://www.universidades.gob.es/wp-content/uploads/2023/06/Principales-resultados_EEU_2022-23.pdf (accedido el 10 de octubre de 2024).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

algunos postes o ponen trols en los caminos, dificultando los avances de las hadas y de las personas que atraviesan el bosque. El juego tiene un trasfondo en el que se puede apreciar el funcionamiento de Internet, con los *routers* como cruces de caminos en el bosque, los caminos como las líneas de comunicaciones que unen esos *routers*, y los pueblos de alrededor como los equipos que intercambian información. Puede verse una representación de la construcción de rutas origen-destino con elementos del juego en la Fig. 2.

Fig. 1. Carta de reto (izquierda), minicartas de programación (arriba-derecha) y carta de solución (abajo-derecha)

Para ampliar las posibilidades del juego, hemos desarrollado una versión competitiva ambientada en una fiesta en el bosque, en el que cada jugadora representa a una ciudad y tiene que conseguir atraer personas hasta ella. Gana quien consiga atraer una persona de cada ciudad vecina antes de que lo haga el resto.

Fig. 2. Imagen de promoción de Net'wor Forest

C. *Alerta Zorro*

Alerta zorro forma parte de la investigación de la tesis doctoral de Andrea de la Fuente. En el juego se utilizan los materiales más frecuentes en la enseñanza-aprendizaje de la probabilidad: dados, urnas, ruleta y cartas. A estos elementos de azar se les han incluido dinámicas, mecánicas y estéticas propias de los juegos para ofrecer una nueva situación de aprendizaje. El objetivo es llenar cuatro cestas con moras rojas y negras. En ciertos momentos, aparece un zorro, ciego y hambriento, al que tendremos que ofrecer una de las cajas para que coja una mora. Sabiendo que le gustan las moras negras, el grupo tendrá que decidir qué caja ofrecerle (¡si coge una mora roja, pierden!). En este diálogo, los estudiantes van desarrollando su razonamiento probabilístico y la comparación intuitiva entre cajas, que conecta el razonamiento proporcional y la probabilidad clásica.

D. *Ctrl+Z (deshacer la infección)*

Ctrl+Z está ambientado en un asentamiento post-apocalíptico en el que tendremos que informar a las patrullas de guardas que protegen los accesos de qué transeúntes de los que llegan son zombis o humanos. A lo largo de la partida, será necesario permitir que entren los humanos con las mercancías necesarias para abastecer la ciudad (comida), defenderla (munición) y poder encontrar una cura (objetivo del juego) pero habrá que fijarse bien en la ropa que llevan, que es lo que nos va a permitir distinguir humanos de zombis. La forma en la que informaremos a las patrullas de guardas es muy similar a las reglas de configuración de los cortafuegos que protegen las redes telemáticas y los transeúntes que entran a la colonia se asemejan a los paquetes de datos, en los que podemos inspeccionar sus cabeceras y su contenido para tomar una decisión sobre si pueden pasar o no.

IV. Difusión del proyecto

A. *Talleres*

Durante 2024 se han realizado diferentes talleres, tanto con profesorado como con estudiantes de diferentes edades, en los que se han dado a conocer los juegos desarrollados en el primer proyecto y se han mostrado los materiales educativos que se han preparado. En el caso de talleres para profesorado, la actividad se inició con una presentación de las posibilidades de la informática "desenchufada" (i.e. sin usar ordenadores) para introducir determinados conceptos en los estudiantes de primaria y secundaria. En este ámbito, los juegos de mesa son una de las mejores herramientas para ello. A continuación, se jugaron varias partidas a ambos juegos para poder conocer sus posibilidades y aprender a utilizarlos en las aulas.

Por su parte, en las actividades con estudiantes se jugó con los juegos, dedicando unos minutos al final de cada partida para descubrirles qué hay detrás de cada juego: programación de ordenadores, funcionamiento de Internet, etc.

B. *Materiales educativos*

En el primer proyecto se han desarrollado diferentes materiales educativos. Estos materiales incluyen una guía didáctica para cada uno de los juegos, con el objetivo de que puedan ser aplicados en diferentes situaciones, así como un estudio de las asignaturas en las que son susceptibles de ser utilizados en las tres comunidades autónomas en las que trabajan los investigadores del proyecto: Castilla-La Mancha, Euskadi y Madrid.

C. *Plan de comunicación de JUEGA-II*

En el segundo proyecto JUEGA se continuará el trabajo mediante talleres presenciales con docentes, donde se

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

presentarán los juegos, se dará formación con un enfoque didáctico, y se proporcionarán todos los materiales para su uso en las aulas. Por otro lado, se organizarán talleres especialmente diseñados para el público infantil y juvenil, donde los niños y niñas no solo tendrán la oportunidad de jugar, sino también de descubrir de manera divertida el potencial educativo de los juegos. También se realizarán versiones de los talleres en familia para disfrutar y aprender de los juegos fortaleciendo vínculos. Como en el caso del proyecto anterior, para aquellas personas interesadas que no puedan acudir presencialmente, se pondrán a su disposición en la web las guías de los juegos, las versiones imprimibles de los juegos (*print and play*) y materiales de ampliación.

Adicionalmente, la participación en foros especializados, tanto en educación como en divulgación científica y juegos de mesa, contribuirá a incrementar el interés por el proyecto y sus resultados, tanto entre el profesorado como entre el estudiantado y sus familias. Además, esperamos captar la atención de alguna editorial de juegos de mesa que pueda amplificar el impacto del proyecto, haciéndolo accesible al gran público. Se realizarán notas de prensa de los eventos más relevantes a los que se asista. Para asegurar nuestra visibilidad en estos espacios, se emplearán *banners* o *rollups* informativos y se distribuirá información y materiales a los asistentes.

D. Concurso de prototipos de juegos de mesa

Con el objetivo de dar mayor visibilidad a más juegos de mesa destinados a facilitar la comprensión y el interés por las materias STEM, hemos convocado un concurso de prototipos de juegos de mesa científicos. Los prototipos finalistas serán mostrados en la página web del concurso y presentados a algunas editoriales que están desarrollando líneas específicas de juegos para el aula.

En el concurso puede participar cualquier autor o autora que tenga un juego en proceso de desarrollo (o terminado) y que no haya sido editado anteriormente, ni esté ligado ya a un contrato editorial (o se haya anunciado algún tipo de financiación colectiva). Esta condición debe cumplirse a lo largo de todo el concurso, hasta el momento en el que se haga público el fallo.

Puede presentarse más de un juego por autor o autora, pero no se podrán seleccionar dos juegos finalistas diseñados por una misma persona (incluso formando parte de un equipo de autores o autoras). Los resultados del concurso se darán a conocer en un evento realizado por el proyecto en el que se presentarán los juegos producidos.

V. Evaluación del impacto

El principal objetivo de estos proyectos es promover la utilización de los juegos diseñados en las aulas de centros educativos. Para medir el impacto en esta línea, hemos recogido datos de cada centro al que se le han entregado copias de los juegos desarrollados en el primer proyecto, incluyendo el número de estudiantes con los que se prevé utilizar el juego en cada curso.

El juego ha tenido una acogida muy favorable en todos los encuentros con profesorado que hemos organizado. El profesorado de los 77 centros que disponen de copias de los juegos prevé utilizarlo con más de 7.200 niños, niñas y adolescentes cada año.

Con respecto al impacto del sitio web, en la fecha en la que se envía este trabajo, ha recibido 1769 visitantes únicos, que han realizado un total de 7918 visitas, incluyendo en estas 3316 descargas de los materiales facilitados, tanto guías didácticas como versiones *print and play* de los juegos.

El impacto cualitativo lo hemos medido por medio de encuestas a las personas que recibían los juegos y/o participaban en los talleres. Todas las valoraciones recibidas han sido muy positivas, por encima de 8 sobre 10.

VI. Conclusiones

A la vista de la buena acogida que han tenido los talleres y materiales realizados en el primer proyecto JUEGA, y las previsiones de uso por parte del profesorado, consideramos que la iniciativa ha contribuido con éxito a favorecer la divulgación científica y la promoción de vocaciones STEM en niñas, niños y adolescentes. Esto ha animado a quienes firman este trabajo a dar continuidad a esta línea por medio del segundo proyecto JUEGA. En estos momentos, el equipo de desarrollo ya está trabajando en nuevos juegos para incorporar conceptos de matemáticas y ciberseguridad. Confiamos en que, de esta manera, podamos aumentar el número de profesionales TIC, al tiempo que colaboramos para reducir la brecha de género existente.

Agradecimientos

Los proyectos JUEGA y JUEGA-II están financiados por el Ministerio de Ciencia, Innovación y Universidades y la Fundación Española para la Ciencia y la Tecnología (FECYT) a través de los proyectos FCT-22-18656 y FCT-24-20718.

Referencias

[1] Cobreros, L., Galindo, J. & Raigada, T. (2024) Mujeres en STEM Desde la educación básica hasta la carrera laboral, ESADE, disponible en https://www.esade.edu/ecpol/wp-content/uploads/2024/03/Mujeres-en-STEM-2024-1.pdf (accedido el 10 de octubre de 2024)

[2] Women in Cybersecurity, An (ISC)^2 Cybersecurity Workforce Report, (ISC)2, 2020. Disponible en https://www.isc2.org/-/media/Project/ISC2/Main/Media/documents/research/ISC2-Women-in-Cybersecurity-Report.pdf (accedido el 10 de octubre de 2024).

[3] Ariffin, M. M., Oxley, A., & Sulaiman, S. (2014). Evaluating game-based learning effectiveness in higher education. Procedia-Social and Behavioral Sciences, 123, 20-27.

[4] López-Fernández, D., Gordillo, A., Alarcón, P. P., & Tovar, E. (2021). Comparing traditional teaching and game-based learning using teacher-authored games on computer science education. IEEE Transactions on Education, 64(4), 367-373.

[5] Pratama, L. D., & Setyaningrum, W. (2018). GBL in math problem solving: Is it effective?. International Journal of Interactive Mobile Technologies, 12(6).

[6] Robaina, J. R. S. (2025). ¿ Qué aprendemos de "lo femenino" con juegos didácticos de mesa de literatura y arte?: análisis crítico del discurso y Ecologías de aprendizaje para una Matriz de Datos en la ES. In *Innovación Educativa en el ámbito de las TIC y las TAC: proyectos, métodos y herramientas para el futuro de la Educación* (pp. 283-293). Dykinson.

[7] Hyde, J. S. (2005). The gender similarities hypothesis. American psychologist, 60(6), 581.

[8] PISA (2022). Programa para la Evaluación Internacional de los Estudiantes. Informe español.

[9] Dondio, P., Gusev, V., & Rocha, M. (2023). Do games reduce maths anxiety? A meta-analysis. *Computers & Education*, *194*, 104650.

[10] Palop, B., Díaz, I., Rodríguez-Muñiz, L. J., & Santaengracia, J. J. (2025). Redefining computational thinking: A holistic framework and its implications for K-12 education. *Education and Information Technologies*, 1-26.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Secure and Integrated LoRaWAN-5G Platform for Unidirectional Communications

Dan Garcia-Carrillo, Ignacio Rodriguez, Diego Arias Sal
Departamento de Informática y Departamento de Ingeniería Eléctrica, Electrónica, de Computadores y Sistemas, Universidad de Oviedo
Campus de Viesques, 33204 Gijon, Asturias, Spain.
garciadan@uniovi.es, irl@uniovi.es, UO263685@uniovi.es

The distinct capabilities of Low-Power Wide-Area Networks (LPWAN) and 5G present a compelling case for their integration. This paper proposes a novel architecture that utilizes a 5G network as a backhaul for LoRaWAN, designed to support massive IoT deployments with minimal configuration overhead. This proposal introduces new message formats, enabling the LoRaWAN network server to communicate directly with the 5G core. It extends the 5G framework to handle LoRaWAN-originated sessions to manage device subscriptions and data transmissions. This integration creates a flexible, multi-technology platform that combines the long-range, low-power benefits of LoRaWAN with the high-performance, reliable connectivity of 5G, paving the way for more sophisticated and scalable IoT solutions.

***Palabras Clave*—5G, Internet of Things (IoT), LoRaWAN, security, integration, convergence**

I. INTRODUCTION

Communication technologies have evolved to meet diverse needs, from low-power IoT networks like LoRaWAN, ideal for long-range and energy-efficient applications, to the high-speed, low-latency capabilities of 5G mobile networks. While each technology excels in specific scenarios, their integration offers a powerful solution that combines their respective strengths. A LoRaWAN-5G integration can create a flexible and efficient platform, supporting applications that range from simple sensor monitoring to those demanding real-time, reliable connectivity.

The growing popularity of the Internet of Things (IoT) and the evolution of fifth-generation (5G) cellular networks have driven significant interest in the integration of low-power wide area network (LPWAN) technologies into the 5G ecosystem [1], [2], [3], [4], [5], [6]. LPWAN technologies, such as LoRaWAN, Sigfox, and NB-IoT, are characterized by their low power consumption, wide communication range, and low cost, making them ideal for massive machine-type communication (mMTC) applications [1], [2], [3]. Although LPWANs are promising solutions for IoT, their efficient integration with 5G networks is essential to leverage the benefits of both technologies [1], [2]

This work proposes a novel integration architecture that uses a 5G network as a backhaul substrate for LoRaWAN. The primary goal is to deliver information from LoRaWAN end-devices to an application server in a massive deployment scenario, while significantly minimizing the need for manual configuration and interaction from a deployment manager.

To achieve this, we introduce new message formats for data exchange between the LoRaWAN network server and the 5G User Plane Function (UPF). The 5G core is enhanced to process these messages, with the Session Management Function (SMF) handling session management, and new databases are introduced in the Unified Data Repository (UDR) to manage integrated subscribers and their transmissions. This architecture leverages the robust 5G infrastructure to provide a scalable and streamlined backhaul solution for large-scale LoRaWAN deployments. This way we will be able to provide seamless unidirectional communication (from device to application server) to LoRaWAN devices through a 5G network with the goal of being suitable for even the most constrained LoRaWAN devices.

II. BACKGROUND AND STATE OF THE ART

A. Background on LoRaWAN and 5G standalone

In this section we briefly cover the basics on LoRaWAN and 5G standalone to address the integration process.

1) LoRaWAN: LoRaWAN provides an architecture of 3 entities, illustrated in Figure 1: i) LoRaWAN End Device, ii) Network Server, and iii) Application Server.

The end device, depending on its class, can have different communication capabilities. For example, Class A allows the device to transmit only at scheduled times or

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

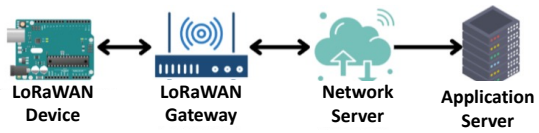


Fig. 1. LoRaWAN architecture

after receiving a signal from the Network Server, making it ideal for applications requiring low energy consumption. Class B enables the device to receive listening windows synchronized by periodic beacons, allowing for more predictable communication. Finally, Class C facilitates continuous bidirectional communication, as the device is always open to receiving messages, which is useful in applications requiring constant and real-time interaction. This proposal suggests working with Class A nodes.

Regarding security, for the LoRaWAN device to join the network managed by the Network Server, it must have pre-installed credentials (ABP) configured or perform the Over-The-Air Activation (OTA) joining process to authenticate and operate as a trusted entity within the network. At this point, the device has the necessary keys to communicate securely with the Application Server.

2) 5G standalone: A common standalone (SA) 5G network, also known as native 5G, is characterized by operating independently of the existing 4G LTE infrastructure. This means it uses a new generation network core (5GC) and a new radio interface (5G NR - New Radio) [7].

The 5GC is based on a modular architecture in which Network Functions (NF) communicate with each other through well-defined interfaces (services). This facilitates flexibility, scalability, and the introduction of new services. Below, we succinctly define the different control functions - illustrated in the figure 2 —involved in this proposal:

- AMF (Access and Mobility Management Function): Manages the connection and mobility of user equipment (UE). It is responsible for authentication, authorization, connection management, mobility management, and UE registration.
- NRF (Network Repository Function): Acts as a central registry where different NFs can discover and communicate with each other.
- AUSF (Authentication Server Function): Handles the authentication of users accessing the network.
- SMF (Session Management Function): Manages the establishment, modification, and release of user sessions (PDU Sessions). It controls IP address allocation, Quality of Service (QoS), and UPF selection.
- UPF (User Plane Function): Responsible for forwarding and routing user data packets. It acts as an anchor point for intra and inter-RAT mobility and can also perform deep packet inspection and apply QoS policies.
- UDM (Unified Data Management): Stores user subscription information and related policies.
- UDR (Unified Data Repository): A database that stores and manages subscriber data, SIM identities, and network service configurations.

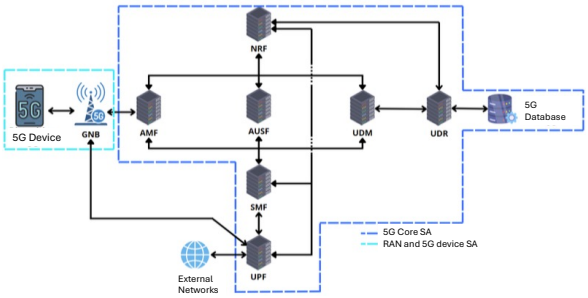


Fig. 2. 5G standalone architecture

B. State of the Art on 5G-LPWAN integration

Various research efforts have addressed the challenges and potential solutions for the integration of LPWAN into 5G, focusing on aspects such as architecture, security, mobility, interoperability, and coexistence with other wireless technologies [2], [3].

1) Challenges of LPWAN-5G Integration: The integration of LPWANs into the 5G System (5GS) presents several key challenges.

LPWAN Heterogeneity: Given the diversity of LPWAN technologies (with distinct architectures, protocol stacks, and radio link characteristics), seamless integration requires considering these differences [1], [3].

Interoperability of Solutions: Integration solutions must work consistently across different LPWAN technologies [1], [2].

Mobility Support: Unlike 5G, many LPWANs are not typically designed to support device mobility, which limits their applicability in certain IoT scenarios [1], [2], [3]. Managing device mobility between different connection points and access technologies is a significant challenge [1].

Scalability: LPWANs are designed for mMTC, which implies many simultaneously connected devices. The integration must scale efficiently to handle this massive volume [1], [2], [3].

Network and Signaling Compatibility: The 5G Core (5GC) architecture is based on a Service-Based Architecture (SBA) with interconnected Network Functions (NFs), whereas LoRaWAN has a simpler architecture. The direct integration of LoRaWAN functionalities into 5GC NFs, as well as signalling compatibility, are significant challenges [1], [2], [3].

Security: Protecting device identity, key derivation and encryption, and unifying authentication procedures are critical challenges, as 5G and LPWANs use different security schemes, which are often simplified in LPWANs due to the cost and energy constraints of the devices [1], [2], [3].

2) Existing Integration Proposals: Several options have been explored to integrate LoRaWAN with cellular networks [2]:

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Integration via 3GPP RAN: In this approach, the LoRaWAN gateway connects to a cellular network's eNodeB (LTE) or gNodeB (5G New Radio) by installing a Universal Subscriber Identity Module (USIM) and an IP stack in the gateway. This allows the gateway to access the cellular network as a User Equipment (UE) and use the cellular infrastructure as a backhaul [2], [5], [6].

Integration via Untrusted Non-3GPP RAN: The LoRaWAN gateway can connect to the 5GC through an evolved Packet Data Gateway (ePDG) function configured in the 5GC, which creates an IPSec tunnel for untrusted non-3GPP access networks (e.g., Wi-Fi) [2].

Modification of the LoRaWAN Gateway to act as an eNodeB: Navarro-Ortiz et al. [4] proposed that the LoRaWAN gateway act as a combination of a UE and an eNodeB from the perspective of the Evolved Packet Core (EPC). This approach does not modify the LoRaWAN end devices or the servers and is transparent to the EPC, as only the LoRaWAN gateway needs to be modified to implement part of the eNodeB protocol stack [4]. This solution ensures end-to-end security and is advantageous in terms of deployment costs compared to the previous approach [2].

Virtual Integration as part of the Core Network: LoRaWAN servers are installed in the cloud as part of the 5GC, using technologies such as Network Functions Virtualization (NFV) and OpenStack cloud platforms [2]. This approach is considered favorable for a convergent core-level integration [2].

The solution proposed in our work differs fundamentally from the reviewed approaches in that it requires no modifications to existing LoRaWAN equipment or the addition of USIM modules to any LoRaWAN devices. Instead, we propose an additional authentication layer implemented at the application level that integrates IoT device identities for authentication within the 5GC while preserving the existing LoRaWAN deployment architecture. Our approach operates independently of the LoRaWAN joining procedure, assuming that standard LoRaWAN device provisioning has already been completed, thereby maintaining seamless communication with the Network Server operating outside the 5GC. This design ensures backward compatibility while enabling enhanced security and integration capabilities within the 5G core network infrastructure. Furthermore, as we are using 5GC as a security enabler, we can choose to locate the LoRaWAN servers, either inside the 5GC as in the Virtual Integration case reviewed, or as an external server as shown in Figure 3 where we illustrate the concept.

3) Security Aspects in Integration: Security is a critical point in integration, given the divergence in the security schemes of 5G and LPWANs [3].

Identity Protection: 5G uses the Subscription Concealed Identifier (SUCI) to protect the Subscription Permanent Identifier (SUPI) and prevent its transmission in cleartext [1], [3]. In contrast, LoRaWAN sends join requests containing identifiers like DevEUI and JoinEUI in cleartext, which can compromise security, allowing for replay attacks, device location exposure, and even Denial-Of-Service (DoS) attacks [2], [3].

Key Derivation and Encryption: Integrating the key hierarchies is complex [2].

Unified Authentication Procedures: 5G defines bidirectional authentication methods such as 5G-AKA, EAP-AKA', and EAP-TLS [1], [3]. LPWANs, due to their limitations, often employ unidirectional authentication methods. Designing a unified authentication method that is compatible with 3GPP standards and the constraints of LPWANs is a challenge [2].

This work goes precisely into this last point by designing an authentication method that can be integrated with 5G and support the constraints of LPWANs such as LoRaWAN.

III. Secure and platform-agnostic LoRaWAN-5G integration

LoRaWAN (Long Range Wide Area Network) and 5G are not competitors but rather complementary technologies. LoRaWAN excels in long-range, low-power, and low-bandwidth communication. This makes it ideal for connecting a massive number of simple, battery-operated sensors over large geographical areas, often at a low cost. Think of applications like smart agriculture (soil sensors), smart cities (parking sensors, waste management), and industrial monitoring (equipment status).

LoRaWAN-5G integration follows a set of key motivations that seek to combine the strengths of both technologies to cover the growing demand brought by the IoT ecosystem. Concretely, in the context of Massive Machine Type Communications (mMTC) [1], typical cellular technologies on their own are not equipped to handle the connectivity of millions of devices per km^2. On the other hand, LPWAN technologies such as LoRaWAN have their own limitations in terms of access to the medium (duty cycle), very low bit rates, and limited payloads, which restrict their applicability in very demanding use cases [2]. 5G, with its performance, flexibility, scalability, agility, and programmability, can provide the connectivity and management functions needed to complement LoRaWAN.

Another interesting benefit of the 5G and LoRaWAN binomial is the enhancement of technical capabilities of LoRaWAN, such as security and privacy. LoRaWAN can benefit from 5G's robust security mechanisms. For instance, managing LoRaWAN servers within the 5G core could help keep data within the operator's network, increasing security and privacy [5].

Lastly, we will highlight the ease of gateway deployment. By using the 5G cellular network as backhaul for LoRaWAN gateways, it enables easier and more flexible deployment almost anywhere with cellular coverage.

These last two points are, in fact, the focus of this proposal: leverage 5G as an enabler, not just for facilitating the deployment of LoRaWAN infrastructure, but to leverage 5G security features to enhance LoRaWAN's security, providing secure and seamless integration between the end device and the application server through 5G.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

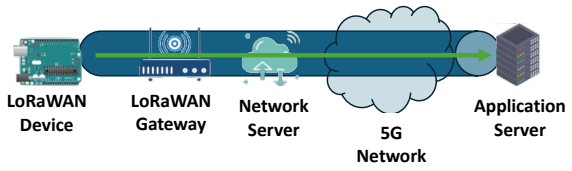


Fig. 3. LoRaWAN-5G integration concept

The concept can be seen in the Figure 3.

The main objective is to leverage 5G to facilitate LoRaWAN deployments and, at the same time, provide an additional layer of security on top of LoRaWAN joining procedure by taking advantage of 5G authentication and security features.

A. LoRaWAN baseline

The LoRaWAN network is assumed to work as expected by establishing the authentication process either through OTAA (Over-the-Air Activation) or by Activation by Personalization (ABP). Even though OTAA offers much better security, as the derived keys can be regenerated if compromised, we will present the additional security layer on top of LoRaWAN that minimizes the effects of the ABP keys being compromised, just allowing the device to send information to the Network server. If the device is deemed compromised by the additional security layer, the device will be blacklisted.

1) LoRaWAN security with ABP: LoRaWAN ABP (Activation By Personalization) is a method for devices to join a LoRaWAN network. The device is manually configured with the network parameters, including the DevAddr (a unique device address), AppSKey (an Application session key), and NwkSKey (a Network session key). Unlike OTAA (Over-The-Air Activation), ABP does not involve a join procedure exchange. The device simply starts transmitting data using the pre-configured parameters.

The device uses DevAddr to identify itself on the network. Message encryption and integrity are ensured using the AppSKey (for application data) and NwkSKey (for network control messages).

Even though this method can seem faster and simpler to deploy since no join process is needed, it is less secure and flexible, as it lacks dynamic session management and key refresh capabilities, which can compromise the deployment if the credentials are discovered.

2) LoRaWAN security with OTTA: Over-the-Air Activation (OTAA) is the preferred and more secure method for enrolling a device into a LoRaWAN network. This process involves a dynamic exchange of credentials between the end device and the network, ensuring that only authorized devices can join and that communication remains encrypted and secure.

At its core, OTAA is a ”join procedure” initiated by the end device. Devices that will perform OTAA are provisioned with a unique set of identifiers. These include a globally unique end-device identifier (DevEUI), an application identifier (AppEUI), and a secret application key (AppKey).

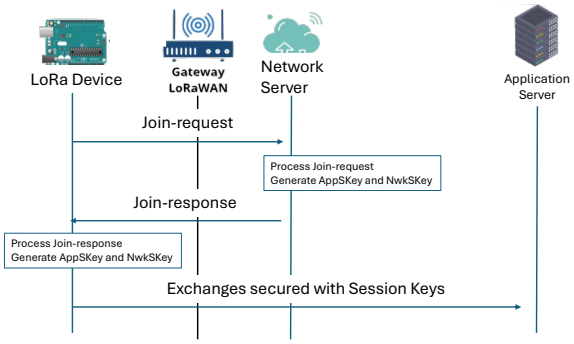


Fig. 4. LoRaWAN OTTA exchange

The activation process. As illustrated in Figure 4, it begins when an end device, wishing to join the network, sends a ”Join-request” message. This message, which is not encrypted, and contains the DevEUI and AppEUI of the device, along with a randomly generated value called a DevNonce. This request is picked up by one or more LoRaWAN gateways, which then forward it to the Network Server. The Network Server, upon receiving the request, first checks the DevNonce to ensure it hasn't been used before. It then uses the AppEUI to identify the appropriate Application Server and, in conjunction with the Join Server, validates the AppKey.

If the device is authorized, the Network Server generates two unique session keys for the device: the Network Session Key (NwkSKey) and the Application Session Key (AppSKey). The NwkSKey is used to secure the communication between the end device and the Network Server, ensuring the integrity of the messages. The AppSKey, on the other hand, provides end-to-end encryption for the application data payload between the end-device and the Application Server, meaning the network provider cannot read the actual application data.

Following the successful generation of the session keys, the Network Server sends a ”Join-accept” message back to the end-device. This message, which is encrypted using the AppKey, contains the newly generated session keys, a unique 32-bit device address (DevAddr) for the current session, and other network parameters.

Once the end-device receives and decrypts the Join-accept message, it securely stores the DevAddr and the two session keys. From this point on, the device is considered trustworthy and can begin securely transmitting data to the network using the established session keys.

B. 5G baseline

A 5G Standalone (SA) network is fundamentally composed of the Radio Access Network (RAN) and the 5G Core (5GC). The RAN features the gNodeB (gNB), which provides wireless connectivity to User Equipment (UE). The 5GC employs a Service-Based Architecture (SBA), where network functions (NFs) offer and consume services via standardized Application Programming Interfaces (APIs).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

1) Core Network Functions and Connectivity: The 5G Core is a collection of specialized Network Functions (NFs). The Network Repository Function (NRF) is central to the SBA, enabling NFs to discover and register their services, which facilitates redundancy and dynamic scaling. Key NFs and their interactions are as follows:

- Access and Mobility Management Function (AMF): The AMF serves as the primary entry point for all control plane signaling from both the UE (via the N1 interface) and the gNB (via the N2 interface). It is responsible for connection and registration management, mobility, and authentication signaling. Crucially, the AMF handles control plane traffic only; it does not process user data.
- Session Management Function (SMF): The SMF is responsible for managing the UE's PDU (Protocol Data Unit) sessions. This includes session establishment, modification, and release. The SMF selects the appropriate User Plane Function (UPF) and instructs it on how to handle traffic via the N4 interface (the control plane interface to the UPF).
- User Plane Function (UPF): The UPF is the system's anchor point for the user plane. It connects the RAN to external Data Networks (DN), such as the internet. Its primary roles include packet routing and forwarding, policy enforcement, and usage reporting. It is the gateway for all user data.
- Unified Data Management (UDM) and Unified Data Repository (UDR): The UDM is responsible for managing user subscription information. It retrieves this data from the UDR, which acts as a centralized, stateful data store. The UDR also holds policy data for the Policy Control Function (PCF) and structured data from other NFs. The UDR serves multiple NFs and is not exclusively connected to the UDM.

2) Communication Flows: Communications within the 5G system are bidirectional and are separated into two distinct planes:

- Control Plane: Signaling messages, such as those for authentication and session setup, flow from the UE to the gNB and are then forwarded to the AMF. The AMF orchestrates further signaling with other NFs like the SMF and UDM to manage the user's state and session policies. The SMF then establishes and manages the session by configuring the UPF over the N4 interface.
- User Plane: Once a PDU session is active, user data flows directly from the UE to the gNB. The gNB forwards this traffic to the selected UPF over the N3 interface. The UPF then routes the data to the target external network (e.g., the internet) via the N6 interface. This direct path from the gNB to the UPF ensures high-throughput, low-latency data transmission, bypassing the control plane functions for established sessions.

3) Basic 5G authentication: When a user initiates communications, they must send an authentication message. This will begin a two-level authentication process in which a set of security mechanisms will be used to authenticate the device using data stored in the subscriber database and the data sent by the device.

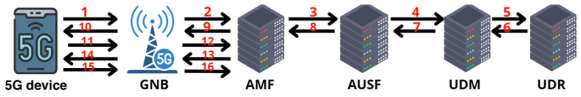


Fig. 5. 5G authentication

The process, illustrated in Figure 5, is as follows:

i. The device initiates communication by sending a registration request to the gNB.
ii. The gNB relays the request to the AMF. The AMF performs an initial verification (first level of authentication) based on the requesting device's identifier.
iii. The AMF forwards the request to the AUSF to begin the second level of authentication.
iv. The AUSF queries the UDM for the subscriber's authentication data (subscriber authentication vector).
v. The UDM requests the data from the UDR.
vi. The UDR retrieves the data and sends it to the UDM.
vii. The UDM relays the data to the AUSF.
viii. The AUSF uses the data to authenticate the device for the derivation of certain keys and relays a set of keys and authentication data to the AMF.
ix. The AMF generates a sequence from the authentication data and sends a challenge to the gNB.
x. The gNB relays the challenge to the device.
xi. The UE generates a response to the challenge and relays it to the gNB.
xii. The gNB forwards the response to the AMF.
xiii. The AMF checks that the response matches the expected response and, if so, sends the gNB a key and parameters that the device should use to encrypt future data messages.
xiv. The gNB relays the key and parameters to the device.
xv. The device replies that it has received the challenge and will begin using the received parameters and key.
xvi. The gNB retransmits the reply to the AMF.

C. 5G-LoRaWAN integration

At this point we assume that LoRaWAN and 5G baseline work as expected. The integration of the two involves some additions to the architecture, which is illustrated in Figure 6.

The authentication and initial session process using bidirectional transmissions between the LoRa device and its LoRaWAN network server remains identical to that explained in the Section A2 and is assumed to be secure. However, once authenticated to its network server, the one-way data messages from the LoRaWAN network, which will be received by the network server and shared with the application server, are now the UPF. This integration is depicted in a green box in Figure 6.

The communication with the 5G network will follow two new formats. These LoRaWAN data messages are divided into integrated network authentication messages

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

and integrated network data messages and are described in C4.

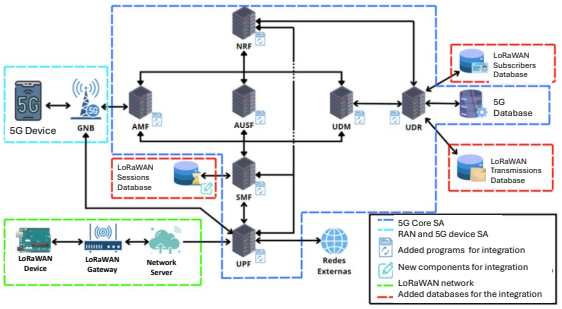


Fig. 6. Integration architecture

There is a prerequisite for users from the LoRaWAN network to subscribe to the integrated network. This registration will be done through an out-of-band process, which is out of the scope of this proposal.

Once the new integration messages can be exchanged between the network server with the UPF, the server will extract the payload from the message, and forward it to the SMF. The SMF, which is responsible for session management in 5G, will detect whether the received message is in the integrated network authentication format or in the data format.

To enable the 5G core to perform these new processes with the data received from the LoRaWAN network, the system will be equipped with three new databases: sessions, located in the SMF; subscribers; and transmissions, located in the UDR. The integration of these databases is depicted in red boxes in Figure 6.

1) 5G-integrated Authentication Message flow: Figure 7 shows the exchange to authenticate and open the data session. The authentication procedure provides a full authentication exchange when the device initiates the process for the first time.

Fig. 7. Authentication message flow

When an authentication message is detected, it is forwarded to the AMF, which will begin an authentication process that mimics the 5G authentication process, which is more robust than LoRaWAN. The one-way nature of this communication, unlike the two-way nature of 5G, does not allow for multiple steps that a 5G authentication would perform. Therefore, integration security focuses on the general concept of each step, mimicking it through processes with similar objectives that conform to one-way communication limitations, rather than replicating them. If authentication is successful, a data session will be opened for the integrated network user.

2) 5G-integrated Data Message flow: Figure 8 shows the exchange to send a data message.

Fig. 8. Data message flow

In the case of a data message, the SMF will check if there is an open session for the user in the integrated network. If so, and the data message is correct, it will be forwarded to the UDM, which in turn will be forwarded to the UDR for storage.

Messages, both authentication and data, will be protected, providing confidentiality, authentication, and integrity.

3) 5G-LoRaWAN integration key management: To ensure the messages are correctly processed and the identity of the user/device ascertained, key management is necessary. The device/subscriber must use these same methods to send and encrypt its messages. Otherwise, the messages will be discarded during their specific message flow, as the identity of the sender cannot be verified.

There is basic information that needs to be pre-shared between the device and the 5G network. This is, at least, a long-term credential, such as a pre-shared key (PSK) and the Device ID, which will be associated in the system.

Key derivation: The PSK will never be used directly. It is considered as a root key, from which we will derive additional keys to secure the communications.

To derive keys, the PBKDF2 algorithm will be used, passing as necessary arguments: the subscriber's pre-shared key (PSK) of 16 bytes, the device/subscriber identifier (DeviceID) of 3 bytes, and the random value generated for derivation (DerivationNonce) of 1 byte. The result is a derived key 32 bytes long.

Ciphering information: The ChaCha20 stream cipher algorithm will be used to encrypt message fields. The following will be passed as arguments: the 32-byte derived key, the 1-byte DerivationNonce, which will be extended to 16 bytes by adding leading zeros, and the fields to be encrypted. The result is the encrypted fields with the same original size.

Using the same DerivationNonce in the encryption as in the derivation does not affect security, since the derived key already provides the necessary "randomness" when used. This saves having to send more bytes for encryption.

Chacha20 requires a nonce of 8, 12, or 16 bytes to operate (depending on the version used), so in this case, the DerivationNonce is extended by adding leading zeros. This stream cipher was chosen because it requires fewer extra bytes to operate than others and does not alter the size of the original message.

4) 5G-integrated Authentication and Data Messages Format: For this integration, we propose two types of messages: authentication and data. The messages described here will travel as the payload of the LoRaWAN message to the 5G core, where they will be processed accordingly.

The messages are designed considering the limitations of LoRaWAN. To ensure the viability of the system ,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

even in the worst conditions, we design the two types of messages to adapt to LoRaWAN Spreading Factor 12 with a bit-rate of 250 bits/sec and a maximum payload of 51 bytes.

Authentication Message: It is responsible for opening a LoRaWAN-5G integration session. This message will ensure that the session is opened by a registered user being authenticated, similar to how it is done in LoRaWAN. The main difference is that, being unidirectional, authentication is only verified if successful, and then the session is activated. The packet structure is illustrated in Figure 9.

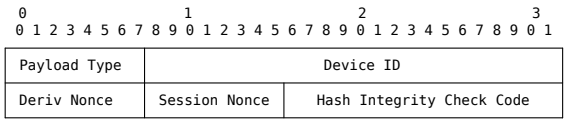


Fig. 9. 5G-integratation authentication message format

The fields are the following:

- PayloadType: Identifies the payload content type for proper handling. In the case of authentication messages, with value 0 (decimal) in 1 byte. This field is not encrypted.
- DeviceID: Unique identifier of the LoRa device sending the message. This is 3 bytes and will not be encrypted.
- DerivationNonce: Randomly generated value that will be used for encryption key derivation along with the DeviceID and the PSK. The field is 1 byte and will not be encrypted.
- SessionNonce: Randomly generated value that will be used for session identification and tracking if authentication is successful. This field is 1 byte and will be encrypted.
- HICC (HASH INTEGRITY CHECK CODE): The value of the first 3 bytes of the MD5 hash of the rest of the authentication message, not including these 3 bytes. The hash is calculated before encrypting the fields. This field is also protected.

In case the authentication fails, we have two options: 1) assume all data belonging to that session will be lost, 2) leverage the interconnection between 5G and LoRaWAN server to, send in in the reception window of the LoRaWAN device after sending the authentication message, a specific protected message indicating authentication failure. How this process is done is out of the scope of this proposal.

Data Message: Once the session is open, messages are sent, protected with credentials specifically derived for the session. This message is illustrated in Figure 10.

- PayloadType: Identifies the payload content type for proper management. Its value will vary depending on the origin, processing, or use of the data to be transmitted. (Example: 0b00000001 - Temperature sensor data). The value 0b00000000 is reserved for authentication messages, leaving 255 values free for specific uses. The field will be 1 byte and not encrypted.

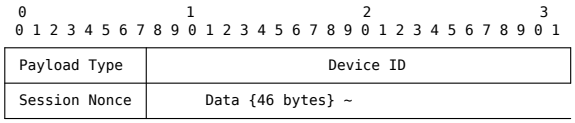


Fig. 10. 5G-integratation data message format

- DeviceID: Unique identifier of the LoRa device sending the message. This field will be 3 bytes and not encrypted.
- SessionNonce: Value used to track the open session. Its first value will be equal to the one used in the same field of the authentication message that opened the current session. It will then be incremented by 1 sequentially (in cycles of 0-255) for each message sent in the same session. This field will be 1 byte and encrypted to avoid replay attacks or encourage devices to forge future messages and generate a possible attack that requires processing the message.
- Data: Data to be transmitted depending on the payload type, use, and purpose. This field has no minimum size. Maximum size varies depending on the SF. Minimum capacity of at least 46 bytes (SF12) and will be protected. To provide additional security to the non-encrypted fields, depending on the use case, we leave it to the application to include a similar field to HICC (HASH INTEGRITY CHECK CODE) in the authentication message format, as part of the data structure, to avoid malicious modification.

IV. Experimental Results

Even though the limits of LoRaWAN, as studied by Ferran et al. [8], indicates the number of messages that can be sent with different number of devices, and different payloads, we test the system here with a higher number of messages to verify its feasibility, putting a higher threshold of 1000 messages to assess the workload of the system can support it and also to asses the time it takes to process each kind of message in the worst case scenario, that is, when the process of authentication messages and data messages do not produce errors that cause the process to stop abruptly.

The tests were performed using the system described in Table I.

O.S.	Ubuntu 22.04.5 LTS
CPU	AMD Ryzen 7 6800H
RAM	15534556 KiB

Tab. I
TESTBED SPECIFICATIONS

After the tests, we see how the system fares against a high message load, illustrated in Figure 11.

After sending 1000 successful authentication messages and 1000 correct data messages (with an open session), the

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

functions illustrated in Figure 11 show that the average processing time for an authentication message is much longer (approximately 32 times) than for a data message. This result is expected due to all the derivation, encryption, etc. processes that occur in the authentication message. This, nonetheless, is positive, since, in a real, typical communication, we will only have one authentication message per session, having the data message have a much more lightweight impact on the system.

Nonetheless, we have corroborated that the system is able to withstand a heavy load under the assumptions of the limits of LoRaWAN [8].

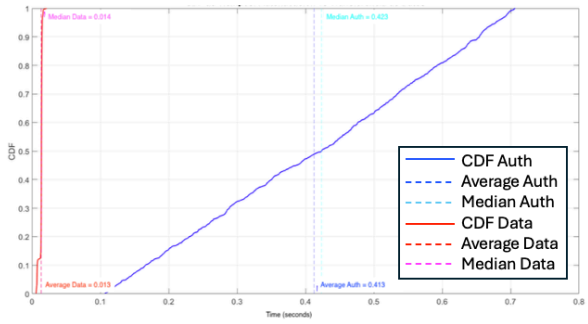


Fig. 11. Cumulative distribution function (CDF) of authentication message and data processing times

V. Discussion

Given that the message flow is designed for unidirectional communication, there are some limitations that have to be considered.

Message loss or corrupted: This affects differently to authentication messages that to data messages. When authentication messages are lost or corrupted, it will affect all data messages sent in the same session. If the application is not critical, this can just be assumed and will only process sessions following a correct authentication message. If we want to prevent this from happening, we can devise an intervention like the one commented on in C4. That a data message is lost or corrupted may have a lesser impact on the system, as just one item of information will be lost. If the application requires reliability, the same, approach can be implemented as to recover from failed authentication messages.

Fragmentation support: Another limitation in this case is the use of fragmentation. Even though it can be implemented, the complexity and time needed to complete the exchange of a fragmented message, accounting for message loss or corruption, do not encourage the use of this feature.

Limited feedback: The use of unidirectional communication in this case limits the information the device can obtain from the 5G network.

For these reasons, we acknowledge that this proposal should be limited to non-critical use cases.

VI. Conclusions and Future Work

This work proposes a novel LoRaWAN-5G integration, where the LoRaWAN device can send messages securely to the application server through a 5G network, enabling unidirectional communication from the LoRaWAN end device to the Application Server. We review why the concept of integration is interesting, and the current work in the area.

The main contribution of this work is the integration proposal and the use of new messages to perform an authentication within the 5G network as an additional layer of security on top of LoRaWAN.

This does not modify how LoRaWAN works, as it is seamless to standard LoRaWAN devices. We proposed adding, as part of the LoRaWAN application layer, a way of interacting securely with the 5G network. This achieves a trustworthy operation, and delivering the payload to the application server in a secure and integrated manner through 5G.

We have tested the performance according to the limits of LoRaWAN to ensure its feasibility.

Future work will focus on the integration with standardized 5G interfaces and explore other alternatives and use cases that could be supported by leveraging 5G features.

Acknowledgements

This work was supported in part by the Spanish Ministry of Science, Innovation, and Universities under Ramon y Cajal Fellowship number RYC-2020-030676-I, funded by MICIU/AEI/10.13039/501100011033 and by the European Social Fund "Investing in your future"; in part by project PID2023-148104OB-C43, funded by MICIU/10.13039/501100011033 and by FEDER, UE; and in part by the Fundación Universidad de Oviedo project FUO-25-147

References

[1] H. Jradi, F. Nouvel, A. E. Samhat, J.-C. Prévotet, and M. Mroue, "A seamless integration solution for LoRaWAN into 5G system," *IEEE Internet of Things Journal*, vol. 10, no. 18, pp. 16 238–16 252, 2023.

[2] Y. Chen, Y. A. Sambo, O. Onireti, and M. A. Imran, "A survey on LPWAN-5G integration: Main challenges and potential solutions," *Ieee Access*, vol. 10, pp. 32 132–32 149, 2022.

[3] J. Sanchez-Gomez, D. G. Carrillo, R. Sanchez-Iborra, J. L. Hernandez-Ramos, J. Granjal, R. Marin-Perez, and M. A. Zamora-Izquierdo, "Integrating LPWAN technologies in the 5G ecosystem: A survey on security challenges and solutions," *IEEE Access*, vol. 8, pp. 216 437–216 460, 2020.

[4] J. Navarro-Ortiz, S. Sendra, P. Ameigeiras, and J. M. Lopez-Soler, "Integration of LoRaWAN and 4G/5G for the Industrial Internet of Things," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 60–67, 2018.

[5] Y. Chen, Y. A. Sambo, O. Onireti, S. Ansari, and M. A. Imran, "LoRaWAN-5G integrated network with collaborative RAN and converged core network," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2022, pp. 1–5.

[6] R. Yasmin, J. Petäjäjärvi, K. Mikhaylov, and A. Pouttu, "On the integration of LoRaWAN with the 5G test network," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017, pp. 1–6.

[7] S. Ahmadi, *5G NR: Architecture, technology, implementation, and operation of 3GPP new radio standards*. Academic Press, 2019.

[8] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of LoRaWAN," *IEEE Communications magazine*, vol. 55, no. 9, pp. 34–40, 2017.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Un Sistema de Detección de Intrusos para IoT basado en DNN y Preprocesamiento Híbrido

Antonio Villafranca, Maria-Dolores Cano

Departamento Tecnologías de la Información y las Comunicaciones,

Universidad Politécnica de Cartagena

30202 Cartagena, España.

antonio.villafranca@upct.es , mdolores.cano@upct.es

La expansión del IoT ha multiplicado la superficie de ataque y exige IDS eficaces en entornos distribuidos y con recursos limitados. Se presenta una revisión del estado del arte y un IDS basado en redes neuronales profundas (DNN) con una tubería estandarizada de preprocesamiento (normalización, balanceo con SMOTE+Tomek y regularización) y validación cruzada estratificada. Se evaluó en CICIDS2017, UNSW-NB15 y BoT-IoT, así como en otros conjuntos públicos representativos de IoT. Para asegurar comparabilidad y reproducibilidad, el preprocesamiento y la evaluación se alinearon con los protocolos de referencia de cada dataset y se aplicaron de forma consistente en todos ellos.

Resultados: todas las evaluaciones superaron el 97% de exactitud; la media entre datasets fue 99,14%, con mejoras notables en clases minoritarias. Los hallazgos indican que una arquitectura densa bien regularizada, combinada con un preprocesamiento cuidadoso, ofrece alta precisión y baja latencia sin necesidad de modelos más complejos, lo que facilita su despliegue en pasarelas IoT.

***Palabras Clave*- IoT, IDS, DNN, SMOTE, Ciberseguridad.**

I. Introducción

La expansión exponencial del Internet de las Cosas (IoT) ha interconectado miles de millones de dispositivos en ámbitos como salud, industria y hogares o ciudades inteligentes, impulsando la automatización y la eficiencia operativa. Sin embargo, esta masiva conectividad también amplía la superficie de ataque. Muchos dispositivos IoT carecen de cómputo, memoria y energía para incorporar medidas de seguridad avanzadas [1]. Además, dependen de protocolos ligeros (p.ej., MQTT, CoAP, RPL) que priorizan la eficiencia sobre la confidencialidad e integridad [2] y, con frecuencia, no disponen de infraestructuras robustas de actualización, lo que prolonga la exposición a vulnerabilidades [3][4].

En este contexto, los sistemas de detección de intrusiones (IDS) resultan esenciales para monitorizar el tráfico y la actividad de los dispositivos, identificar patrones maliciosos mediante firmas o modelos de comportamiento, y reaccionar en tiempo real con alertas y acciones de mitigación.

Diversos estudios han analizado técnicas basadas en *Machine Learning* (ML) y *Deep Learning* (DL) para IDS en IoT, mostrando que las redes neuronales profundas (DNN) pueden capturar representaciones jerárquicas de las características del tráfico, mejorando la detección de ataques desconocidos [5][6]. Sin embargo, la heterogeneidad y desequilibrio de los datos, donde el tráfico benigno supera con creces al malicioso, plantean retos de sobreajuste y altas tasas de falsos positivos. Para mitigarlo han aparecido enfoques híbridos, generativos (p.ej., *Generative Adversarial Network*, GAN, *autoencoders*) y distribuidos (p.ej., *Federated Learning*, FL, *Edge Artificial Intelligence*), cada uno con sus ventajas en privacidad, latencia o precisión, pero con limitaciones en recursos y generalización [7][8].

Se presenta una revisión crítica del estado de la técnica en IDS para IoT y un IDS basado en redes neuronales profundas (DNN) diseñado para operar con baja latencia en pasarelas y nodos *edge*. La elección de DNN se justifica por su capacidad para modelar relaciones no lineales y jerárquicas en flujos de red heterogéneos, manteniendo un coste de inferencia moderado con arquitecturas densas. Para mitigar el desequilibrio de clases extremo típico del tráfico IoT, se integra una estrategia de balanceo con SMOTE y Tomek Links, junto con normalización y regularización por *dropout*, y se emplea validación cruzada estratificada.

La contribución principal no reside en la novedad aislada de cada técnica, sino en su integración controlada y reproducible dentro de una tubería estandarizada que alinea el preprocesamiento y la evaluación con los protocolos de referencia de cada conjunto de datos. El modelo se evaluó en CICIDS2017 [9], UNSW-NB15 [10] y BoT-IoT [11]; adicionalmente, se aplicó la misma tubería a CSE-CIC-IDS2018 [12], MQTT-IoT-IDS2020 [13], NSL-KDD [14], ToN_IoT [15], entre otros. Los resultados demuestran que todas las evaluaciones superaron el 97 % de exactitud; la media entre datasets fue 99,14%, mejorando significativamente la detección de clases minoritarias y superando de forma consistente a modelos especializados, evidenciando su eficacia, escalabilidad y robustez en escenarios diversos de IoT.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

El resto del trabajo se organiza así: en Estado del arte se revisan las líneas principales y se propone una taxonomía de IDS para IoT; Desafíos actuales sintetiza los retos operativos y Tendencias emergentes resume las líneas que están respondiendo a dichos retos. Propuesta de IDS basado en DNN para IoT describe el pipeline y el diseño del sistema (Preprocesamiento, Arquitectura de DNN y Regularización y validación). Evaluación experimental y comparativa detalla datasets, protocolo y resultados (incluyendo ablaciones y métricas operativas), además de las comparativas con el estado del arte. Finalmente, Discusión analiza implicaciones, limitaciones y operación (tasa de falsos positivos (FPR), umbrales y deriva de concepto) y Conclusiones cierra con los hallazgos y líneas futuras.

II. Estado del Arte

En la última década, la investigación sobre IDS en entornos IoT ha pasado de trabajos aislados a revisiones sistemáticas que integran cientos de estudios para ofrecer una visión global de los avances, desafíos y oportunidades. La taxonomía de la Fig. 2 organiza los IDS para IoT en cinco dimensiones: (1) Despliegue/ubicación (centralizado, distribuido; nube, fog, borde/edge), (2) método de detección (basados en firmas, basados en anomalías, basados en especificaciones e híbridos), (3) evaluación del rendimiento (precisión, recall, F1, tasa de falsos positivos, tiempo de detección/latencia, consumo energético), (4) escenarios de aplicación (ciudades inteligentes, industria, salud, defensa, finanzas, energía, agricultura) y (5) metodologías (ML, DL, FL, XAI, blockchain).

En método de detección, "basados en especificaciones" cubre enfoques que verifican el cumplimiento de reglas/contratos de comportamiento; cuando la literatura usa "behavior-based/BIDS", se mapea a anomalías o especificaciones según el criterio del autor. En despliegue/ubicación, se distinguen centralizado vs. distribuido y nube/fog/borde como planos complementarios.

La revisión del estado del arte revela que la mayoría de las soluciones convergen en arquitecturas híbridas que combinan aprendizaje profundo con reglas o firmas, con el fin de equilibrar la capacidad de descubrir ataques desconocidos y la eficiencia computacional. Además, muestran un claro énfasis en escenarios industriales y de ciudades inteligentes, donde el volumen de datos y la criticidad del servicio demandan detección en tiempo real con mínimos falsos positivos. El tipo de implementación más habitual es el enfoque distribuido, que delega capacidades de detección en *gateways* o nodos *edge* para reducir la latencia y el tráfico hacia la nube. Ejemplos recientes incluyen FLADEN, un marco de aprendizaje federado que alcanza más del 99,9 % de precisión mientras preserva la privacidad de los datos locales [16]. En contraste, los esquemas centralizados mantienen modelos complejos en servidores potentes a costa de mayor latencia y suelen usarse en entornos donde los dispositivos IoT tienen poca capacidad de cómputo.

Asimismo, las propuestas basadas en *blockchai*n se emplean para asegurar la integridad de los registros de intrusión. Li *et al*. [8] demostraron que un *framework* colaborativo sobre *blockchain* reduce en un 18 % las manipulaciones de *logs* y mejora la detección en *smart cities*. En cuanto a conjuntos de datos, la evaluación de la mayoría de los trabajos se realiza sobre conjuntos sintéticos como UNSW-NB15 y NSL-KDD, así como específicos de IoT como BoT-IoT, N-BaIoT e IoT-Traffic2020, aunque su falta de realismo y desequilibrios internos limitan la generalización de los resultados y han impulsado estudios recientes que fusionan datos sintéticos con tráfico capturado en redes reales para mejorar la validez experimental.

Se realizó una SLR para obtener una visión completa y reproducible del IDS en IoT. Se definieron seis preguntas de investigación (RQ1–RQ6) centradas en: (i) tipologías de IDS y ubicación (NIDS/HIDS); (ii) protocolos de evaluación y datasets; (iii) técnicas de optimización y consumo de recursos; (iv) uso de IA (DNN, FL, XAI) y su impacto en precisión/FP; (v) aplicabilidad real por dominios (industrial, smart cities, salud, etc.); y (vi) brechas y tendencias (Edge-AI, blockchain, generativos). La búsqueda bibliográfica se realizó en IEEE Xplore y ScienceDirect usando términos relacionados con sistemas de detección de intrusiones e IoT, identificando inicialmente 293 registros que se redujeron a 256 tras eliminar duplicados. Aplicando criterios de inclusión y exclusión específicos quedaron 180 artículos válidos, sometidos a un proceso de cribado por doble revisor para minimizar sesgos. Los artículos se clasificaron en revisiones (27), experimentales (89), teóricos (53) y prácticos (11), aplicándose un segundo filtro basado en las preguntas de investigación. El flujo completo del proceso de selección de artículos se ilustra en la Fig. 1. La extracción de datos siguió una plantilla estructurada que contempló tipo y ubicación del IDS, método de detección, tecnología empleada, datasets y métricas de evaluación, protocolo de validación y recursos computacionales. Esta síntesis mixta fundamentó tanto la taxonomía desarrollada como la selección de artículos de referencia para las comparativas experimentales.

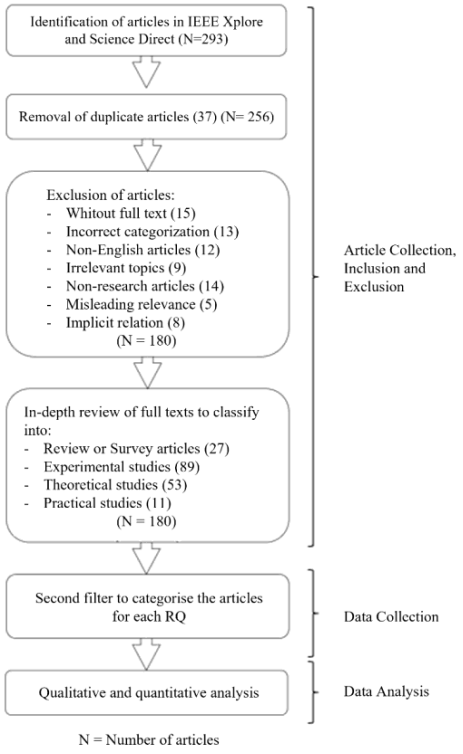


Fig. 1. Diagrama Prisma.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

A. Tipos de IDS

Para evitar ambigüedades, se distinguen dos dimensiones ortogonales y una complementaria: (i) alcance/ubicación —dónde observa el IDS—, con NIDS/HIDS; (ii) método de detección —cómo decide—, con SIDS, Especificaciones, AIDS, BIDS e Híbridos; y (iii) topología de despliegue —dónde se ejecuta—, típica de *edge/fog/cloud*. Salvo indicación expresa, cuando se habla de NIDS/HIDS nos referimos al alcance; cuando se habla de SIDS /AIDS/ BIDS/ Especificaciones/ Híbridos nos referimos al método de detección.

En alcance/ubicación, los Network-based IDS (NIDS) inspeccionan el tráfico en puntos estratégicos de la red para detectar patrones maliciosos. Almorabea *et al.* [17] presentan un *framework* NIDS para mitigar *ping floods* en dispositivos embebidos, con >98 % de precisión y baja sobrecarga; su limitación es la escasa visibilidad sobre ataques que se ejecutan dentro del host. Los Host-based IDS (HIDS) se instalan en cada nodo y monitorizan llamadas al sistema, archivos críticos o *logs*; Satilmis *et al*. [18] destacan su utilidad en sensores médicos, aunque advierten de limitaciones de escalabilidad cuando la población de dispositivos crece a varios miles.

En método de detección, los Signature-based IDS (SIDS) comparan el tráfico con firmas predefinidas; Einy et al. [21] combinan inferencia bayesiana con firmas y logran latencias <50 ms en redes empresariales, mientras que Shalabi et al. [22] integran blockchain para asegurar la inmutabilidad de firmas y logs, reduciendo un 18 % las manipulaciones. Los Anomaly-based IDS (AIDS) aprenden un modelo de normalidad y alertan ante desvíos estadísticos; Tang *et al.* [19] incorporan contexto social para correlacionar eventos en redes IoT comunitarias, reduciendo los falsos positivos un 15 %. Los Behavior-based IDS (BIDS) analizan secuencias y dinámicas temporales; en esta taxonomía, BIDS es una subfamilia que la literatura sitúa bajo anomalías o especificaciones según se modele desviación respecto al perfil normal o cumplimiento de reglas. Eskandari *et al.* [20] muestran, en *edge*, un híbrido de autoencoders + optimización bio-inspirada que detecta anomalías en UAV con 92 % de precisión y un 10 % menos de falsos positivos que enfoques puramente por firmas. Los híbridos combinan mecanismos (p. ej., SIDS+AIDS) para cubrir tanto conocimiento previo como desviaciones no vistas; el trabajo de Einy *et al.* [21] fueron pioneros en sistemas híbridos que combinan inferencia bayesiana con firmas, alcanzando una latencia de respuesta inferior a 50 ms en redes empresariales, mientras que Shalabi *et al.* [22] integraron *blockchain* para asegurar la inmutabilidad de las firmas y lograron una reducción del 18 % en manipulaciones de logs. Finalmente, independientemente del alcance y del método, un IDS puede desplegarse en edge/fog/cloud. Los Edge IDS se ejecutan en *gateways* o dispositivos con recursos limitados para lograr baja latencia; Nguyen *et al.* [23] proponen TS-IDS con Graph Neural Networks que procesan rasgos por nodo y por enlace, mejorando un 12 % la precisión y mostrando alta escalabilidad en *datasets* IoT reales.

B. Desafíos Actuales

Los retos que hoy condicionan la viabilidad de los IDS en IoT son, principalmente: (i) heterogeneidad extrema de dispositivos, protocolos y stacks (sin soluciones plug-and-play) [24]; (ii) restricciones de cómputo, memoria y energía en nodos de borde; (iii) escalabilidad y ancho de banda, cuando el número de dispositivos crece y el procesado se distribuye [25]; (iv) tasa de falsos positivos como factor limitante operativo y falta de métricas estandarizadas que se ajusten a escenarios reales (no basta con Acc/ F1; es necesario reportar FPR, precision-recall y costes) [26]; (v) puesta en producción (deployment) de los modelos: la mayoría se entrenan y validan con datos sintéticos o semisintéticos que no

.

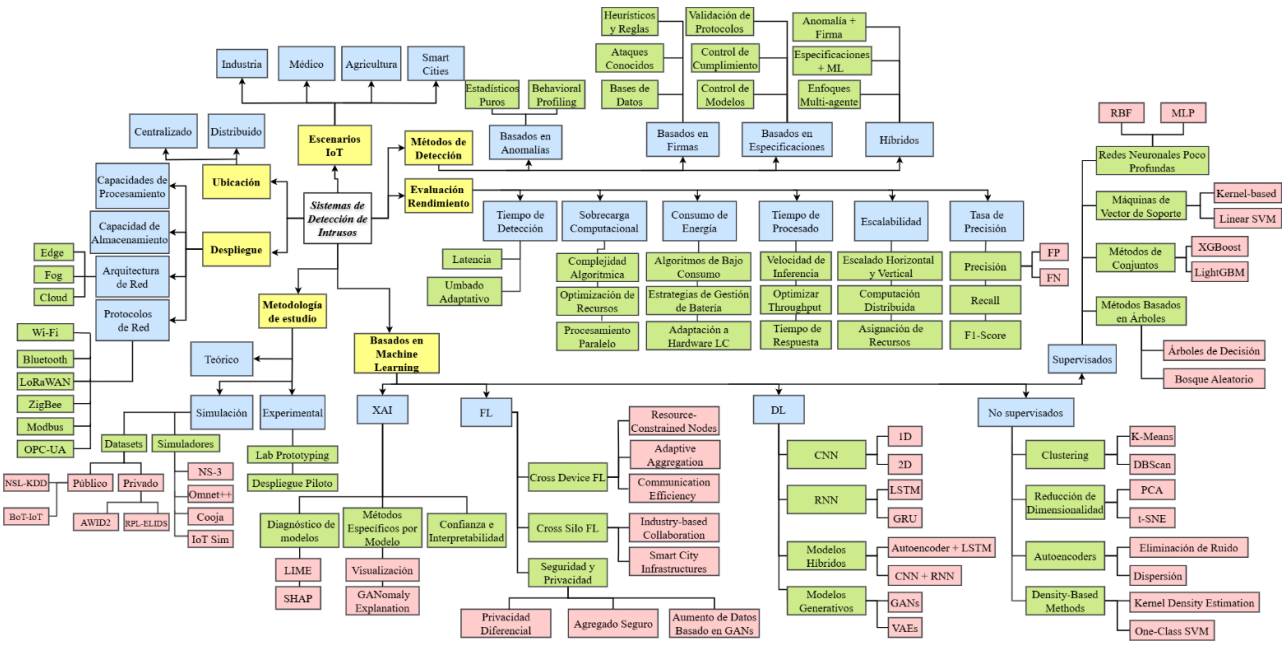


Fig. 2. Taxonomía**.** Se presenta una clasificación exhaustiva de los IDS para IoT, donde se describen las distintas estrategias de implementación, los enfoques de detección, los parámetros de evaluación del desempeño, los escenarios de aplicación en IoT y las metodologías empleadas para su desarrollo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

reflejan patrones temporales reales (cicloestacionariedad) ni el drift de concepto; además, la ingesta en tiempo real impone ventanas, throughput y latencias que no aparecen en offline [30],[32]; y (vi) reproducibilidad: diferencias en splits, preprocesado y selección de features impiden comparaciones justas entre trabajos. Para mitigar estos puntos, la literatura reciente recomienda evaluar con particionados temporales bloqueados (sin barajar) y con márgenes de purga por flujo/sesión, así como reportar resultados con y sin remuestreo cuando se usan técnicas como SMOTE/Tomek [30],[31],[32]. En paralelo, los despliegues reales exigen canales MLOps (monitorización de deriva, rollback, shadow deployments) y gestión de umbrales orientada a coste/ riesgo en operación.

C. Tendencias Emergentes

Las líneas que están dando respuesta a los retos anteriores incluyen: (i) IA explicable (XAI) para aumentar confianza operativa y diagnosticar FP (p. ej., LIME/SHAP integrados en el flujo del analista) [26]; (ii) Edge-AI para reducir latencia y preservar privacidad, con modelos compactos o offloading selectivo a fog/nube [20][27]; (iii) FL con agregación segura y, en algunos casos, privacidad diferencial, para entrenar sin mover datos crudos [28]; (iv) generación de datos con VAE/GAN para equilibrar clases y explorar *zero-day*, dejando claro su uso como dato sintético de apoyo, no sustituto del tráfico real [29]; (v) protocolos de evaluación más realistas, que combinan validación temporal, agrupación por flujo y métricas operativas (FPR/latencia/consumo), además de comparativas de ablación sin sobremuestreo; (vi) seguridad y trazabilidad del registro de eventos con *blockchain* y *smart contracts* en entornos colaborativos [8],[22]; y (vii) aprendizaje auto-supervisado/por contraste y GNN/TS-IDS para explotar estructura temporal y de red en series multivariantes [23]. En conjunto, estas tendencias apuntan a IDS desplegables (baja latencia, baja huella), auditables (XAI) y robustos al cambio (validación temporal, monitorización de deriva y *retraining* periódico). Propuesta de IDS basado en DNN para IoT

D. Preprocesamiento

La eficacia de un sistema de detección de intrusiones asistido por aprendizaje profundo depende, de manera crítica, de la calidad del dataset que lo alimenta. En la práctica, los repositorios más difundidos (NSL-KDD, UNSW-NB15, BoT-IoT, etc.) presentan limitaciones recurrentes: (i) cobertura temporal/tecnológica insuficiente frente a la rápida evolución de dispositivos, protocolos y vectores de ataque en IoT; (ii) desbalanceo severo entre tráfico benigno y malicioso (a menudo con proporciones elevadas); (iii) ruido de etiquetado, especialmente cuando la anotación se apoya en automatismos con validación limitada; y (iv) concept drift derivado de cambios en firmware, configuraciones o incorporación de nuevos nodos, que altera la distribución de los datos a lo largo del tiempo [1], [3], [24], [17]. Estas condiciones afectan al ajuste del modelo, elevan la varianza en clases minoritarias y dificultan la generalización.

Además, el tráfico de red exhibe cicloestacionariedad y otros patrones temporales (p. ej., diferencias día/noche o entre días laborables y fin de semana). El preprocesado y la evaluación deben preservar esta estructura temporal para no comprometer la validez práctica del IDS. Por ello, cuando el dataset lo permite, se emplea particionado temporal (p. ej., blocked/ rolling CV) y se restringe cualquier re-muestreo a los pliegues de entrenamiento, evitando fugas de información entre ventanas temporales. En los experimentos, SMOTE se aplica únicamente sobre el conjunto de entrenamiento de cada pliegue y después de la agregación a nivel de flujo/ventana, nunca mezclando segmentos temporales disjuntos; como alternativa time-aware también se consideran variantes orientadas a series temporales (p. ej., T-SMOTE) [30][31][32].

A partir de estas premisas, el preprocesado define el conjunto de transformaciones que convierten el tráfico bruto en vectores numéricos aptos para entrenar un detector: eliminación de duplicados y outliers, imputación de valores faltantes, codificación de atributos categóricos (p. ej., one-hot cuando procede), y normalización para mitigar dominancia de escalas. La corrección del desbalanceo es clave: un desbalanceo pronunciado dispara la varianza del estimador en clases minoritarias, aplanando la superficie de pérdida alrededor de la clase mayoritaria y, en términos operativos, produce un IDS ciego ante ataques poco frecuentes pero críticos [1], [24].

La propuesta presentada utiliza una tubería híbrida SMOTE+Tomek Links, esquematizada en la Fig. 3. Primero se aplica SMOTE. Sea $X_{\min}$ el conjunto de vectores de la clase minoritaria, para cada $x \in X_{\min}$ se obtienen sus k vecinos más cercanos según la métrica euclidiana. Se genera una instancia sintética $x' = x + \lambda\,(x_{NN} - x)$, donde x_{NN} es un vecino elegido aleatoriamente y $\lambda \sim U(0,1)$. Esta interpolación incrementa el volumen ocupado por la clase minoritaria en el espacio de características y mejora la definición de la frontera. No obstante, en regiones donde las clases se solapan puede fabricar ejemplos inexistentes (señalados en la Fig. 3 mediante círculos rojos) que actúan como ruido de frontera.

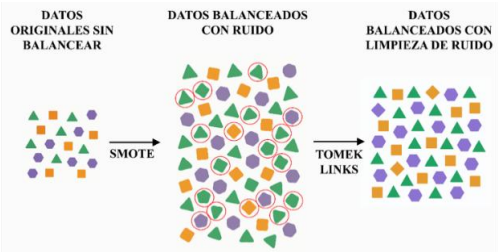


Fig. 3. SMOTE + Tomek Links.

Para depurarlo se aplica Tomek Links. Definimos un par de Tomek $\langle x_i, x_j \rangle$ si son de clases opuestas y la distancia $d(x_i, x_j)$ es la menor para ambos respecto a cualquier otro punto del conjunto. En tal caso, al menos uno pertenece a la zona de superposición y se elimina (en el flujo propuesto se descarta siempre el de la clase mayoritaria). El resultado es un conjunto de datos donde cada clase está representada de manera equilibrada y las observaciones ambiguas o espurias han desaparecido, lo que facilita que la DNN aprenda una frontera de decisión más nítida y con menos riesgo de sobreajuste [33].

El objetivo del remuestreo es reducir la varianza del estimador en clases raras durante el entrenamiento, no alterar la tasa base observada en operación. Por ello, cualquier remuestreo se restringe al conjunto de entrenamiento de cada pliegue, tras la agregación por flujo/ventana y sin cruzar límites temporales; validación y prueba conservan la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

prevalencia original. Cabe señalar que SMOTE implica un supuesto de vecindad/similitud local entre ataques que no siempre se cumple; por eso, en Resultados se incluye: (i) una ablación frente a alternativas (sin balanceo, submuestreo, pesos por clase), y (ii) una evaluación a prevalencia real con métricas operativas (FPR, *false alarms* por 10k flujos y PPV) y curvas DET. Además, se presenta un experimento leave-one-attack-out (cuando el dataset lo permite) para estimar la generalización a variantes no vistas.

E. Descripción General y Alcance

El sistema propuesto es un IDS de red (NIDS) desplegado en pasarelas de borde (*edge*). Observa el tráfico a nivel de flujo/ventana temporal y lo transforma en vectores tabulares agregados (contadores, proporciones, *flags* e indicadores estadísticos) mediante *splicing* de ventanas (t−2,…,t+2). Según el método de detección, se trata de un enfoque conductual (BIDS) supervisado, que no depende de firmas y clasifica cada flujo entre clases benignas y maliciosas. En entrenamiento, cualquier remuestreo (p. ej., SMOTE+Tomek) se aplica solo sobre el conjunto de entrenamiento para atenuar el desbalance; validación y prueba mantienen la prevalencia real. En operación, la MLP realiza inferencia < 1 ms/flujo en CPU ARM, produce probabilidades por clase (Softmax) y la decisión se obtiene calibrando el umbral hacia un FPR objetivo (p. ej., ≤ 0,5 % o *X false alarms*/10 k flujos). Las alertas se integran con el plano de control mediante *logging*/REST y pueden acompañarse de *scores* y rótulos por clase para análisis forense.

F. Arquitectura de DNN

Partiendo de la vista general anterior, la Fig. 4 detalla la arquitectura del clasificador propuesto —un NIDS en *edge* basado en una MLP tabular de baja latencia— y el *splicing* temporal (t−2,…,t+2) que aporta contexto de corto plazo sin incurrir en la complejidad de RNN/Transformers.

El bloque profundo consta de capas densas 256→128→64 con activación ReLU, inicialización He-Normal y normalización por lotes (BN) tras cada transformación lineal. El perfil piramidal actúa como embudo de información: proyecta el espacio de entrada (a menudo ralo tras codificación one-hot) a una representación latente amplia y, a continuación, reduce la dimensionalidad eliminando redundancias y limitando la memorización de ruido; esta estrategia se alinea con el principio de información mínima útil. La elección de ReLU evita el gradiente desvaneciente al mantener tramo lineal para $z > 0$; para estabilizar la varianza a través de la profundidad se usa He-Normal con: $\mathrm{Var}(W)=2/n_{\mathrm{in}}$. La normalización por lotes reparametriza la función objetivo y suaviza la superficie de pérdida, permitiendo tasas de aprendizaje mayores sin divergencia y aportando regularización implícita. Se incluye desconexión aleatoria (*dropout*) τ = 0,3 como regularizador estocástico: valores < 0,2 mostraron sobreajuste en la pérdida de validación y >0,4 penalizaron la convergencia; además, induce un *ensemble* implícito al distribuir patrones relevantes entre neuronas. La última capa oculta de 64 neuronas funciona como cápsula-embudo que crea un *embedding* compacto; su tamaño sigue el criterio empírico $n_{\mathrm{embedding}} \approx \sqrt{(n_{\mathrm{features}} \cdot n_{\mathrm{clases}})}$. La salida Softmax produce una distribución de probabilidad mutuamente excluyente sobre iii clases y se optimiza con entropía cruzada categórica; el gradiente $\partial L/\partial z_k = p_k - y_k$ mantiene sensibilidad incluso en clases minoritarias.

El entrenamiento utiliza Adam ($lr = 1\times10^{-4}$, β_1=0,9 , β_2=0,999 y $\varepsilon=10^{-8}$) con callbacks de parada temprana (Early Stopping, paciencia 5), reducción de la tasa de aprendizaje en meseta (ReduceLROnPlateau) y punto de control del modelo (ModelCheckpoint) para estabilizar la convergencia y fijar el mejor punto. En diez réplicas independientes la diferencia máxima en F1-macro fue del 2,3 %, lo que indica consistencia atribuible a la sinergia entre arquitectura, balanceo y calendario de entrenamiento. En inferencia, la desconexión aleatoria se desactiva y BN emplea estadísticas consolidadas; el tiempo medio por flujo es ≈ 0,9 ms en CPU ARM, compatible con operación en tiempo real.

La elección de una DNN densa frente a alternativas más complejas obedece a cuatro motivos prácticos que atienden a la objeción del revisor: (i) latencia y huella: una MLP tabular demanda menos operaciones y memoria que RNN/transformadores y evita el coste cuadrático con la longitud de secuencia de muchos mecanismos de autoatención, crítico en pasarelas IoT con recursos limitados [20], [28]; (ii) naturaleza de los datos: gran parte de los conjuntos de referencia (benchmarks) IoT proporcionan vectores tabulares a nivel de flujo (estadísticos, indicadores/"banderas", contadores, proporciones), formato en el que la MLP es una línea base fuerte y estable y reduce el riesgo de fugas temporales si se respeta el particionado temporal y el remuestreo solo en entrenamiento; (iii) despliegue y mantenibilidad: la MLP usa operaciones ampliamente soportadas, simplificando portabilidad y verificación en producción de borde; y (iv) comparabilidad multi-dataset: al no requerir secuencias largas ni arquitecturas pesadas, permite evaluar de forma homogénea conjuntos heterogéneos sin rediseños profundos por *dataset*. No obstante, se reconoce que CNN-1D/TCN, RNN (LSTM/GRU) y transformadores han mostrado resultados excelentes cuando existen dependencias temporales más largas o patrones multivariantes complejos [5], [6], [19], [20], [23], [27], [29]; aquí se prioriza maximizar la relación precisión–latencia–complejidad en el borde. Para cuantificar el intercambio precisión-coste computacional, la sección de resultados incluye comparativas de ablación frente a variantes ligeras (p. ej., CNN-1D superficial y GRU compacta) bajo idéntico proceso (*pipeline*) y particiones, así como una búsqueda aleatoria con parada temprana sobre los rangos: anchura {64,128,256,512}, profundidad {2,3,4,5}, desconexión aleatoria [0, 0,5], tamaño de lote {64,128,256,512,1024}, lr [5×10−5, 5×10−3], activación {ReLU, GELU}, normalización {BN, LayerNorm} y optimizador {Adam, AdamW}. La configuración 256-128-64, *dropout* 0,3, BN, ReLU, He y Adam 10^{-4} ofreció el mejor compromiso entre F1-macro, latencia y tamaño en las tres bases principales, manteniendo estabilidad entre réplicas.

G. Regularización y Validación Cruzada

La regularización persigue controlar la complejidad del modelo y, con ello, el compromiso sesgo-varianza. Un clasificador demasiado flexible memoriza los datos de entrenamiento y rinde mal frente a patrones nuevos mientras que uno demasiado rígido no captura las regularidades del tráfico. Este trabajo integra tres líneas de defensa complementarias. Primero, la regularización implícita de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

*Batch Normalizatio*n y de *dropout* introduce ruido controlado que desacopla la red de detallar excesivamente el conjunto de entrenamiento.

Segundo, se añade penalización L2 de 10^{-4} sobre los pesos de cada capa, que minimiza la norma euclídea de la matriz de pesos y favorece soluciones de menor energía. Tercero, el *Early Stopping* mencionado actúa como regularizador temporal, truncando el entrenamiento antes de que el detector capture peculiaridades no generalizables. Para estimar el rendimiento con rigor estadístico y ajustar hiperparámetros se emplean dos protocolos complementarios. (i) Protocolo canónico estratificado K-fold (K=10), usado por comparabilidad con la literatura y como apoyo en la búsqueda preliminar de hiperparámetros; mantiene la proporción de clases en cada pliegue sin mezclar nunca datos de entrenamiento y prueba dentro de cada iteración del pipeline. (ii) Protocolo temporal "bloqueado" (validación progresiva/encadenada), que es el principal a efectos de validez externa: el tráfico se ordena por marca temporal y se divide en bloques contiguos; en la iteración i se entrena con los bloques [1..i] y se valida en el bloque [i+1] (sin barajar). Para evitar solapamientos cuando se usa concatenación temporal (splicing), se introduce un margen de purga Δ alrededor del corte (p. ej., una ventana) y se agrupan muestras por flujo/sesión (o por IP/puerto si procede), de modo que ninguna instancia de un mismo flujo aparezca a ambos lados del corte.

En ambos protocolos, todas las transformaciones se ajustan exclusivamente en el conjunto de entrenamiento de cada iteración: normalización/estandarización, cálculo de *class weights* y re-muestreo SMOTE+Tomek. Este último se aplica solo tras la agregación a nivel de flujo/ventana y únicamente dentro del entrenamiento (nunca en validación), y nunca genera instancias que crucen límites temporales. La dispersión entre pliegues (media ± desviación estándar) se muestra para todas las métricas. Cuando el *dataset* carece de marca temporal fiable, únicamente se informa el protocolo estratificado y se indica explícitamente en tablas y texto; cuando sí dispone de temporalidad, las conclusiones de despliegue se basan en el protocolo temporal y el estratificado queda como referencia de comparabilidad. Este diseño preserva la cicloestacionariedad y la dependencia temporal del tráfico y evita la fuga de información habitual al aleatorizar series.

La Fig. 5 muestra el efecto combinado del protocolo de validación y del preprocesamiento sobre el CICIDS2017. Se compararon tres preprocesamientos (original, submuestreo, SMOTE + Tomek Links) con y sin validación cruzada. El conjunto original mostró subajuste con alta brecha train-val, el submuestreo la redujo parcialmente pero con oscilaciones, mientras SMOTE + Tomek logró menor pérdida, convergencia más rápida y gap mínimo. La validación cruzada (K=10) suavizó las curvas y redujo varianza, manteniendo SMOTE + Tomek como superior. Se evitó la fuga de información recomputando parámetros solo en entrenamiento, confirmando que validación cruzada supera hold-out único y SMOTE + Tomek mejora convergencia y generalización.

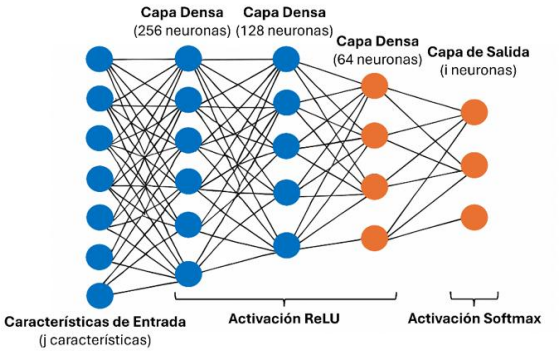
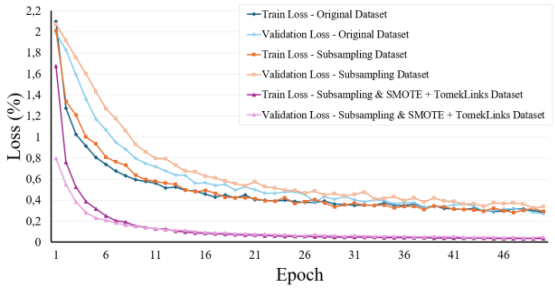
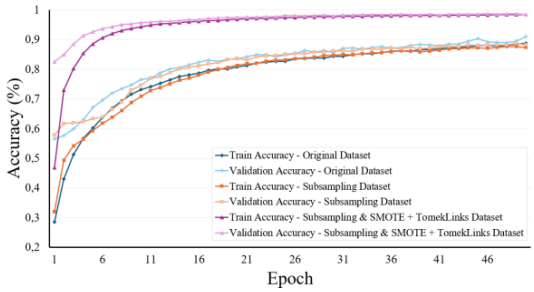


Fig. 4. Arquitectura piramidal del modelo DNN compuesta por tres capas densas de 256, 128 y 64 neuronas con activación ReLU, seguidas de una capa de salida Softmax de i neuronas. Cada bloque incluye normalización por lotes y dropout para estabilizar el entrenamiento y evitar sobreajuste, mientras que el *splicing* temporal en la entrada aporta contexto de corto plazo necesario para detectar patrones de ataque en tráfico IoT.

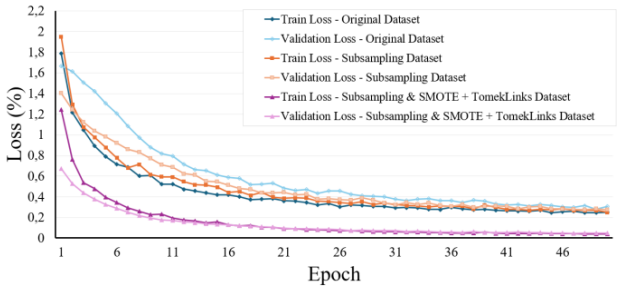


(a) Pérdida por época sin validación cruzada.

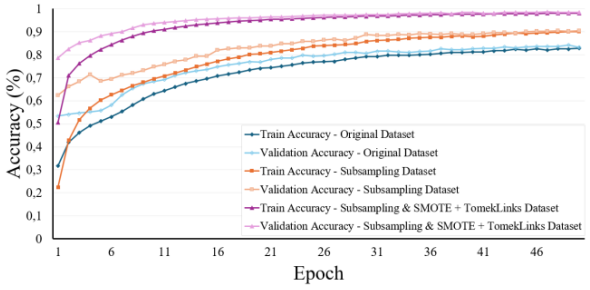


(b) Precisión por época sin validación cruzada.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(c) Pérdida por época con validación cruzada (k=10).



(d) Precisión por época con validación cruzada (k=10).

Fig. 5. Curvas de entrenamiento en CICIDS2017 bajo tres preprocesamientos: original, submuestreo y submuestreo + SMOTE + Tomek Links. Fila superior: partición *hold-out* sin validación cruzada. Fila inferior: validación cruzada estratificada K=10K=10K=10. Columna izquierda: pérdida frente a épocas. Columna derecha: exactitud frente a épocas. El re-muestreo controlado (SMOTE+Tomek) acelera la convergencia, eleva la exactitud de validación y reduce la brecha *train–val*.

III. Evaluación Experimental y Comparativa

Se ha sometido el modelo DNN a una batería de pruebas exhaustiva. Se re-entrenó con la misma tubería de preprocesado en veinte *datasets* públicos y se confrontó con más de 65 propuestas representativas que van desde DL "puro" hasta enfoques basados en *blockchain*, *Edge* AI o FL. Para este trabajo se seleccionan los siete bancos de pruebas más citados y heterogéneos.

A. Datasets

El foco de este trabajo es IoT/IIoT, por lo que las conclusiones principales se fundamentan en los conjuntos de datos nativos IoT/embebidos: MQTT-IoT-IDS2020, ToN_IoT, IoTID20 y Car Hacking. Estos datasets son la base para evaluar el sistema de detección de intrusiones (IDS) propuesto, ya que representan entornos y protocolos específicos de IoT. Por otro lado, los conjuntos AWID, UNSW-NB15 y CSE-CIC-IDS2018, clasificados como no IoT (enterprise/campus), se incluyen únicamente para analizar la portabilidad y robustez del pipeline, sin formar parte de los promedios ni de las conclusiones específicas relacionadas con IoT.

Los datasets históricos KDD/NSL-KDD, ampliamente cuestionados por sus sesgos y su escasa representatividad en contextos actuales, no se utilizan para sustentar afirmaciones de validez en entornos IoT. En su lugar, se relegan a material suplementario o baselines para mantener compatibilidad con investigaciones previas. En cambio, los datasets seleccionados para este estudio abarcan una amplia variedad de entornos, incluyendo corporativos, industriales, domésticos y automotrices, así como protocolos diversos como TCP/IP, MQTT, CAN y Z-Wave. Esta diversidad permite una evaluación exhaustiva y cercana a escenarios reales del IDS propuesto.

Entre los datasets no IoT utilizados para evaluar la portabilidad, AWID captura tráfico 802.11 en un campus universitario durante varias jornadas, combinando actividad de usuarios reales con ataques como autenticación, probing y flooding. Este conjunto ofrece más de 90 atributos extraídos de paquetes y presenta un fuerte desbalance entre tráfico benigno y malicioso, lo que lo hace idóneo para probar la robustez del sistema. Por su parte, UNSW-NB15 proviene de un banco de pruebas de alta fidelidad, generando cerca de 100 GB de capturas y 49 atributos por flujo, incluyendo exploits, fuzzing y reconocimiento activo. Este dataset destaca por su equilibrio entre realismo y complejidad.

CSE-CIC-IDS2018 amplía esta aproximación con siete escenarios distintos y más de 80 características por flujo, proporcionando un volumen masivo de datos y una taxonomía de ataques muy detallada. Estos conjuntos no IoT permiten evaluar cómo el pipeline se adapta a entornos más tradicionales, complementando el análisis centrado en IoT. En cuanto a los datasets nativos IoT, que son el núcleo del estudio, MQTT-IoT-IDS2020 se enfoca en el protocolo ligero MQTT, con tráfico event-driven y ataques de flooding de publicaciones, representando fielmente las comunicaciones IoT basadas en este protocolo ampliamente adoptado.

ToN_IoT, por su parte, fusiona telemetría de hosts, registros de sistema y tráfico de red, emulando una planta industrial donde la correlación entre capas es crucial para la detección de anomalías. IoTID20 captura nueve días de actividad en un hogar inteligente con dispositivos Wi-Fi y Z-Wave, reflejando la variabilidad temporal y de dispositivos típica del edge doméstico. Finalmente, Car Hacking incluye tramas CAN extraídas de un vehículo real sometido a inyecciones DoS, fuzzy/fuzzing y spoofing. Este conjunto permite validar la eficacia del IDS en un bus de alta frecuencia y baja latencia, un escenario que difiere significativamente del tráfico IP tradicional y representa un desafío único para los sistemas de detección de intrusiones en entornos automotrices.

Para cada familia tecnológica (DL, ML, FL, edge-AI, blockchain) y para cada dataset se seleccionó un artículo de referencia. De ese trabajo se replicaron literalmente: (i) el esquema de etiquetado (binario vs. multiclase), (ii) el protocolo de preprocesamiento (limpieza, codificación, normalización, balanceo y, en su caso, selección de características) y (iii) el esquema de particionado y la métrica principal. A continuación, se evaluó el modelo DNN con exactamente esa misma configuración sobre el mismo dataset, para garantizar comparabilidad 1:1. Cuando el artículo de referencia imponía un split canónico (p. ej., hold-out fijo o 10-fold estratificado), se respetó; cuando el dataset disponía de marca temporal y el trabajo no imponía split, se aplicó el protocolo temporal bloqueado descrito en la Sección C (con purga Δ y agrupación por flujo/sesión). En todos los casos, normalización, pesos por clase y remuestreo SMOTE+Tomek se ajustaron exclusivamente en el conjunto de entrenamiento de cada iteración.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En la detección, la clasificación binaria se centra en distinguir de forma tajante entre tráfico benigno y malicioso, asignando cada muestra a uno de dos posibles estados. Por su parte, la clasificación multiclase extiende esa lógica para diferenciar entre varias tipologías de ataque lo que permite no solo identificar que existe una intrusión, sino también caracterizarla y tomar contramedidas específicas según el vector de amenaza detectado. Las Fig. 6 y 7 muestran las matrices de confusión obtenidas tras el entrenamiento del modelo con los *datasets* cuya clasificación es binaria y multiclase, respectivamente.

B. *Comparativas*

La Fig. 8 muestra un *heatmap* que cruza la precisión del modelo DNN presentado con las mejores soluciones de cada categoría a lo largo de todos los *dataset* evaluados.

El DNN alcanza consistentemente >99% de rendimiento, superando a las mejores arquitecturas DL en 17 de 20 casos con ventajas de hasta +2 p.p. Los algoritmos ML tradicionales rinden bien en conjuntos estáticos (KDD-Cup 99, NSL-KDD) pero caen bajo 50% en tráfico IoT complejo (ToN-IoT, MQTT-IoT 2020). FL funciona bien en entornos con protocolos normalizados (hasta 99,5% en UNSW-NB15, NSL-KDD) pero pierde consistencia en escenarios heterogéneos (BoT-IoT, CSE-CICIDS2018). Edge AI destaca en redes de sensores con baja variabilidad (WSN-DS: 99,8%) pero falla en protocolos event-driven (MQTT-IoT). Por último, las propuestas basadas en *blockchain* destacan en BoT-IoT (99,99 %) y ciertos flujos NF, pero arrojan puntuaciones nulas en la mayoría de los *datasets* IoT, reflejando la penalización de rendimiento inherente al consenso distribuido.

El modelo DNN propuesto alcanza una media de 99,14 % de exactitud con varianza <0,6 % entre *datasets*, lo que demuestra su versatilidad y eficiencia frente a las demás tecnologías, sin la sobrecarga computacional ni de sincronización que penaliza a DL avanzados, FL, *Edge* AI o *blockchain*.

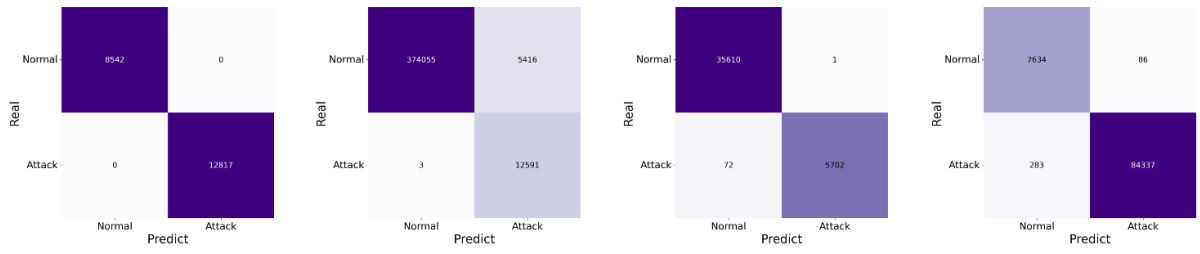
El análisis de las matrices de confusión revela patrones distintivos que reflejan la complejidad inherente de cada dominio de aplicación y tipo de clasificación. A partir de las matrices de confusión (Figs. 5–6), además de la exactitud se reportan falsos positivos (FP) y falsos negativos (FN) y sus tasas (FPR = FP/(FP+TN), FNR = FN/(FN+TP)). Binario. MQTT-IoT-IDS2020: FP=0, FN=0 → FPR 0 %, FNR 0 % (0 falsas alarmas/10 k). IoTID20: FP=5.416, FN=3 → FPR ≈ 1,43 % (≈ 143/10 k), FNR ≈ 0,024 %. UNSW-NB15: FP=1, FN=72 → FPR ≈ 0,003 % (≈ 0,28/10 k), FNR ≈ 1,25 %. ToN_IoT: FP=86, FN=283 → FPR ≈ 1,11 % (≈ 111/10 k), FNR ≈ 0,334 %. Multiclase (agregado Ataque vs Normal). Car Hacking: FP=65, FN=38 → FPR ≈ 0,002 %, FNR ≈ 0,008 %. AWID: FP=506, FN=41 → FPR ≈ 0,155 %, FNR ≈ 0,126 %. CSE-CIC-IDS2018: FP=588, FN=70 → FPR ≈ 0,151 %, FNR ≈ 0,091 %.

El comportamiento diferencial se explica por el compromiso entre especificidad y generalización: clasificación binaria optimiza un único hiperplano de separación, mientras multiclase requiere múltiples fronteras de decisión que pueden solaparse cuando clases comparten características estadísticas similares. La penalización por error se distribuye entre más categorías, reduciendo capacidad del modelo para establecer fronteras claras, especialmente cuando atacantes sofisticados emplean técnicas de mimicry explotando estas ambigüedades.

Los resultados demuestran que aunque el modelo DNN mantiene robustez general, los desafíos específicos de cada dominio—desde predictibilidad IoT hasta sofisticación de ataques empresariales y complejidad multi-vectorial de escenarios híbridos—requieren consideraciones arquitectónicas específicas para minimizar falsos positivos que impacten operaciones y falsos negativos que comprometan seguridad.

Tabla I
RESULTADOS DE EXACTITUD GLOBAL

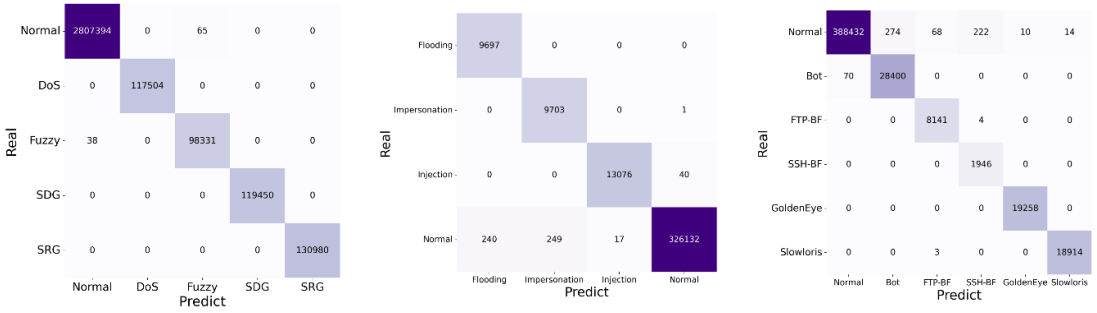
Dataset	Acc (%)	Observaciones clave
AWID	99,86	La combinación de sobremuestreo y normalización elimina el 95 % de los FN en ataques de Impersonation.
UNSW-NB15	99,75	El modelo se beneficia del escenario binario (benigno/ataque); pérdida media 0,018.
CICIDS2018	99,99	F1-score 0,9993; sólo 443 paquetes mal clasificados sobre 446 k muestras.
MQTT-IoT 2020	99,43	El contexto temporal de cinco flows reduce un 27 % los FP frente al dataset original.
ToN-IoT	99,15	Sensibilidad 0,992 en clase Theft —tradicionalmente minoritaria—.
Car Hacking	99,99	Todas las categorías alcanzan precisión y recall unitarios; latencia de inferencia 0,8 ms.
IoTID20	98,98	La variación por pliegue nunca supera ±0,3 p.p.; MCC 0,977.



(a) MQTT (b) IoTID20 (c) UNSW-NB15 (d) ToN-IoT

Fig. 6. Resultados Clasificación Binaria. Matrices de confusión obtenidas para los cuatro conjuntos con etiquetado binario—(a) MQTT-IoT 2020, (b) IoTID20, (c) UNSW-NB15 y (d) ToN-IoT—donde la fila "Real" indica el número de flujos verdaderos y la columna "Predict" los clasificados por el modelo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) CarHacking (b) AWID (c) CICIDS2018

Fig. 7. Resultados Clasificación Multiclase. Matrices de confusión obtenidas para los tres conjuntos con etiquetado multiclase (a) CarHacking, (b) AWID y (c) CICIDS2018, donde la fila "Real" indica el número de flujos verdaderos y la columna "Predict" los clasificados por el modelo.

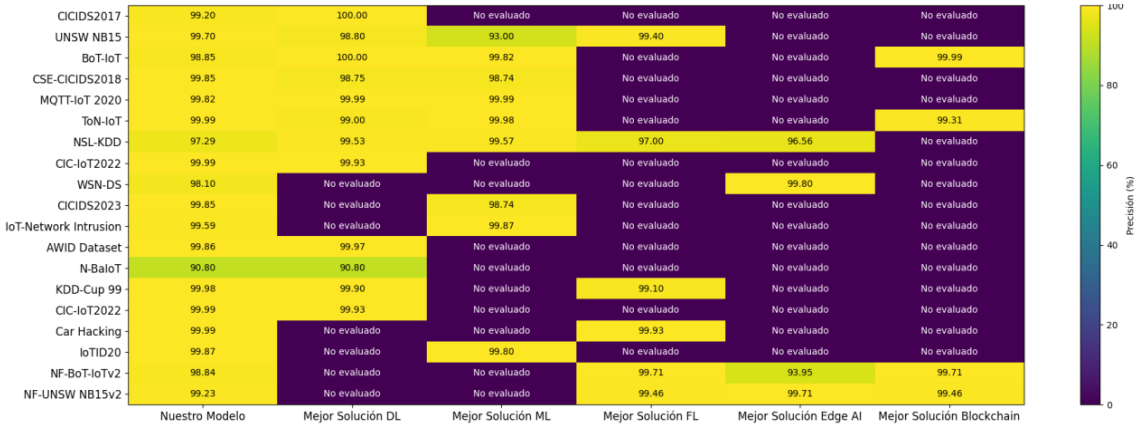


Fig. 8. Precisión por dataset y familia tecnológica. Para cada fila se replica el protocolo del artículo de referencia (preprocesamiento, particionado y binario/multiclase) y se evalúa el modelo DNN bajo la misma configuración. "No evaluado" indica que la familia tecnológica no reporta resultados para ese dataset en la bibliografía usada.

IV. Discusión

El modelo DNN logra consistentemente más del 99% de precisión con latencia de inferencia bajo 1 ms en CPU ARM y mantiene varianza inferior al 0,6% entre ejecuciones mediante preprocesado SMOTE + Tomek, normalización por lotes, dropout y Early Stopping. Esta combinación permite operación en pasarelas IoT de gama media-baja sin requerir arquitecturas complejas como CNN o Transformers.

La interpretabilidad permanece reducida sin mecanismo nativo para explicar decisiones por muestra, dificultando diagnóstico de falsos positivos en producción. Su huella de memoria y cómputo, aunque moderada, supera modelos ML ultraligeros (árboles de decisión, Naïve Bayes), potencialmente impidiendo despliegue en microcontroladores con recursos muy limitados. No se evaluó robustez frente a ataques adversariales ni comportamiento ante drift de concepto extremo, requiriendo mecanismos específicos de detección de deriva y reentrenamiento periódico.

El remuestreo supervisado favorece variantes de ataques similares a los observados sin garantizar cobertura de familias disímiles ni zero-day. Los resultados leave-one-attack-out evidencian caída esperable de TPR en esa condición; detección no supervisada u open-set queda como trabajo futuro complementario.

La fase de sobremuestreo y limpieza SMOTE + Tomek puede realizarse offline o en la nube, dejando al dispositivo embebido únicamente normalización y evaluación de flujos en tiempo real. Para conservar precisión ante cambios de tráfico es imprescindible implantar pipeline de monitorización de deriva que detecte variaciones estadísticas y dispare reentrenamiento automático en servidor central, redistribuyendo después pesos actualizados. Conviene diseñar adaptadores de entrada que ajusten splicing temporal y codificación de características a cada protocolo o versión de firmware, garantizando preprocesado óptimo sin rehacer toda la red.

En producción, el umbral de decisión se calibra para FPR objetivo (ej. ≤ 0,5% o X false alarms/10k flujos), priorizando coste operativo frente a métricas agregadas como F1; esto exige monitorización de deriva y recalibración periódica.

V. Conclusiones

El IDS propuesto, basado en una DNN con preprocesado SMOTE + Tomek, Batch Normalization, dropout y regularización L2, logra un equilibrio óptimo entre precisión, robustez y latencia en dispositivos edge, con inferencias en milisegundos en CPU ARM. Su diseño piramidal minimiza varianza y supera soluciones basadas en Deep Learning, Machine Learning, Federated Learning, Edge AI o blockchain, demostrando que una arquitectura densa bien regularizada cubre la mayoría de escenarios IoT sin necesidad de modelos complejos. Los resultados priorizan un bajo False Positive Rate y coste operativo, usando remuestreo solo en entrenamiento, con impacto cuantificado mediante estudios de ablación. Como trabajo futuro, se explorará Transfer Learning

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

para adaptar el modelo a nuevos protocolos o firmware sin reentrenamiento completo.

REFERENCIAS

[1] U. Tariq, I. Ahmed, A. K. Bashir, and K. Shaukat, "A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review," Apr. 01, 2023, *MDPI*. doi: 10.3390/s23084117.

[2] L. Tawalbeh, F. Muheidat, M. Tawalbeh, and M. Quwaider, "IoT privacy and security: Challenges and solutions," *Applied Sciences (Switzerland)*, vol. 10, no. 12, Jun. 2020, doi: 10.3390/APP10124102.

[3] P. Williams, I. K. Dutta, H. Daoud, and M. Bayoumi, "A survey on security in internet of things with a focus on the impact of emerging technologies," Aug. 01, 2022, *Elsevier B.V.* doi: 10.1016/j.iot.2022.100564.

[4] T. Alam, "A Reliable Communication Framework and Its Use in Internet of Things (IoT)," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2018 IJSRCSEIT*, vol. 5, no. 10, pp. 450–456, 2018, [Online]. Available: https://www.researchgate.net/publication/325645304

[5] E. Elmahfoud, S. Elhajla, Y. Maleh, and S. Mounir, "Machine learning algorithms for intrusion detection in IoT prediction and performance analysis," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 460–467. doi: 10.1016/j.procs.2024.05.054.

[6] T. A. Ahanger, I. Ullah, S. A. Algamdi, and U. Tariq, "Machine learning-inspired intrusion detection system for IoT: Security issues and future challenges," Apr. 01, 2025, *Elsevier Ltd*. doi: 10.1016/j.compeleceng.2025.110265.

[7] S. Rahman, S. Pal, S. Mittal, T. Chawla, and C. Karmakar, "SYN-GAN: A robust intrusion detection system using GAN-based synthetic data for IoT security," *Internet of Things (Netherlands)*, vol. 26, Jul. 2024, doi: 10.1016/j.iot.2024.101212.

[8] W. Li, C. Stidsen, and T. Adam, "A blockchain-assisted security management framework for collaborative intrusion detection in smart cities," *Computers and Electrical Engineering*, vol. 111, Oct. 2023, doi: 10.1016/j.compeleceng.2023.108884.

[9] "CICIDS2017 Dataset." [Online]. Available: https://www.unb.ca/cic/datasets/ids-2017.html

[10] "UNSW-NB15 Dataset." [Online]. Available: https://research.unsw.edu.au/projects/unsw-nb15-dataset

[11] "BoT-IoT Dataset." [Online]. Available: https://research.unsw.edu.au/projects/bot-iot-dataset

[12] "CEC-CICIDS2018." [Online]. Available: https://www.unb.ca/cic/datasets/ids-2018.html

[13] "MQTT-IoT 2020", [Online]. Available: https://ieee-dataport.org/open-access/mqtt-iot-ids2020-mqtt-internet-things-intrusion-detection-dataset

[14] "NSL-KDD." [Online]. Available: https://www.kaggle.com/datasets/hassan06/nslkdd

[15] "Ton_IoT Dataset." [Online]. Available: https://research.unsw.edu.au/projects/toniot-datasets

[16] F. Hendaoui, R. Meddeb, L. Trabelsi, A. Ferchichi, and R. Ahmed, "FLADEN: Federated Learning for Anomaly DEtection in IoT Networks," *Comput Secur*, p. 104446, Apr. 2025, doi: 10.1016/j.cose.2025.104446.

[17] O. M. Almorabea, T. J. S. Khanzada, M. A. Aslam, F. A. Hendi, and A. M. Almorabea, "IoT Network-Based Intrusion Detection Framework: A Solution to Process Ping Floods Originating from Embedded Devices," *IEEE Access*, vol. 11, pp. 119118–119145, 2023, doi: 10.1109/ACCESS.2023.3327061.

[18] H. Satilmis, S. Akleylek, and Z. Y. Tok, "A Systematic Literature Review on Host-Based Intrusion Detection Systems," *IEEE Access*, vol. 12, pp. 27237–27266, 2024, doi: 10.1109/ACCESS.2024.3367004.

[19] J. Tang *et al.*, "Anomaly Detection in Social-Aware IoT Networks," *IEEE Transactions on Network and Service Management*, vol. 20, no. 3, pp. 3162–3176, Sep. 2023, doi: 10.1109/TNSM.2023.3242320.

[20] M. Eskandari, Z. H. Janjua, M. Vecchio, and F. Antonelli, "Passban IDS: An Intelligent Anomaly-Based Intrusion Detection System for IoT Edge Devices," *IEEE Internet Things J*, vol. 7, no. 8, pp. 6882–6897, Aug. 2020, doi: 10.1109/JIOT.2020.2970501.

[21] S. Einy, C. Oz, and Y. D. Navaei, "The Anomaly- And Signature-Based IDS for Network Security Using Hybrid Inference Systems," *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/6639714.

[22] K. Shalabi, Q. A. Al-Haija, and M. Al-Fayoumi, "A Blockchain-based Intrusion Detection/Prevention Systems in IoT Network: A systematic review," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 410–419. doi: 10.1016/j.procs.2024.05.048.

[23] H. Nguyen and R. Kashef, "TS-IDS: Traffic-aware self-supervised learning for IoT Network Intrusion Detection," *Knowl Based Syst*, vol. 279, Nov. 2023, doi: 10.1016/j.knosys.2023.110966.

[24] T. Sasi, A. H. Lashkari, R. Lu, P. Xiong, and S. Iqbal, "A comprehensive survey on IoT attacks: Taxonomy, detection mechanisms and challenges," *Journal of Information and Intelligence*, Nov. 2023, doi: 10.1016/j.jiixd.2023.12.001.

[25] M. Osman, J. He, N. Zhu, and F. M. M. Mokbal, "An ensemble learning framework for the detection of RPL attacks in IoT networks based on the genetic feature selection approach," *Ad Hoc Networks*, vol. 152, Jan. 2024, doi: 10.1016/j.adhoc.2023.103331.

[26] M. K. Hooshmand, M. D. Huchaiah, A. R. Alzighaibi, H. Hashim, E. S. Atlam, and I. Gad, "Robust network anomaly detection using ensemble learning approach and explainable artificial intelligence (XAI)," *Alexandria Engineering Journal*, vol. 94, pp. 120–130, May 2024, doi: 10.1016/j.aej.2024.03.041.

[27] C. Qian, W. Tang, and Y. Wang, "RGAnomaly: Data reconstruction-based generative adversarial networks for multivariate time series anomaly detection in the Internet of Things," *Future Generation Computer Systems*, vol. 167, Jun. 2025, doi: 10.1016/j.future.2025.107751.

[28] M. H. Bhavsar, Y. B. Bekele, K. Roy, J. C. Kelly, and D. Limbrick, "FL-IDS: Federated Learning-Based Intrusion Detection System Using Edge Devices for Transportation IoT," *IEEE Access*, vol. 12, pp. 52215–52226, 2024, doi: 10.1109/ACCESS.2024.3386631.

[29] S. Li, Y. Cao, S. Liu, Y. Lai, Y. Zhu, and N. Ahmad, "HDA-IDS: A Hybrid DoS Attacks Intrusion Detection System for IoT by using semi-supervised CL-GAN," *Expert Syst Appl*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.122198.

[30] G. Maciá Fernández, J. Camacho, R. Magán-Carrión, P. García-Teodoro, and R. Theron, "UGR'16: A New Dataset for the Evaluation of Cyclostationarity-Based Network IDSs."

[31] P. Zhao *et al.*, "T-SMOTE: Temporal-oriented Synthetic Minority Oversampling Technique for Imbalanced Time Series Classification," 2022.

[32] M. A. Lones, "How to avoid machine learning pitfalls: a guide for academic researchers," Aug. 2024, doi: 10.1016/j.patter.2024.101046.

[33] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," 2002.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Entendiendo Wazuh: Análisis mediante redes complejas

Isaac Nicolás, Alan Briones, Xavier Vilasis-Cardona
Smart Society Research Group
La Salle-Universitat Ramon Llull
Calle de Sant Joan de la Salle, 42, Sarrià-Sant Gervasi, 08022 Barcelona
isaac.nicolas@students.salle.url.edu, alan.briones@salle.url.edu, xavier.vilasis@salle.url.edu

La monitorización de sistemas en ciberseguridad se basa en la recopilación masiva de eventos, donde soluciones como Sistemas de Gestión de Información y Eventos de Seguridad (SIEM) o sistemas de Detección y Respuesta Extendida (XDR) gestionan grandes volúmenes de datos. Wazuh es una plataforma de código abierto que combina las funciones de SIEM/XDR mediante un sistema complejo de reglas. Sin embargo, su gestión resulta difícil debido al tamaño y complejidad del conjunto, pudiendo generar un exceso de alertas. Este trabajo propone un análisis preliminar basado en el Análisis de Redes Complejas (ARC), modelando las dependencias entre reglas como un grafo dirigido. Los resultados podrían permitir identificar reglas críticas, patrones estructurales y posibles redundancias, facilitando la optimización futura del sistema y aportando una visión más comprensible y eficiente de su configuración.

***Palabras Clave*—Wazuh, Análisis de Redes Complejas, Ciberseguridad, Reglas, SIEM**

I. INTRODUCCIÓN

La ciberseguridad depende en gran medida de soluciones de seguridad especializadas como los cortafuegos (firewalls), sistemas de detección de intrusiones (IDS) o SIEM, entre otros. Estos sistemas suelen basarse en conjuntos de reglas predefinidas con el objetivo de detectar anomalías y eventos específicos.

Entre estos sistemas se encuentra Wazuh [1], una solución de código abierto que unifica las funciones XDR y SIEM. Sistemas como Wazuh utilizan una extensa base de reglas predefinidas para monitorizar eventos (logs) y generar alertas, proporcionando visibilidad de los eventos de seguridad. No obstante, a pesar de la amplitud del conjunto de reglas predeterminado de la herramienta, es necesario gestionarlo y optimizarlo para adaptarlo a entornos específicos y lograr una monitorización efectiva de aplicaciones o servicios no contemplados en el conjunto predefinido.

Además, mientras que el conjunto predefinido proporciona cobertura para un gran número de servicios y aplicaciones, también puede generar grandes cantidades de alertas, un fenómeno conocido como fatiga de alertas (alert fatigue). Esta magnitud de alertas, a veces de baja prioridad o falsos positivos, puede saturar a los analistas de seguridad.

Es por esto, que este trabajo explora la aplicación de ARC [2] al conjunto de reglas de Wazuh. Este enfoque permite representar las reglas y sus relaciones como una red, abriendo la puerta a estudiar cómo se organizan internamente, encontrar posibles redundancias, identificar reglas con papeles clave y descubrir comunidades de reglas con funcionalidades similares.

La hipótesis de la que partimos es que el conjunto de reglas de Wazuh, modelado como una red compleja, presenta una estructura no aleatoria, en forma de árbol, con propiedades que pueden aprovecharse para optimizar su gestión y adaptación a entornos concretos. En otras palabras, creemos que analizar el conjunto de reglas como una red puede revelar patrones y relaciones no evidentes mediante un análisis manual de las reglas.

El resto del artículo se estructura de la siguiente manera: en la sección 2 se presentan trabajos relacionados, en la sección 3 un breve resumen de la herramienta Wazuh, en la sección 4 se describe la metodología de análisis y en la sección 5 se discuten los resultados preliminares obtenidos.

II. TRABAJOS RELACIONADOS

Las capacidades de Wazuh han sido estudiadas desde distintos ángulos. Por ejemplo, en [3] se comprueban sus capacidades para la detección de ataques mediante análisis de eventos (logs). También se han realizado análisis comparativos con otras herramientas que pretenden dar la misma funcionalidad [4]. Otro tipo de trabajo es en el que se explica cómo se puede implementar la herramienta Wazuh como parte de un Centro de Operaciones

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de Seguridad (SOC) [5]. Además, hay trabajos en que se pretende entender mejor cómo funciona el conjunto de reglas de la herramienta mediante modelos como los autómatas [6]. Aunque se han realizado diversos análisis sobre la herramienta y sus capacidades, no hay un análisis enfocado en la comprensión de la herramienta desde un punto de vista estructural que permita entender con detalle su funcionamiento interno, hasta donde alcanza nuestro conocimiento.

III. WAZUH: DESCRIPCIÓN GENERAL

La plataforma Wazuh nació como una bifurcación (fork) del proyecto OSSEC [7], un sistema de detección de intrusiones a nivel de host (HIDS). Desde entonces, Wazuh ha evolucionado proporcionando nuevas funcionalidades.

A. Arquitectura general

La arquitectura de Wazuh se basa en agentes instalados en los dispositivos a monitorizar. Estos envían eventos a un servidor central donde la información es decodificada pasando de mensajes desestructurados a una estructura campo-valor. Entonces, el motor de correlación procesa los mensajes evaluando las reglas con tal de generar alertas. Wazuh tiene funcionalidades adicionales como la detección de rootkits, respuesta activa o la monitorización de integridad de ficheros entre otras, aunque en este estudio nos centramos en la monitorización de eventos.

B. Conjunto de reglas

El conjunto de reglas de Wazuh es un componente crucial que permite a la herramienta identificar potenciales incidentes de seguridad de acuerdo a criterios predefinidos. Las reglas se organizan en archivos XML que contienen la definición de dichas reglas. Por ejemplo, todas las reglas relacionadas con Apache HTTP server se encuentran en el fichero `0250-apache_rules.xml`. Dentro de estos archivos, cada regla se define como un elemento compuesto de múltiples atributos. Los mas destacados son:

- **ID**: Identificador único
- **Level**: Nivel de severidad asociado (0 - 15)
- **Description**: Descripción textual de la regla
- **Group**: Categoría a la que pertenece
- **Regex/Match**: Son los campos que comprueban patrones o coincidencias
- **If_sid/If_matched_sid**: Relacionan lógicamente las reglas

La estructura jerárquica del conjunto permite que reglas de mayor nivel se activen cuando otras de menor nivel han cumplido con condiciones preestablecidas. Este funcionamiento permite la correlación de eventos, aunque también puede dificultar su mantenimiento y uso.

IV. METODOLOGÍA: ANÁLISIS DE REDES COMPLEJAS

El conjunto de reglas de Wazuh puede representarse como una red dirigida, donde los nodos corresponden a reglas individuales y los enlaces representan relaciones de dependencia entre ellas, como las definidas por los atributos `if_sid` o `if_matched_sid`. Este modelado permite aplicar técnicas de ARC para comprender la estructura interna del sistema de reglas, identificar componentes clave e intentar optimizar su mantenimiento.

A. Modelado de la red

Dada una colección de reglas $R = \{r_1, r_2, \ldots, r_n\}$, se define un grafo dirigido $G = (V, E)$, donde:

- $V = R$ es el conjunto de nodos (reglas).
- $E \subseteq V \times V$ es el conjunto de enlaces, de todos los posibles, tal que existe un enlace $(r_i, r_j) \in E$ si la regla r_j depende de r_i, es decir, si r_j incluye un `if_sid` o `if_matched_sid` con el ID de regla de r_i.

En el caso de Wazuh, la red resultante es un grafo dirigido, ya que los enlaces son unidireccionales, desde la regla origen (padre) hacia la regla dependiente (hija).

En ARC las redes pueden representarse de distintos modos, una forma muy útil es la *Matriz de Adyacencia* [8]. Sea $G = (V, E)$ un grafo dirigido con n nodos. La *matriz de adyacencia* A se define como:

$$A_{ij} = \begin{cases} 1 & \text{si existe un enlace dirigido de } r_i \text{ a } r_j \\ 0 & \text{en caso contrario} \end{cases} \quad (1)$$

En el caso de grafos dirigidos, la matriz A no es necesariamente simétrica.

B. Métricas y algoritmos

Una vez modelada la red, es posible aplicar diversas métricas y algoritmos para caracterizarla. Desde el punto de vista de la centralidad, exploramos el *Grado* (Degree), la *Intermediación* (Betweenness) y *Cercanía* (Closeness).

a) Grados de entrada y salida (Degree Centrality): En las redes dirigidas los nodos tienen dos grados, de entrada y de salida. El grado de entrada es el número de enlaces conectados a un nodo. El grado de salida es el número de enlaces salientes de un nodo. Estas métricas permiten identificar reglas altamente influyentes (con muchas dependencias salientes) o altamente dependientes (con muchas dependencias entrantes):

$$\deg^-(r_j) = \sum_{i=1}^{n} A_{ij} \quad (2)$$

$$\deg^+(r_i) = \sum_{j=1}^{n} A_{ij} \quad (3)$$

b) Intermediación (Betweenness Centrality): Es otra forma de centralidad que mide en que grado un nodo se encuentra en el camino que une otros nodos. Permite identificar nodos que actúan como puentes críticos dentro de los flujos que tienen lugar en un sistema:

$$C_B(r) = \sum_{s \neq r \neq t} \frac{\sigma_{st}(r)}{\sigma_{st}} \quad (4)$$

donde σ_{st} es el número total de caminos geodésicos entre los nodos s y t, y $\sigma_{st}(r)$ es el número de esos caminos que pasan por el nodo r.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

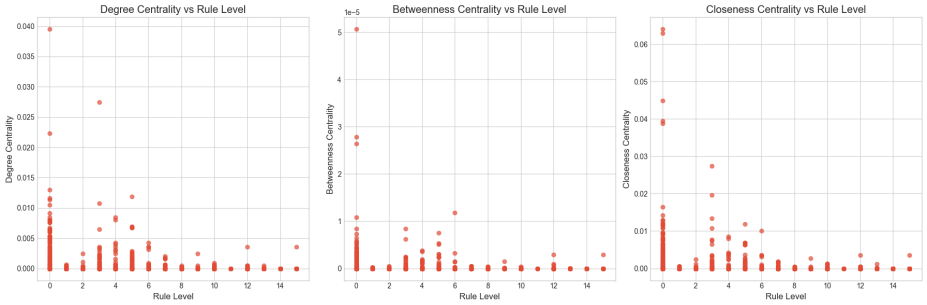


Fig. 1. Métricas de centralidad (grado, betweenness, closeness) en función del nivel de la regla.

c) Cercanía (Closeness Centrality): La cercanía mide la distancia media de un nodo a otro. La centralidad de cercanía de un nodo r se define como:

$$C_C(r) = \frac{n-1}{\sum_{t \in V, , t \neq r} d(r,t)} \tag{5}$$

donde $d(r,t)$ representa la longitud del camino más corto desde el nodo r hasta el nodo t, y la suma se realiza sobre todos los nodos alcanzables distintos de r.

V. ANÁLISIS DE WAZUH BAJO LA PERSPECTIVA DE REDES COMPLEJAS

A partir del modelado del conjunto de reglas de Wazuh como una red compleja, se han obtenido los primeros resultados preliminares. A continuación se presentan distintas secciones sobre estructura, centralidad, comunidades funcionales y análisis en infraestructura real para mostrar el potencial del enfoque.

A. Estructura general de la red

La red generada es un grafo dirigido y está compuesta por 4482 nodos (reglas) y 4323 enlaces (relaciones entre reglas). Aunque no se puede apreciar en la Fig. 2. debido a la magnitud de la red, esta sigue una estructura jerárquica. Comprobamos que la red tiene estructura de árbol y no está totalmente conectada, indicando la presencia de componentes o subredes dentro del conjunto global.

B. Análisis de centralidad

Con el objetivo de caracterizar la estructura interna de las dependencias entre reglas del conjunto se han estudiado las métricas de centralidad de *Grado*, *Intermediación* y *Cercanía* en función del nivel de criticidad de la regla, como se puede ver en la Fig. 1. La distribución de las métricas observadas revela una mayor concentración de valores altos en reglas de bajo nivel, especialmente en el nivel 0, lo cual indica su papel central en la red.

C. Componentes y Comunidades funcionales

El análisis de conectividad, segmentando la red en sus componentes débilmente conectados, revela que se compone de 420 subredes. El número de reglas por subred es muy variable, con componentes de un solo nodo hasta un componente con 1033 nodos. El tamaño medio es de 10 nodos. Esta distribución nos indica que en el conjunto

Fig. 2. Estructura de la red del conjunto de reglas de Wazuh.

de reglas de Wazuh coexisten múltiples grafos pequeños junto a grafos de mayor tamaño y complejidad.

Algunas subredes agrupan reglas procedentes de distintos archivos del conjunto de reglas, como se puede ver en la Tabla 1. Esto pone en evidencia dependencias lógicas entre reglas del conjunto que van más allá de la división lógica en archivos por la que se caracteriza Wazuh.

Tabla I
RESUMEN DE LAS 10 MAYORES SUBREDES DEL CONJUNTO

Subred	Tamaño	Multiarchivo	Archivos
S_1	1033	✓	18
S_2	321		1
S_3	198	✓	7
S_4	174		1
S_5	136	✓	2
S_6	128		1
S_7	128	✓	2
S_8	99		1
S_9	76		1
S_{10}	71	✓	4

D. Reglas aisladas y posibles errores de configuración

Del análisis de los componentes se descubren una serie de reglas aisladas. Estas reglas generan alertas de modo

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

directo una vez un evento cumple sus condiciones por lo que es interesante ver como se caracterizan a nivel de severidad.

De la distribución del nivel de estas reglas vemos que la mayoría se sitúan en niveles de severidad bajos, aunque existe un pico en el nivel 12. El estudio de estas reglas será relevante en el futuro ya que desde un punto de vista operativo son susceptibles de generar grandes volúmenes de alertas y generar ruido.

Por otro lado, también se ha encontrado una regla, concretamente la 91821, que forma parte del subgrafo más grande del conjunto, y que se auto referencia como condición de correlación mediante el campo `if_sid`, lo cual podría ser un error de configuración como se puede ver en la Fig. 3. La misma regla está relacionada con la 91820 también por lo que únicamente generará una alerta cuando esta se cumpla.

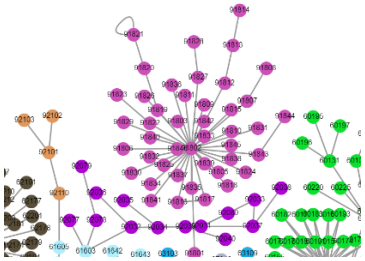


Fig. 3. Auto referencia de la regla 91821 (parte superior izquierda), parte del mayor subgrafo de la red.

E. Análisis en infraestructura real

Con el objetivo de validar empíricamente los resultados obtenidos, se ha iniciado un análisis de las alertas generadas por Wazuh durante 1 año en un entorno productivo que incluye un total de 150 millones de alertas.

La frecuencia de las reglas activadas revela un comportamiento desequilibrado, en el que un conjunto pequeño de reglas, entorno a un total de 10, son responsables de mas de un 90% de las alertas generadas como se puede apreciar en la Fig. 4.

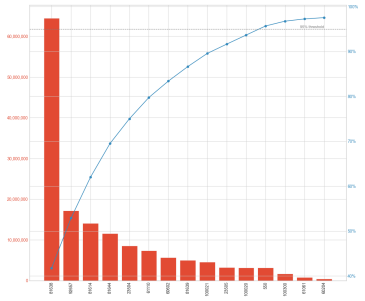


Fig. 4. Activación y porcentaje acumulado de reglas.

Esto confirma, al menos para este entorno, la existencia de un núcleo de reglas que tienen un impacto extremo en la operativa, reforzando la necesidad de implementar mecanismos de revisión y optimización continua de las reglas del conjunto.

VI. CONCLUSIONES

En este trabajo se ha presentado un análisis preliminar de la estructura interna del conjunto de reglas de Wazuh mediante técnicas de ARC.

El análisis sobre infraestructura real ha puesto de manifiesto la elevada concentración de alertas en un reducido subconjunto de reglas, lo que refuerza la importancia de estudiar el impacto operativo de las reglas centrales y su relación con el volumen de alertas generadas.

Aunque el trabajo se encuentra en progreso, los resultados obtenidos evidencian el potencial de los enfoques basados en ARC como herramienta de soporte para exploración de conjuntos de reglas en sistemas de ciberseguridad. En el futuro, este enfoque podría emplearse para recomendar la activación o desactivación de reglas en función del entorno, detectar la necesidad de nuevas reglas o combinar varias reglas con tal de reducir redundancias. Aunque el análisis se ha realizado sobre Wazuh, la metodología propuesta podría adaptarse a otras plataformas SIEM/XDR/IDS como Suricata o Snort.

Como líneas de trabajo futuro se plantea ampliar el análisis de correlación entre métricas de centralidad, estudiar en profundidad la estructura de comunidades y validar de forma empírica la relación entre la centralidad de las reglas y su comportamiento operativo en entornos productivos, así como la realización de análisis temporales sobre las alertas generadas.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por la Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya en el marco del programa de Doctorado Industrial.

This work has been partially funded by the European Cybersecurity Competence Centre (ECCC) under grant agreement No 101145874

REFERENCIAS

[1] Wazuh, "Wazuh - open source XDR. open source SIEM," 2025. [Online]. Available: https://wazuh.com/

[2] E. Estrada, *The structure of complex networks: theory and applications*. American Chemical Society, 2012.

[3] S. Stanković, S. Gajin, and M. I. R. Petrović, "A review of wazuh tool capabilities for detecting attacks based on log analysis," 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:254589228

[4] D. Šuškalo, Z. Morić, J. Redžepagić, and D. Regvart, "Comparative analysis of IBM QRADAR and Wazuh for security information and event management." *34TH DAAAM INTERNATIONAL SYMPOSIUM ON INTELLIGENT MANUFACTURING AND AUTOMATION*.

[5] N. Akshai Sankar and K. A. Fasila, "Implementation of SOC using ELK with integration of wazuh and dedicated file integrity monitoring," in *2023 9th International Conference on Smart Computing and Communications (ICSCC)*, 2023, pp. 350–354.

[6] A. Haydar and M. Ramparison, "Modeling wazuh rules with weighted timed automata," *Procedia Computer Science*, vol. 251, pp. 75–82, 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1877050924033209

[7] A. Hay, D. Cid, and R. Bray, "OSSEC host-based intrusion detection guide." [Online]. Available: www.sans.edu

[8] M. E. J. Newman, *Networks : an introduction*. Oxford University Press, 2017.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Prueba de concepto de comunicaciones semánticas para aplicaciones inmersivas

Juan L. Arenas-Sánchez[1], Jesús Galeano-Brajones[1], Francisco Díaz-Barrancas[1], Pedro J. Pardo[1], Francisco Luna[2,3], Javier Carmona-Murillo[1]
[1]Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos, Universidad de Extremadura, Mérida, 06800
[2]Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga, E.T.S.I. Informática, Málaga, 29071
[3]ITIS Software, Universidad de Málaga, Málaga, 29071
juanluisas@unex.es, jgaleanobra@unex.es, frdiaz@unex.es, pjpardo@unex.es, flv@lcc.uma.es, jcarmur@unex.es

Las comunicaciones semánticas representan un paradigma prometedor para optimizar la transmisión de información en aplicaciones inmersivas como la realidad virtual. Sin embargo, la mayoría de los trabajos existentes se centran en una perspectiva teórica o en simulaciones, sin explorar suficientemente su aplicación práctica en entornos inmersivos. Este trabajo tiene como objetivo desarrollar una prueba de concepto sobre comunicaciones semánticas extremo a extremo. Específicamente, se transmite la información semántica esencial de un vídeo para reconstruirlo como una experiencia inmersiva con un uso muy reducido de ancho de banda. La metodología propuesta se basa en un *pipeline* que codifica información semántica a partir de vídeo y la decodifica para generar representaciones en un espacio virtual tridimensional. Los resultados muestran que es posible representar con fidelidad escenas en tiempo real, manteniendo un contexto constante y con una latencia inicial asociada a la generación visual. Estos hallazgos abren nuevas posibilidades para sistemas eficientes en el ámbito de las aplicaciones inmersivas y metaverso.

***Palabras Clave*—comunicaciones semánticas, redes de próxima generación, realidad virtual, 6G, inteligencia artificial**

I. INTRODUCCIÓN

La evolución hacia la sexta generación de redes móviles (6G) ha impulsado un cambio fundamental desde el paradigma tradicional, basado en la transmisión exacta de bits, hacia enfoques más inteligentes, eficientes y conscientes del contexto de la comunicación [1]. En este escenario, emergen las comunicaciones semánticas (SC, *Semantic Communications*), cuyo objetivo es transmitir de manera eficiente el significado esencial para realizar tareas específicas, priorizando el impacto efectivo sobre el receptor por encima de la precisión sintáctica [2], [3]. Este paradigma, potenciado por el auge exponencial de la inteligencia artificial (IA) y el internet de las cosas (IoT), tiene diversas aplicaciones, desde la transmisión adaptativa de vídeo multicanal hasta redes vehiculares y sistemas autónomos, mejorando la eficiencia y confiabilidad mediante inferencia semántica [4], [5]. Sin embargo, la implementación efectiva de las SC plantea desafíos importantes, como la cuantificación del valor semántico y la interoperabilidad en entornos heterogéneos [6].

La Figura 1 muestra un esquema del funcionamiento típico de una SC. El transmisor emplea un codificador para realizar conjuntamente la codificación de fuente, la codificación del canal y la modulación. Por su parte, el receptor cuenta con decodificadores que realizan demodulación, decodificación del canal y de la fuente, y una decodificación semántica apoyada en técnicas de IA. En este proceso, el codificador semántico transmite únicamente la información imprescindible que, apoyada en el contexto compartido, permite al receptor interpretar y reconstruir un mensaje acorde con la intención comunicativa. Por ejemplo, si el objetivo es transmitir una imagen, en lugar de enviarla completa se podría enviar una descripción textual de sus elementos clave. El decodificador semántico utilizaría entonces dicha descripción y técnicas de IA para generar una imagen similar, preservando únicamente el significado para la tarea específica.

En la literatura reciente, se ha evidenciado un creciente interés en la aplicación de comunicaciones semánticas para la transmisión eficiente de contenido visual, especialmente en escenarios exigentes como redes inalámbricas ruidosas, realidad virtual (VR) y transmisión en tiempo real. Algunos ejemplos de este interés se reflejan en trabajos como el de *Miao et al.* en [7], que proponen un sistema de comunicación semántica basado en *vector quantization* para vídeos de VR, denominado SeVQ-VR. El modelo introduce una técnica de segmentación semántica en mosaicos que permite asignar distintos niveles de prioridad a las regiones de la imagen en función de su relevancia

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

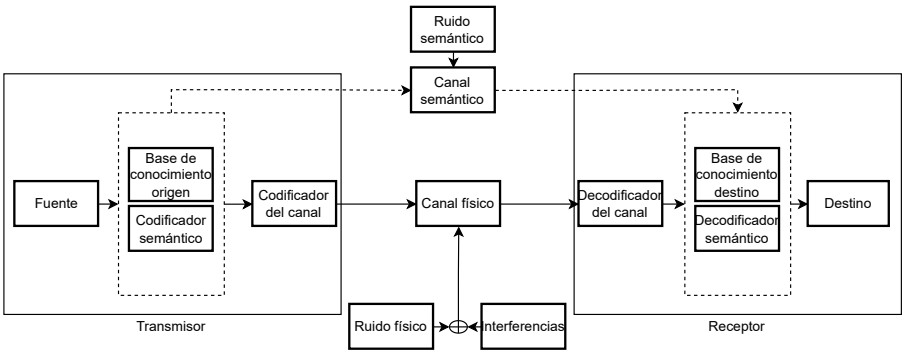


Fig. 1: Diagrama de una comunicación semántica con sus elementos.

para la experiencia inmersiva. Otro trabajo relacionado es el de *Qi et al.* en [8], donde presentan WAFI-VSC, un sistema de comunicación semántica de vídeo basado en interpolación de fotogramas adaptativa. Su objetivo es habilitar una transmisión de vídeo de ultra baja latencia en entornos inalámbricos mediante la predicción de frames intermedios usando modelos semánticos y redes neuronales convolucionales.

En este contexto, este trabajo presenta una prueba de concepto sobre el uso de SC para la transmisión de vídeo en tiempo real, incorporando además VR para permitir al usuario interactuar dinámicamente con la escena generada. Entre las principales contribuciones destacan:

- La validación experimental del uso de SC para reducir significativamente la carga de red al transmitir vídeos en tiempo real.
- La integración de tecnologías de VR en el decodificador semántico para situar dinámicamente personas en entornos virtuales.
- La descripción detallada del *pipeline* semántico implementado, desde la captura en la fuente de datos hasta la visualización final en escenarios inmersivos.

Este trabajo está estructurado de la siguiente manera. En la Sección II, se describe detalladamente la solución propuesta. A continuación, en la Sección III, se presentan y discuten los resultados obtenidos. Finalmente, en la Sección IV, se exponen las conclusiones alcanzadas y futuras líneas de trabajo.

II. SOLUCIÓN PROPUESTA

En esta sección se presenta en profundidad la solución propuesta para la implementación de SC en la transmisión de vídeo en tiempo real, detallando su diseño conceptual, desarrollo e implementación práctica.

En el contexto de las SC, un trabajo relevante tomado como referencia es el de *Yang et al.* en [9], donde se abordan los desafíos en redes inalámbricas derivados de los grandes volúmenes de datos y las frecuentes actualizaciones en escenarios como la videovigilancia remota. En este sentido, las investigaciones existentes suelen centrarse en dos enfoques principales: (i) extracción semántica, usando técnicas basadas en el aprendizaje profundo para compresión, y (ii) muestreo semántico, orientado por la antigüedad de la información (AoI). Aunque estos enfoques han demostrado mejoras en eficiencia energética, generalmente consideran solo atributos intrínsecos de la información, como la compresibilidad [10]. Sin embargo, suelen descuidar el papel crítico que desempeña la información contextual específica de la tarea. De hecho, solo la información semántica estrictamente relevante para el objetivo final es realmente necesaria en destino, lo que sugiere que integrar información intrínseca y extrínseca podría maximizar significativamente la eficiencia [11].

Para abordar esta limitación, *Yang et al.* proponen un novedoso framework modular denominado comunicación semántica generativa dirigida por agente (A-GSC, por sus siglas en inglés). Este framework plantea un rediseño integral de los procesos involucrados en el procesamiento, muestreo, transmisión y reconstrucción de información. A diferencia de estudios previos, que tratan de forma aislada la extracción o el muestreo semántico, A-GSC integra ambos enfoques considerando simultáneamente los atributos intrínsecos de la fuente y el contexto específico de la tarea.

Inspirándonos en este enfoque, proponemos un framework modular para transmitir vídeo desde una cámara de vigilancia en una sala cerrada, donde se detectan personas en movimiento y se visualizan en un entorno de VR). La solución se estructura en varios módulos interconectados, como se muestra en la Figura 2.

En nuestro sistema propuesto, el codificador semántico comienza a partir de imágenes de vídeo capturadas por una cámara, sobre las cuales se aplica el modelo YOLOv11 para detectar y localizar a las personas presentes en cada frame. A partir de estas detecciones, se genera un archivo JSON que contiene las coordenadas espaciales de cada individuo identificado, constituyendo así la información semántica a transmitir. En el extremo receptor, el decodificador semántico utiliza esta información para generar, mediante el modelo Shap-E, un modelo 3D por cada nueva persona detectada, que posteriormente es incorporado al escenario tridimensional desarrollado en Unity. Este escenario constituye el contexto semántico compartido, y permanece constante durante toda la comunicación. Su geometría ha sido diseñada para reflejar las dimensiones físicas reales del entorno, lo que permite ubicar a cada persona en el espacio virtual de forma coherente con la realidad.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

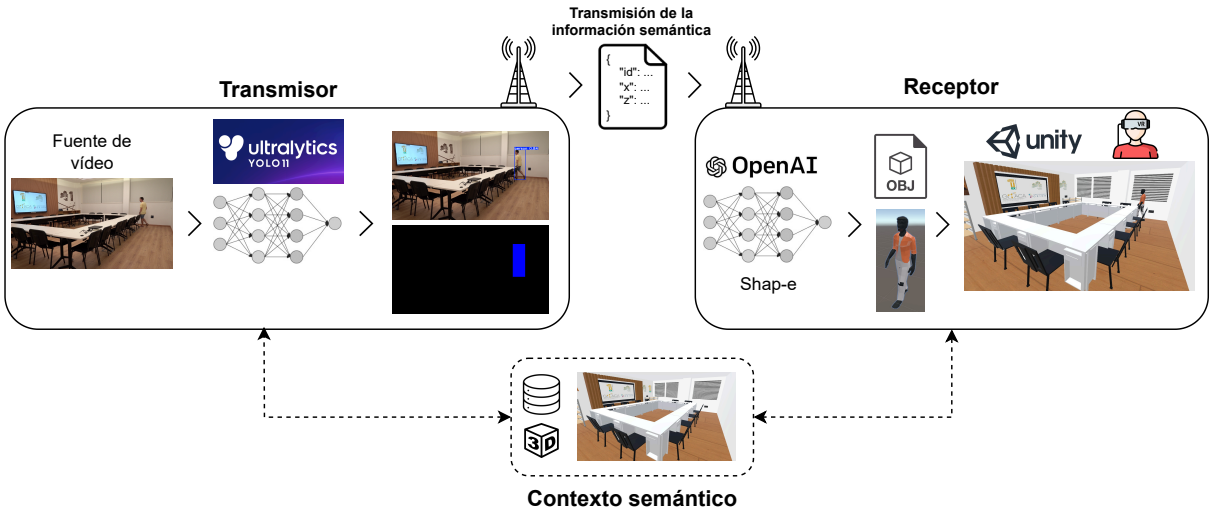


Fig. 2: *Pipeline* de comunicación semántica propuesto.

III. RESULTADOS Y DISCUSIÓN

En esta sección primero se detalla la metodología experimental para asegurar la reproducibilidad del trabajo. Después, se analizan y discuten los resultados.

A. Metodología experimental

En lo que respecta al modelo YOLOv11, se han ajustado diversos parámetros para adaptarlo a nuestro entorno específico y necesidades particulares. Este modelo ya está previamente entrenado para la detección de múltiples objetos, pero en nuestro caso particular se ha limitado exclusivamente a la clase "persona". La infraestructura utilizada para realizar la inferencia del modelo ha consistido en un equipo con una GPU NVIDIA A100.

Para la utilización del modelo Shap-E se han considerado diferentes parámetros clave que permiten optimizar la relación entre la calidad de los modelos generados y la carga computacional necesaria, así como los tiempos de respuesta. Es importante destacar que hay que mantener un equilibrio adecuado entre estos parámetros, ya que influye directamente en la calidad visual y la fidelidad al texto original (*prompt*), así como en la carga computacional requerida y la creatividad del modelo.

B. Resultados

A continuación, se presentan los principales resultados obtenidos a partir de la implementación del sistema de comunicación semántica propuesto, evaluando su eficiencia en términos de compresión, latencia, fidelidad visual y rendimiento de transmisión.

1) Tasa de compresión: La Figura 3 muestra la evolución de la tasa de compresión, definida como el cociente entre el tamaño del *frame* original y el del archivo JSON resultante de la codificación semántica. Los valores oscilan alrededor de 2.2×10^5 con una desviación típica de 7×10^3, lo que implica que el volumen de datos transmitido en la comunicación semántica representa menos del 0.00046 % del tamaño original de las imágenes del vídeo. Las únicas caídas apreciables corresponden a los primeros *frames*, cuyo menor tamaño se debe a la cercanía de la persona a la cámara.

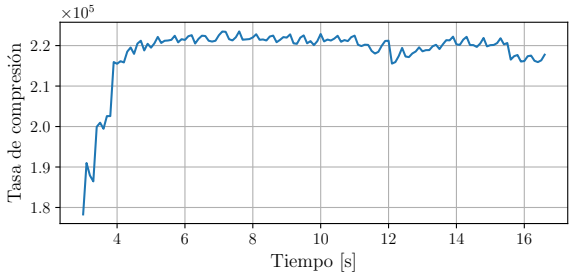


Fig. 3: Tasa de compresión entre el tamaño de cada *frame* y el tamaño de los JSON.

2) Latencia y tamaño del modelo 3D: El modelo Shap-E genera cada objeto 3D en unos 5 s sobre una GPU NVIDIA A100 (40 GB VRAM). Para equilibrar latencia y calidad se emplea una resolución media–baja, obteniendo ficheros `.obj` de $\sim$2 MB. Aunque el retardo inicial en la generación del modelo 3D es significativo ($\sim$ 5 s), se ha gestionado integrándolo de forma coherente en la escena virtual: la línea temporal del entorno inmersivo se alinea con la del mundo real, pero desplazada en el tiempo por dicho retardo. Esto permite mantener la fidelidad temporal de la experiencia sin comprometer la coherencia narrativa ni la percepción del usuario.

3) Throughput: En la Figura 4 se compara el *throughput* medio de la transmisión tradicional y la semántica. Es decir, se compara la transmisión de *frames* del vídeo con la transmisión de la información semántica en los ficheros JSON. La diferencia de un orden de magnitud confirma la eficiencia de la comunicación semántica, mientras que la estabilidad de ambas curvas indica un rendimiento consistente a lo largo del tiempo.

4) Correspondencia visual: En la fuente y receptor de la Figura 2 se ejemplifica un *frame* de entrada y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

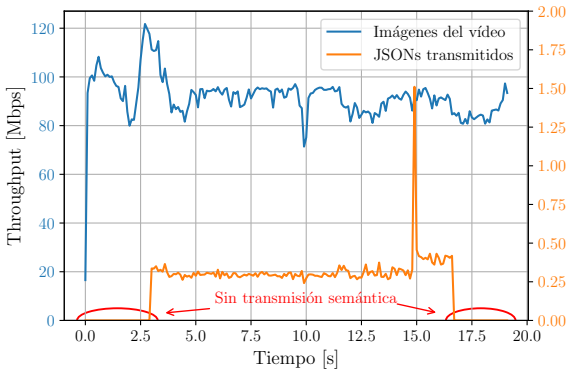


Fig. 4: Comparación de *throughput* en el envío de *frames* y en el envío de JSON.

su reconstrucción en Unity. El modelo 3D se posiciona con precisión gracias a las coordenadas semánticas, evidenciando la coherencia espacial entre entorno físico y virtual.

C. Discusión

Durante el desarrollo del sistema se han identificado varios desafíos clave que abren líneas prometedoras para trabajos futuros. En primer lugar, la latencia en la generación de objetos 3D supone una limitación importante para alcanzar una experiencia verdaderamente en tiempo real. Resulta prioritario explorar técnicas que aceleren este proceso manteniendo un equilibrio entre calidad visual y rendimiento.

Asimismo, se plantea la necesidad de enriquecer la interacción dentro del entorno virtual, permitiendo que los personajes puedan realizar acciones complejas como manipular objetos o interactuar con dispositivos físicos, lo que aumentaría tanto el realismo como las posibilidades de aplicación. Otro aspecto a mejorar es la virtualización precisa de espacios físicos a entornos 3D, aún limitada por la elevada carga computacional y los requisitos de fidelidad espacial.

Además, integrar información semántica más rica (como gestos, posturas corporales o expresiones faciales) podría incrementar notablemente la inmersión y precisión de la experiencia virtual. Por último, es fundamental evaluar el sistema en condiciones de red variables (por ejemplo, redes móviles con interferencias y latencias fluctuantes) para validar su robustez y adaptabilidad en escenarios operativos reales.

IV. CONCLUSIONES

En este trabajo se ha probado la viabilidad de aplicar comunicaciones semánticas a la transmisión de vídeo en tiempo real mediante una arquitectura modular que conecta desde la captura de información hasta su visualización en entornos inmersivos. A través de la integración de técnicas de inteligencia artificial generativa, se han generado objetos 3D a partir de descripciones semánticas, incorporándolos con precisión en un escenario virtual.

La metodología propuesta permite una reducción drástica en la cantidad de datos transmitidos, optimizando el uso del ancho de banda sin comprometer la fidelidad espacial de la escena. Como resultado, se ha logrado recrear un entorno virtual que refleja fielmente el movimiento de personas en un espacio físico real, con un consumo de recursos de red mínimo. Esta solución abre nuevas oportunidades para experiencias de realidad virtual más eficientes, especialmente relevantes en el contexto del metaverso y otras aplicaciones inmersivas.

AGRADECIMIENTOS

Este trabajo ha sido cofinanciado por el Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación (MICIU/AEI/10.13039/501100011033) bajo el proyecto PID2023-151462OB-I00, cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER); por la Unión Europea, Fondo Europeo de Desarrollo Regional, y la Junta de Extremadura (autoridad de gestión: Ministerio de Hacienda) mediante las ayudas GR24177 y GR24030; y la cátedra INCIBE-UEx-CUMe (C110.23) en el marco "European Union NextGenerationEU/PRTR".

REFERENCIAS

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2019.

[2] E. E. Aydın, E. Çetin, T. T. Sarı, and G. Seçinti, "From Theory to Practice: A Testbed for DL-Based Semantic Communication Architecture," in *2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC)*. IEEE, 2025, pp. 1–2.

[3] S. Guo, A. Zhang, Y. Wang, C. Feng, and T. Q. Quek, "Semantic-Enabled 6G Communication: A Task-oriented and Privacy-preserving Perspective," *IEEE Network*, 2025.

[4] H. Gao, M. Sun, X. Xu, X. Cheng, S. Han, and P. Zhang, "Adaptive Cross-Modal Super-Resolution Semantic Communication for Mobile AI-Generated Panoramic Video," *IEEE Transactions on Cognitive Communications and Networking*, 2025.

[5] T. T. Win, A. B. Wondmagen, T. M. T. Nguyen, J. Oh, G. Kim, and S. Cho, "A Survey on Energy-Efficient in Semantic Communication: Techniques, Challenges, and Future Directions," in *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, 2025, pp. 0001–0006.

[6] S. Sharif, F. Khandaker, and W. Ejaz, "Semantic communication: Implication for resource optimization in 6G networks," in *2024 IEEE International Conference on Advanced Telecommunication and Networking Technologies (ATNT)*, vol. 1. IEEE, 2024, pp. 1–4.

[7] Y. Miao, D. Hu, and Y. Wang, "An Efficient Vector Quantization-Based Semantic Communication System for Virtual Reality Video Transmission," *IEEE Wireless Communications Letters*, 2025.

[8] X. Qi, N. Ma, Z. Bao, Y. Liu, C. Dong, and X. Xu, "WAFI-VSC: Wireless Adaptive Frame Interpolation Video Semantic Communication," in *2024 16th International Conference on Wireless Communications and Signal Processing (WCSP)*. IEEE, 2024, pp. 1503–1508.

[9] W. Yang, Z. Xiong, Y. Yuan, W. Jiang, T. Q. Quek, and M. Debbah, "Agent-driven Generative Semantic Communication with Cross-Modality and Prediction," *IEEE Transactions on Wireless Communications*, 2024.

[10] S. Ma, W. Qiao, Y. Wu, H. Li, G. Shi, D. Gao, Y. Shi, S. Li, and N. Al-Dhahir, "Task-oriented explainable semantic communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9248–9262, 2023.

[11] F. Pase, S. Kobus, D. Gündüz, and M. Zorzi, "Semantic communication of learnable concepts," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 731–736.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Herramienta de Preprocesamiento de Tráfico IoMT para Generación Configurable de *Datasets*

Olga León, Isabel V. Martin-Faus, Josep Pegueroles, Jordi Doménech
Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya (UPC),
c. Jordi Girona 1-3, 08034, Barcelona, Spain.
{olga.leon, isabel.martin-faus, josep.pegueroles, jordi.domenech.fons}@upc.edu

La incorporación de dispositivos IoT médicos (*IoMT, Internet of Medical Things*) está transformando el sector sanitario, mejorando la monitorización de pacientes y optimizando la gestión hospitalaria. También está incrementando la superficie de ataque, exponiendo datos sensibles y poniendo en riesgo la seguridad del paciente. Entre los métodos para proteger los entornos IoMT destacan los sistemas de detección y prevención de intrusiones (IDPS) basados en Inteligencia Artificial (IA). Estos requieren de bases de datos de calidad para entrenar los modelos de aprendizaje automático y lograr una eficiencia elevada. En este trabajo presentamos una herramienta que, a partir de tráfico de red IoMT y la extracción configurable de atributos relevantes, facilita la generación de *datasets* y, en particular, obtener *datasets* específicos que permitan entrenar el IDPS según el modelo de IA usado. La herramienta incluye hasta 103 características distintas de tráfico de red, esto es el número más elevado tenido en cuenta hasta la fecha. La flexibilidad en la selección de parámetros mejora la calidad de los datos utilizados e incrementará la eficiencia de los IDPS. La herramienta ha sido validada dentro del *testbed* (MIoTTA) desarrollado en la UPC para obtener tráfico de red según el comportamiento de una habitación hospitalaria con sistema de telemetría. Su diseño modular ofrece una solución versátil para investigadores y profesionales que buscan generar *datasets* robustos y realistas para la detección proactiva de amenazas en entornos médicos conectados.

***Palabras Clave*—IoMT, Sistemas de Detección de Intrusión, Datasets, Testbed, Aprendizaje Automático.**

I. INTRODUCCIÓN

La Internet de las Cosas (IoT, *Internet of Things*) está transformando el sector sanitario, permitiendo una atención médica más eficiente, personalizada y en tiempo real. Los dispositivos médicos conectados (IoMT, *Internet of Medical Things*), que tanto monitorizan constantes vitales como automatizan la gestión hospitalaria, mejoran significativamente la calidad de los servicios de salud [1]. Sin embargo, los avances en digitalización también plantean importantes desafíos de ciberseguridad exponiendo datos sensibles a vulnerabilidades y amenazas, que en ocasiones incluso pueden comprometer la integridad física del paciente [2].

La protección de redes IoMT frente a ciberataques es especialmente crítica en dispositivos activos (por ejemplo, bombas de infusión, ventiladores mecánicos, desfibriladores implantables y marcapasos), cuyas acciones automatizadas pueden afectar severamente la salud de los pacientes [3]. Los Sistemas de Detección y Prevención de Intrusiones (IDPS, *Intrusion Detection and Prevention System*) son una buena solución para identificar las amenazas y ofrecer una respuesta rápida ante ataques [4].

Entre las tecnologías usadas para el diseño de IDPS en IoMT, la inteligencia artificial (IA) es una de las más prometedoras [5], [6]. La IA permite analizar grandes volúmenes de datos en tiempo real, identificar patrones anómalos y predecir amenazas con precisión y rapidez. Los modelos de aprendizaje automático (ML, *Machine Learning*) son los más explorados actualmente para reducir falsos positivos y optimizar la respuesta en sistemas críticos de salud [7].

Es conocido que la IA requiere de *datasets* amplios, representativos y fiables para su correcto entrenamiento, validación y prueba. La calidad de los datos en estos tres aspectos es crucial e influye directamente en la precisión y capacidad de generalización del modelo. En el ámbito de la ciberseguridad, además, la adquisición de datos debe incluir ataques lo más diversos posibles, así como comportamientos normales para garantizar que el sistema pueda identificar y clasificar amenazas con alta fiabilidad.

Este trabajo tiene como objetivo proporcionar una her-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

ramienta software para la generación de *datasets* (en formatos CSV y *nPrint*), a partir de ficheros PCAP de capturas de tráfico de red.

El resto del artículo está organizado de la siguiente manera: La Sección II analiza trabajos similares en *datasets* existentes, y destaca sus limitaciones. La Sección III introduce brevemente el *pipeline* de preprocesado y lo ubica en el proceso global de funcionamiento de un IDPS. Las Secciones IV y V, respectivamente, describen la herramienta en sí con sus funcionalidades y muestran un caso de uso en el *testbed* desarrollado por el equipo investigador. Finalmente, la Sección VI presenta conclusiones y líneas futuras.

II. TRABAJOS RELACIONADOS

En el ámbito de la ciberseguridad se conocen distintos trabajos que emplean herramientas de preprocesamiento de tráfico de red, en particular para entornos IoT e IoMT, con el objetivo de entrenar modelos de detección de intrusiones basados en inteligencia artificial. Las herramientas suelen convertir archivos de captura de paquetes de red (PCAP) en representaciones estructuradas que contienen atributos específicos sobre el tráfico observado. No obstante, la mayoría de las herramientas existentes presentan limitaciones importantes en términos de configurabilidad y granularidad. Además, la mayoría de ellas no están disponibles públicamente. Esto restringe su aplicabilidad a experimentos personalizados o necesidades específicas para la investigación. Los detalles de estas herramientas se describen en la Tabla I.

Por ejemplo, los *datasets* CICIoT2023 [12] y CICIoMT2024 [13], desarrollados por el Canadian Institute for Cybersecurity (CIC), utilizan herramientas internas para la extracción de atributos de red. Estas herramientas, las cuales están disponibles parcialmente de forma pública, extraen atributos a través de ventanas de paquetes y flujos. El CIC, además, tiene desarrollada una herramienta interna llamada CICFlowMeter [8] para la extracción de características. Aunque esta herramienta sí que está disponible públicamente, solo opera exclusivamente a nivel de flujo, con un conjunto predefinido de atributos, sin ofrecer al usuario la posibilidad de configurar el tipo de extracción o la selección de atributos.

El *dataset* IoT-Healthcare [14] también utiliza una herramienta propia, disponible públicamente, para convertir archivos PCAP a CSV. Los archivos finales (en formato CSV) incorporan únicamente 50 atributos, y éstos sólo se extraen a través de paquetes individuales.

Una situación similar se observa en otros *datasets* como IoMT-TrafficData [15] y WUSTL-EHMS [16]. En éstos las herramientas de preprocesamiento son cerradas, poco documentadas o limitadas a un número muy reducido de características. Además, el tipo de extracción de los atributos es limitado a flujos y paquetes, o a flujos y datos biométricos, respectivamente.

En contraste, la herramienta propuesta en este estudio ofrece una solución integral y altamente configurable. Es capaz de extraer atributos de red a nivel de paquete, flujo, ventana de paquetes o mediante la combinación de todos estos, permitiendo generar hasta 103 atributos distintos, todos configurables por el usuario. La herramienta también es capaz de extraer el contenido de los paquetes bit a bit, mediante el estándar *nPrint* [17]. Esta flexibilidad convierte nuestra propuesta en una herramienta única y novedosa de preprocesamiento de tráfico de red, especialmente orientada a la generación de *datasets* personalizados para la detección de amenazas en entornos IoMT.

III. PIPELINE PARA LA RECOLECCIÓN Y PROCESADO DE DATOS

Típicamente, en escenarios como el que nos ocupa, la recolección y procesado de datos sigue una *pipeline* estructurada, ilustrada en la Fig. 1, que permite capturar, procesar y almacenar el tráfico de red generado por los dispositivos IoMT en distintos formatos. La captura inicial de datos se realiza en formato PCAP, que posteriormente puede ser convertido a CSV o a *nPrint* [17] para entrenar y evaluar modelos IA de detección de ciberataques. Estos modelos nos permiten detectar comportamientos de red anómalos de forma precisa y automatizada, evitando así filtraciones de datos médicos o ataques que puedan afectar a la salud de los pacientes.

El objetivo principal de esta *pipeline* es proporcionar una metodología para desarrollar *datasets* configurables y flexibles, que permitan evaluar múltiples enfoques de modelado en el ámbito de la ciberseguridad en entornos médicos. Gracias a la herramienta de preprocesado propuesta, los investigadores podrán generar datos con distintas configuraciones para poder entrenar algoritmos de inteligencia artificial y estudiar su comportamiento. En esta *pipeline* también describimos el *testbed* MIoTTA-UPC que estamos desarrollando en la UPC como parte de la recolección de datos de dispositivos IoMT. Este *testbed* se describe en más detalle en la sección V.

A. Recolección de datos

El tráfico de red generado en el *testbed* se captura utilizando *Wireshark* [18] y se almacena en ficheros PCAP. Esta captura incluye las comunicaciones entre los dispositivos IoMT físicos y simulados, y el *gateway*, permitiendo la observación de patrones de tráfico bajo condiciones normales y de ataque.

Las capturas pueden realizarse de forma continua o en intervalos de tiempo predefinidos, dependiendo del experimento en cuestión. Se registran los paquetes que circulan por la red, pertenecientes a de MQTT y CoAP, así como otros protocolos utilizados en la comunicación (p. ej., ARP, TCP, etc.).

B. Preprocesamiento de los datos

Una vez capturado el tráfico en formato PCAP, se utiliza la herramienta configurable de preprocesado, descrita en detalle en la sección IV, para convertir los datos en formatos más adecuados para el entrenamiento de modelos de inteligencia artificial; como son CSV y *nPrint*.

El formato CSV se obtiene extrayendo estadísticas del tráfico de red a partir de la captura PCAP, y bajo tres

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I: Tabla comparativa de herramientas de preprocesado en entornos IoMT conocidos.

Herramienta/Dataset	Año	Tipo de Extracción	Formato de Datos	Configurable	Nº de atributos	Pública	Referencia
CICFlowMeter	2016	Flujos	CSV	No	76	Sí	[8]
WUSTL-EHMS	2020	Flujos	CSV	No	34	No	-
IoT-Healthcare	2021	Paquetes	CSV	No	50	Sí	[9]
CICIoT2023	2023	Flujos, Ventanas	CSV	No	47	Sí	[10]
CICIoMT2024	2024	Flujos, Ventanas	CSV	No	45	No	-
IoMT-TrafficData	2024	Flujos, Paquetes	CSV	No	64	No	-
Herramienta propuesta	**2025**	**Flujos, Paquetes, Ventanas**	**CSV, nPrint**	**Sí**	**Hasta 103**	**Sí**	[11]

enfoques diferentes: análisis por paquetes, análisis por flujos y análisis por ventana de paquetes. Este enfoque flexible permite analizar el tráfico a distintos niveles de granularidad, facilitando la extracción de patrones relevantes para la detección de ciberataques.

El segundo formato en el que se convierten los datos es *nPrint* [17], el cual proporciona una representación binaria de los paquetes capturados. En este caso, el script transforma directamente el contenido de los campos de cada paquete en una secuencia de bits. Esto permite que modelos de aprendizaje profundo (*deep learning*) puedan analizar la estructura de los paquetes sin la necesidad de aplicar ingeniería de características de red de forma manual.

C. Entrenamiento y evaluación de algoritmos de detección

La última fase de la *pipeline* consiste en entrenar y evaluar los algoritmos de detección de intrusiones basados en IA, utilizando los datos capturados y procesados en los pasos anteriores. Asimismo, se pueden utilizar distintos tipos de algoritmos, como por ejemplo árboles de decisiones (*Decision Tree*, *Random Forest* o *XGBoost*), y redes neuronales (*FeedForward Neural Networks* o redes neuronales basadas en algoritmos de *Time Series Classification*), entre otros. Se ha podido comprobar el buen rendimiento de estos algoritmos en estudios anteriores, dando resultados excelentes en la detección de ciberataques en redes IoMT [19].

IV. HERRAMIENTA PROPUESTA

En esta sección presentamos la herramienta propuesta: una herramienta flexible y extensible desarrollada para convertir archivos de red en formato PCAP, a representaciones numéricas en formato CSV o *nPrint*. Está diseñada específicamente para algoritmos de detección de intrusiones basados en inteligencia artificial, con un enfoque especial a redes de dispositivos IoMT.

A. Descripción de la herramienta

La herramienta se ha diseñado con el objetivo de ofrecer un preprocesamiento altamente configurable, permitiendo a los investigadores o desarrolladores seleccionar tanto el nivel de granularidad de los datos, como el conjunto exacto de atributos a extraer. Asimismo, la herramienta desarrollada permite obtener un conjunto de ficheros en formato CSV y *nPrint* a partir de capturas de red en formato PCAP. Entre las principales funcionalidades de la herramienta, destacan:

- Conversión de tráfico PCAP a CSV estructurado con múltiples estadísticas.
- Soporte para tres niveles de agregación:
  - Nivel de paquete: se extraen los atributos de cada paquete de manera individual. De este modo, cada paquete de la captura PCAP da lugar a una línea del fichero CSV.
  - Nivel de flujo: se identifican los paquetes asociados a un mismo flujo (i.e., aquellos que contienen las mismas direcciones IP y puertos, tanto origen como destino), y se extraen los atributos por flujo. En este caso, el número de líneas del fichero de salida dependerá del número de flujos identificados en la captura.
  - Nivel de ventana: los paquetes de la captura PCAP se agrupan en ventanas de N paquetes, y se extraen los atributos promediando los valores de todos los paquetes de la ventana, dando lugar a una línea del fichero de salida por cada N paquetes a la entrada.
  - Nivel combinado (flujos y ventanas): se agrupan los paquetes de la captura en ventanas de N paquetes, se identifican los flujos dentro de dicha ventana y se extraen los atributos por flujo. De esta manera, los atributos extraídos serán una adición de los atributos a nivel de flujo y a nivel de ventana.
- Exportación alternativa a formato *nPrint*, donde se extrae la información bit a bit de cada uno de los paquetes de red capturados.

En la Fig. 2 podemos ver el diagrama de flujo de la herramienta, donde la entrada es un fichero PCAP. En cuanto a la salida, esta puede ser un fichero CSV con los atributos de red especificados por el usuario, o un fichero cuyo formato sea *nPrint*.

B. Extracción de atributos de tráfico de red

La conversión de PCAP a CSV se realiza mediante la extracción de atributos estadísticos y estructurales de cada paquete, flujo o ventana. Para hacerlo, se utiliza la herramienta *tshark* [20]: *tshark* nos ayuda a extraer las estadísticas básicas de los paquetes, para, posteriormente, poder calcular los atributos específicos del tráfico de red. Estos atributos permiten representar de forma numérica el comportamiento de la red, facilitando su procesamiento por modelos de inteligencia artificial. La herramienta ofrece hasta 103 atributos de red distintos, que pueden ser activados o desactivados por el usuario, agrupados en las

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

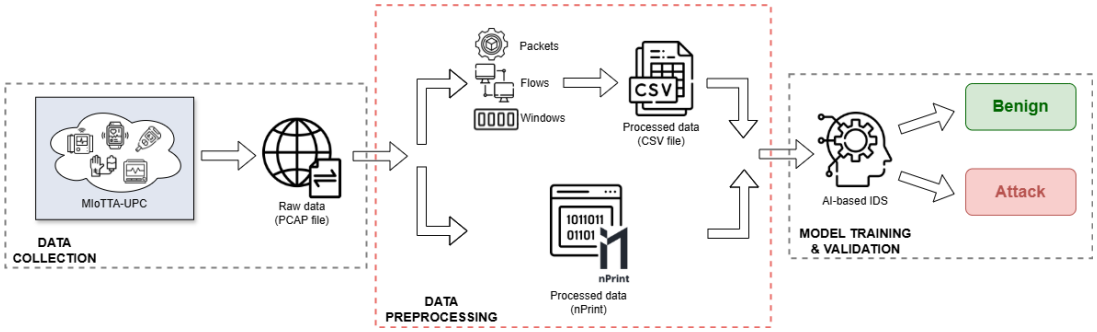


Fig. 1: *Pipeline* para la recolección y preprocesado de datos del *testbed* MIoTTA-UPC.

categorías de paquete (25 atributos), flujo (42 atributos) o ventana (36 atributos). Estos atributos y su significado se presentan en la Tabla II del Apéndice.

Además de la flexibilidad de la herramienta, cabe destacar que, a diferencia de otras herramientas existentes como CICFlowMeter [8] o IoMT-TrafficData [15], que únicamente extraen un conjunto reducido de características básicas de flujos o flujos y paquetes, respectivamente; nuestra herramienta proporciona una extracción mucho más exhaustiva y flexible de características de red, lo que la convierte en una solución más completa para entornos de análisis avanzados y detección de intrusiones. Asimismo, la herramienta también ha sido diseñada para permitir añadir nuevos atributos en el futuro mediante una arquitectura modular.

C. Soporte para formato nPrint

Además de la conversión a CSV con los atributos de tráfico de red, la herramienta ofrece la opción de transformar los paquetes PCAP al formato *nPrint* [17], una representación binaria de paquetes diseñada para alimentar directamente algoritmos de deep learning. En este modo, los paquetes son trasformados en vectores de bits mediante un proceso configurable:

- Posibilidad de truncar a los primeros N bytes del paquete.
- Inclusión de encabezados específicos, como pueden ser Ethernet, IP, TCP, UDP, ICMP y ARP.
- Opciones para mantener o eliminar la carga útil (payload).
- Formato de salida compatible con frameworks como PyTorch y TensorFlow.

Esta funcionalidad es especialmente útil en modelos de aprendizaje profundo, tales como modelos basados en redes neuronales convolucionales (CNN) o modelos híbridos, que analizan directamente la estructura binaria del tráfico. Además, la flexibilidad de la herramienta, que permite al usuario seleccionar de forma personalizada los atributos a emplear, facilita una reducción significativa del tamaño de los conjuntos de datos utilizados en estos modelos. En particular, el formato original de *nPrint* genera representaciones altamente detalladas y voluminosas, que pueden resultar difíciles de procesar en entornos con recursos computacionales limitados. Con la herramienta propuesta, el usuario puede configurar de manera eficiente los atributos extraídos, seleccionando únicamente aquellos más relevantes para su caso de estudio, lo que mejora la escalabilidad y manejabilidad del proceso de análisis.

V. USO DE LA HERRAMIENTA EN UN ENTORNO IOMT

En esta sección se describe el entorno de validación (*testbed*) empleado para comprobar la utilidad práctica de la herramienta desarrollada, centrado en un contexto realista del Internet de las Cosas Médicas (IoMT). Asimismo, se explica su uso, destacando las funcionalidades clave mediante ejemplos gráficos extraídos de su ejecución.

A. Entorno experimental: MIoTTA-UPC

La arquitectura del *testbed* MIoTTA-UPC [21], [22] representada en la Fig. 3, está diseñada para replicar un entorno realista de dispositivos IoMT en un contexto hospitalario, simulando una habitación de hospital, donde uno o más pacientes están conectados a distintos dispositivos médicos. Esta arquitectura combina dispositivos físicos reales con dispositivos simulados para garantizar la diversidad y realismo de los datos obtenidos. Esto permite el estudio de ciberataques en un entorno configurable, mejorando así la robustez de los algoritmos de detección de intrusiones basados en inteligencia artificial.

1) Infraestructura de Red y Flujo de Datos: La arquitectura del *testbed* MIoTTA-UPC se organiza en diferentes capas funcionales que simulan el flujo de datos en un entorno hospitalario:

- Dispositivos IoMT (físicos y simulados) en la habitación del hospital, los cuales generan datos biomédicos y los envían al *gateway* utilizando protocolos ligeros como MQTT o CoAP.
- *Gateway* que actúa como intermediario en la comunicación; se encarga de recibir los datos de los sensores y reenviarlos al *Router/Access Point (AP)* utilizando comunicación WiFi (*IP-based*).
- *Router/AP* que gestiona la conectividad de la red hospitalaria, asegurando la transmisión de los datos hacia el servidor *proxy*.
- Servidor *Proxy* que permite que el personal sanitario acceda en tiempo real a la información del paciente sin comprometer la seguridad del servidor central.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

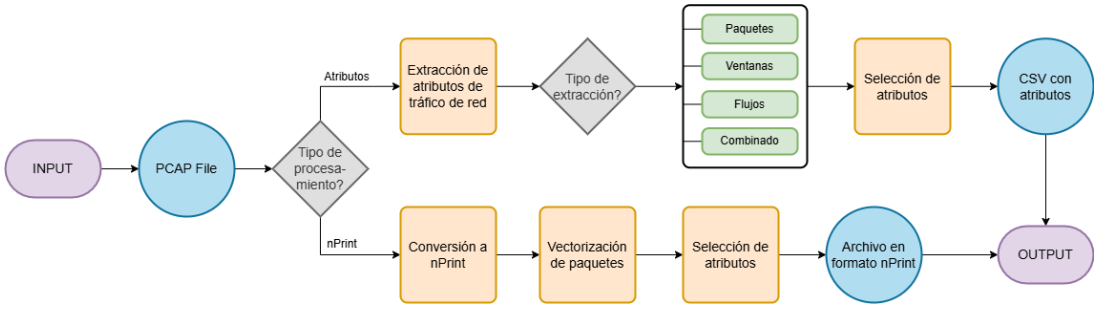


Fig. 2: Diagrama de flujo de la herramienta propuesta.

- Servidor del Hospital que almacena la información médica relevante para el análisis centralizado de los datos.

Estas capas funcionales proporcionan un entorno controlado, flexible y realista para la evaluación de diferentes estrategias de detección de ciberataques.

2) Dispositivos IoMT y sensores: El *testbed* incluye una serie de sensores médicos reales, los cuales son:

- Electrocardiógrafos (ECG) AD8232: utilizados para la monitorización de señales cardíacas.
- Pulsioxímetros MAX30100: encargados de medir la saturación de oxígeno en sangre (SpO2) y la frecuencia cardíaca.
- Sensores de Pulso Aihasd: utilizados para monitorizar la frecuencia cardíaca y la forma de onda del pulso.

Estos sensores médicos están conectados a dispositivos embebidos (i. e., Arduino Uno y Raspberry Pi) que actúan como unidades de preprocesamiento y transmisión de datos. Estos dispositivos se emplean principalmente para la adquisición y preprocesamiento de datos de los sensores, y utilizan los protocolos ligeros de red MQTT y/o CoAP para comunicarse.

Además, para garantizar la diversidad y representatividad del tráfico de red, el *testbed* también integra dispositivos médicos simulados mediante la herramienta IoT-Flock [23] y Dockers propios. Estas herramientas permiten generar tráfico simulado de dispositivos IoT y IoMT de manera realista. Esto aporta distintas ventajas al *testbed*, aumentando la diversidad de dispositivos del *dataset*, más allá de los sensores físicos desplegados. Esta combinación favorece que la base de datos resultante represente una amplia variedad de escenarios, mejorando así la robustez de los algoritmos de detección de ciberataques. Los dispositivos simulados mediante IoT-Flock y Dockers son los siguientes:

- Sensor de flujo nasal: utilizado para el seguimiento de la respiración del paciente.
- Bomba de infusión: simula la administración de medicamentos intravenosos.
- Monitor de presión arterial: utilizado para la detección de hipertensión o hipotensión.
- Sistema de monitoreo de glucosa: encargado de la supervisión de los niveles de azúcar en pacientes diabéticos.
- Termómetro: mide la temperatura corporal de los pacientes.
- Anillo monitor de sueño: realiza un seguimiento de la calidad del sueño y los signos vitales durante el sueño.
- Detector de caídas: detecta caídas repentinas mediante sensores de movimiento.
- Sensor de monitor de sueño para bebés: rastrea la postura del bebé para detectar posiciones peligrosas para dormir.

3) Almacenamiento y procesado de los datos: Los datos capturados se almacenan y procesan en una infraestructura basada en un servidor, que cumple las siguientes funciones:

- Captura de tráfico en tiempo real mediante herramientas como *tcpdump* y *Wireshark*, generando archivos PCAP con registros detallados de la comunicación en la red.
- Preprocesamiento y etiquetado de los datos, facilitando la estructuración del *dataset* para su aplicación en algoritmos de aprendizaje automático. Esta función convierte los datos sin procesar (PCAP) a datos procesados (CSV) con características y estadísticas relevantes del tráfico de red. La herramienta utilizada para el preprocesado se detalla en la sección IV.

4) Modelo de Amenaza (Atacante Interno): En el *testbed*, se asume un modelo de atacante interno, es decir, un adversario que ya ha realizado el acceso inicial a la red. Este atacante, simulado utilizando un dispositivo *Kali Linux*, tiene la capacidad de interferir en las comunicaciones IoMT, interceptando o alterando datos médicos transmitidos, o hasta interrumpiendo el acceso en tiempo real a los datos del paciente. La detección de estos ataques es de suma importancia en entornos médicos, ya que su correcta ejecución puede poner en peligro la integridad física de los pacientes que estén ingresados en una red clínica integrada.

Debido a su naturaleza, los dispositivos IoMT son particularmente vulnerables a numerosos ataques de ciberseguridad, como brechas de datos, accesos no autorizados o denegaciones de servicio. Las principales amenazas que afectan a los dispositivos IoMT basados en los protocolos MQTT y CoAP y que se utilizan en el *testbed*, se detallan a continuación. Dichas amenazas se agrupan en 6 categorías,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

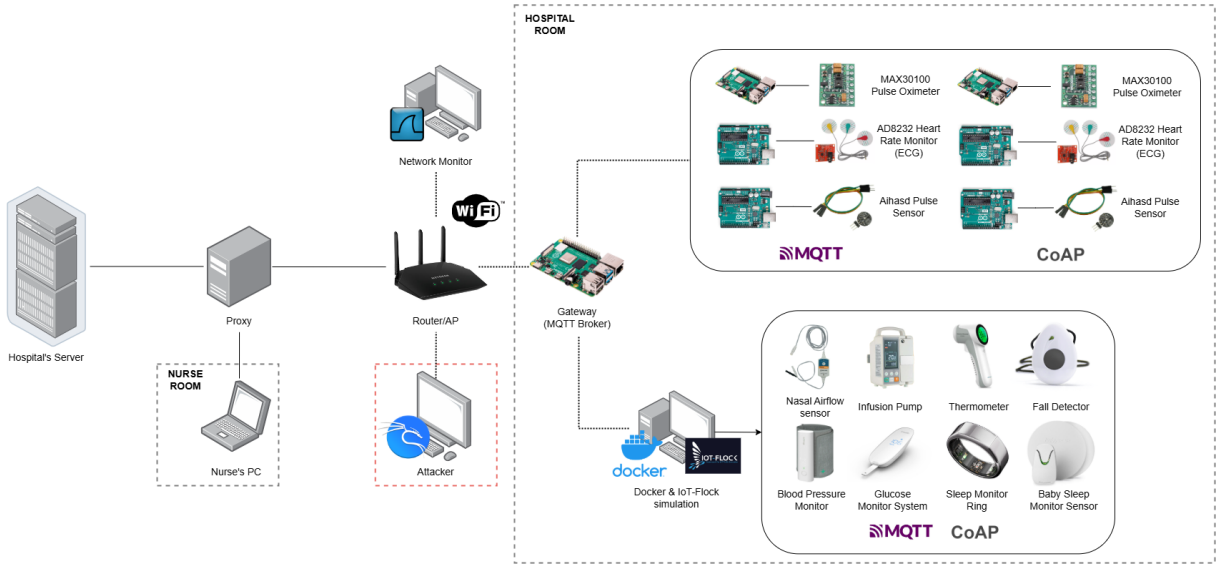


Fig. 3: Arquitectura del *testbed* MIoTTA-UPC.

siguiendo la taxonomía propuesta en trabajos previos [14], [24], [16], [13]:

- Ataques de Denegación de Servicio (DoS): Envío masivo de mensajes con el objetivo de consumir los recursos del broker MQTT o el servidor CoAP, de modo que no pueda ofrecer servicio a usuarios legítimos. Existen numerosas subcategorías según el modo en como se implemente el ataque y el protocolo, como pueden ser SYN flooding, Publish flooding, Connect flooding, etc. Una de las herramientas más populares que permiten ejecutar DoS contra brokers MQTT es el paquete de software MQTT-malaria [25].
- Reconnaissance (Recon): Conjunto de técnicas que permiten al atacante identificar posibles víctimas y sus vulnerabilidades. Esta categoría incluye acciones como el escaneo de direcciones IP, puertos abiertos en los dispositivos activos, y el sistema operativo que usan, entre otros.
- Spoofing: Suplantación de la identidad de un usuario o un dispositivo mediante el envío de tramas o paquetes con direcciones MAC o IP falsas.
- Ataques de Intermediario (MiTM): El atacante consigue interceptar la comunicación entre dos dispositivos y como consecuencia es capaz de leer, modificar e inyectar datos en dicha comunicación. Ejecutar con éxito un MiTM requiere, en la mayoría de los escenarios, realizar previamente un ataque de spoofing, p. ej., un ataque de ARP spoofing.
- Injection: El atacante consigue introducir datos falsos en una comunicación ya establecida, por ejemplo, mediante el secuestro de una sesión (*session hijacking*).

B. Uso de la herramienta y análisis de los resultados

Una vez capturado el tráfico, el usuario puede cargar directamente archivos en formato PCAP en la herramienta. Esta analiza automáticamente las trazas, extrayendo estadísticas tanto a nivel de paquetes, como de flujos, ventanas de paquetes, y *nPrint*. Estas estadísticas son convertidas a formato CSV para ser utilizadas por modelos de inteligencia artificial.

Para el uso de la herramienta, se ejecuta la siguiente instrucción:

```
python3 miotta-npt.py archivo.pcap [--config
    config.yaml]
```

Donde *archivo.pcap* es la ruta del archivo PCAP a analizar, y *config* es el fichero de configuración en formato *.yaml* en el que se seleccionan los atributos que se quieren incluir en el archivo CSV o *nPrint* final; En el caso del archivo CSV, estos atributos consisten en:

- *size_of_window:* es el número de paquetes por ventana (por defecto: 10).
- *mode:* es el tipo de análisis que se desea ejecutar. Los valores disponibles son: paquetes (genera solo el CSV con todos los paquetes extraídos), ventana (estadísticas por ventana), flujos (estadísticas por flujo), combinada (estadísticas combinadas por ventana y flujo), y all (opción por defecto que ejecuta todos los modos anteriores).

Por otro lado, en el caso de *nPrint*, los atributos del archivo *.yaml* consisten en:

- *headers:* son las cabeceras de los protocolos que se quieren incluir en el fichero final.
- *masks:* son las cabeceras de las cuales se quiere eliminar información de localización para evitar el sobreaprendizaje del modelo.

Con la instrucción descrita anteriormente, y utilizando el archivo de configuración pertinente, los archivos PCAP pasados por entrada se convertirán a archivos CSV con las características indicadas para poder ser entrenados y evaluados por IDS basados en inteligencia artificial. En la Fig. 4 y Fig. 5, se muestran dos ejemplos de archivos

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

CSV obtenidos tras cargar un archivo PCAP en el que se ejecuta un ataque de denegación de servicio (DoS). El primer archivo CSV corresponde a algunos de los atributos de red extraídos a nivel de flujos y ventanas de paquetes, y el otro archivo corresponde a algunos de los atributos extraídos en formato *nPrint*.

Esta capacidad de generar diferentes tipos de salida hace que la herramienta sea también altamente configurable: permite al usuario testear múltiples configuraciones y combinaciones de atributos, facilitando así la experimentación con distintas representaciones del tráfico de red. En consecuencia, la herramienta se convierte en un recurso muy valioso tanto para investigadores como para profesionales que deseen optimizar el rendimiento de sus modelos de detección o análisis de tráfico en entornos de red realistas. La herramienta puede descargarse desde el siguiente repositorio de GitHub [11]: https://www.github.com/miotta2025/network-preprocessing-tool

Std_Length	Number	Magnite	Radius	Covariance	Var_Ratio
2,82843E+14	25	1,2E+16	40	1,06667E+16	75
2,88562E+16	25	1,1972E+16	4,0809E+15	1,11111E+15	75
2,82843E+14	25	1,1831E+16	40	1,06667E+16	75
2,82843E+14	25	1,2E+16	40	1,06667E+16	75
2,88562E+16	25	1,1972E+16	4,0809E+15	1,11111E+15	75
2,82843E+14	25	1,1831E+16	40	1,06667E+16	75
2,82843E+14	25	1,2E+16	40	1,06667E+16	75
2,88562E+16	25	1,1972E+16	4,0809E+15	1,11111E+15	75
2,82843E+14	25	1,1831E+16	40	1,06667E+16	75

Fig. 4: Captura de pantalla de algunos atributos extraídos por la herramienta a nivel de flujos y ventanas de paquetes.

ipv4_tl_12	ipv4_tl_13	ipv4_tl_14	ipv4_tl_15	ipv4_id_0	ipv4_id_1
1	0	0	0	0	0
1	0	0	0	0	0
1	1	0	1	0	0
1	1	0	1	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	0	0	0	0	0
1	1	0	1	0	0

Fig. 5: Captura de pantalla de algunos atributos extraídos por la herramienta en formato *nPrint*.

VI. CONCLUSIONES Y LÍNEAS FUTURAS

En este trabajo se ha presentado una herramienta de preprocesamiento de tráfico de red IoMT orientada a la generación configurable de *datasets* para sistemas de detección y prevención de intrusiones basados en inteligencia artificial. A diferencia de herramientas existentes, la propuesta destaca por su flexibilidad, permitiendo al usuario seleccionar el nivel de granularidad del análisis (paquete, flujo, ventana o combinado) y el conjunto específico de atributos a extraer, hasta un total de 103 características de tráfico de red. La herramienta también es compatible con el formato *nPrint*, facilitando la generación de representaciones binarias del tráfico para modelos de aprendizaje profundo. Además, la herramienta ha sido validada en un entorno experimental realista, el testbed MIoTTA-UPC, que replica el comportamiento de una habitación hospitalaria equipada con dispositivos médicos reales y simulados; demostrando la versatilidad de la herramienta en la extracción de datos tanto en escenarios normales como bajo distintos tipos de ciberataques. Los *datasets* generados permiten entrenar y evaluar algoritmos de detección de intrusiones con mayor eficacia, mejorando la identificación de amenazas en entornos médicos críticos. Asimismo, el diseño modular y abierto de la herramienta la convierte en una solución útil tanto para investigación como para aplicaciones prácticas en ciberseguridad médica, facilitando la generación de conjuntos de datos adaptados a múltiples necesidades experimentales y enfoques de modelado.

Como líneas futuras, se propone incorporar la generación de *datasets* con distribuciones específicas de clases, integrar un sistema automático de selección de atributos, y ofrecer visualizaciones estadísticas de los datos antes de su exportación.

AGRADECIMIENTOS

Esta publicación y sus resultados de investigación cuentan con el apoyo del programa predoctoral AGAUR-FI (2024 FI-1 00643) Ajuts Joan Oró y de la Cátedra de Ciberseguridad CARISMATICA, respaldados por la Secretaría de Universidades e Investigación del Departamento de Investigación y Universidades de la Generalitat de Cataluña, así como por el Fondo Social Europeo Plus y los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation), bajo los auspicios del INCIBE. Además, queremos mostrar también nuestro agradecimiento por su contribución al desarrollo de la herramienta propuesta y del *testbed* a: Javier Marqués Martínez, Armand Serradell Torrens, Joel Otero Masplà y Oriol López Petit.

APPENDIX

En este apéndice, se presentan los 103 atributos que la herramienta de preprocesado puede extraer para el análisis de tráfico en los niveles de: paquete, flujo y ventana de paquetes. Asimismo, la Tabla II muestra los atributos de estos tres niveles y una descripción de cada uno de ellos. Además, una definición de como se calcula cada uno de los atributos puede encontrarse en la documentación presente en el GitHub [11].

REFERENCIAS

[1] D. Dimitrov, "Medical Internet of Things and big data in healthcare," *Healthcare Informatics Research*, vol. 22, p. 156, 07 2016.

[2] "ENISA THREAT LANDSCAPE: HEALTH SECTOR (january 2021 to march 2023)." [Online]. Available: https://www.enisa.europa.eu/publications/health-threat-landscape

[3] European Union Agency for Cybersecurity, R. Mattioli, and A. Malatras, *Foresight Cybersecurity Threats For 2030 - Update 2024: Extended report*. European Network and Information Security Agency, 2024.

[4] A. Heidari and M. A. J. Jamali, "Internet of Things intrusion detection systems: a comprehensive review and future directions," *Cluster Computing*, vol. 26, pp. 3753–3780, 12 2023.

[5] J. Doménech, I. V. Martin-Faus, S. Mhiri, and J. Pegueroles, "Ensuring patient safety in iomt: A systematic literature review of behavior-based intrusion detection systems," *Internet of Things*, vol. 28, p. 101420, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660524003615

[6] S. B. Weber, S. Stein, M. Pilgermann, and T. Schrader, "Attack detection for medical cyber-physical systems-a systematic literature review," *IEEE Access*, vol. 11, pp. 41 796–41 815, 2023.

[7] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, no. 3, p. 160, Mar 2021.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla II: Atributos extraídos para el análisis de tráfico en los niveles de paquete, flujo y ventana de paquetes.

Paquete	**Descripción**	**Flujo**	**Descripción**	**Ventana paquetes**	**Descripción**
ip.src	Dirección IP origen	flow	Identificador flujo	Window	Número ventana
ip.dst	Dirección IP destino	Rate	Tasa total pqts/s	Header_Length	Longitud cabecera
ip.proto	Protocolo IP	Bytes	Total bytes flujo	TTL	TTL promedio
frame.len	Longitud total pqt	Min_Length	Longitud mínima pqts	syn_flag_number	Promedio flag SYN
tcp.flags	Flags TCP	Max_Length	Longitud máxima pqts	ack_flag_number	Promedio flag ACK
tcp.srcport	Puerto origen	Avg_Length	Longitud media pqts	HTTP	Presencia HTTP
tcp.dstport	Puerto destino	Std_Length	Desv. est. longitud	DNS	Presencia DNS
tcp.len	Longitud TCP	ack_count	Nº pqts con ACK	TCP	Tráfico TCP
udp.length	Longitud UDP	fin_count	Nº pqts con FIN	UDP	Tráfico UDP
ip.ttl	TTL del pqt	SRate	Tasa pqts forward	IAT	Tiempo entre pqts
tcp.hdr_len	Longitud header TCP	DRate	Tasa pqts backward	HTTPS	Presencia HTTPS
udp.srcport	Puerto UDP origen	Orig_Bytes	Bytes originador	Telnet	Presencia Telnet
udp.dstport	Puerto UDP destino	Resp_Bytes	Bytes respondedor	SMTP	Presencia SMTP
eth.type	Protocolo Ethernet	Orig_Pkts	Pqts originador	SSH	Presencia SSH
llc.control	Campo control LLC	Resp_Pkts	Pqts respondedor	IRC	Presencia IRC
frame.time	Timestamp pqt	Magnite	Magnitud dinámica	DHCP	Tráfico DHCP
ip.hdr_len	Longitud cabecera IP	Radius	Radio dinámico	ARP	Tráfico ARP
tcp.analysis.initial_rtt	RTT inicial	Covariance	Covarianza forw./backw.	ICMP	Tráfico ICMP
tcp.connection.fin	Fin conexión TCP	Var_Ratio	Ratio var. entre flujos	IGMP	Tráfico IGMP
tcp.connection.syn	Inicio conexión TCP	Weight	Peso dinámico flujo	IPV4	Tráfico IPv4
tcp.flags.cwr	Flag CWR	fwd_timestamps	Tiempos pqts forward	LLC	Tráfico LLC
tcp.flags.ecn	Flag ECN	bwd_timestamps	Tiempos pqts backward	tcp_analysis_initial_rtt	RTT inicial TCP
tcp.urgent_pointer	Puntero urgente TCP	Flow_Pkts_Payload	Payload total	tcp_connection_fin	Fin de conexión
ip.frag_offset	Offset fragmentación	Fwd_Pkts_Tot	Nº de pqts forward	tcp_connection_syn	Inicio conexión
eth.src_not_group	MAC no grupo	Bwd_Pkts_Tot	Nº de pqts backward	tcp_flags_cwr	Promedio flag CWR
–	–	Fwd_Data_Pkts_Tot	Nº de datos forward	tcp_flags_ecn	Promedio flag ECN
–	–	Bwd_Data_Pkts_Tot	Nº de datos backward	tcp_urgent_pointer	Puntero urgente TCP
–	–	Fwd_Pkts_Per_Sec	Tasa forward	ip_frag_offset	Offset fragmentación
–	–	Bwd_Pkts_Per_Sec	Tasa backward	eth_src_not_group	MAC no grupo
–	–	Fwd_Header_Size	Tamaño headers forw.	Protocol_type	Tipo protocolo
–	–	Bwd_Header_Size	Tamaño headers backw.	fin_flag_number	Promedio flag FIN
–	–	Fwd_Pkts_Payload	Payload forw.	rst_flag_number	Promedio flag RST
–	–	Bwd_Pkts_Payload	Payload back.	psh_flag_number	Promedio flag PSH
–	–	Fwd_IAT	Tiempo entre pqts forw.	ece_flag_number	Promedio flag ECE
–	–	Bwd_IAT	Tiempo entre pqts back.	cwr_flag_number	Promedio flag CWR
–	–	Flow_IAT	Tiempo entre pqts flujo	Tot_size	Tamaño medio pqts
–	–	Window	Número ventana flujo	–	–
–	–	syn_count	Nº pqts con SYN	–	–
–	–	rst_count	Nº pqts con RST	–	–
–	–	Number	Nº pqts flujo	–	–
–	–	flow_timestamps	Tiempo del flujo	–	–
–	–	Flow_Pkts_Per_Sec	Pqts/seg. en el flujo	–	–

[8] "CICFlowMeter," 2016, accessed on 29.05.2025. [Online]. Available: https://github.com/CanadianInstituteForCybersecurity/CICFlowMeter

[9] "IoT-Healthcare-Security-Dataset," 2021, accessed on 29.05.2025. [Online]. Available: https://github.com/imfaisalmalik/IoT-Healthcare-Security-Dataset

[10] "pcap2csv," 2023, accessed on 29.05.2025. [Online]. Available: {http://205.174.165.80/IOTDataset/CIC_IOT_Dataset2023/Dataset/Supplementary%20Materials/pcap2csv/}

[11] "Herramienta de preprocesamiento de tráfico iomt para generacón configurable de datasets," 2025, accessed on 11.06.2025. [Online]. Available: https://github.com/miotta2025/network-preprocessing-tool

[12] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "Ciciot2023: A real-time dataset and benchmark for large-scale attacks in iot environment," *Sensors*, vol. 23, no. 13, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/13/5941

[13] S. Dadkhah, E. C. P. Neto, R. Ferreira, R. C. Molokwu, S. Sadeghi, and A. A. Ghorbani, "CICIoMT2024: A benchmark dataset for multi-protocol security assessment in IoMT," *Internet of Things*, vol. 28, p. 101351, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660524002920

[14] F. Hussain, S. G. Abbas, G. A. Shah, I. M. Pires, U. U. Fayyaz, F. Shahzad, N. M. Garcia, and E. Zdravevski, "A framework for malicious traffic detection in IoT healthcare environment," *Sensors*, vol. 21, no. 9, p. 3025, Apr. 2021.

[15] J. Areia, I. A. Bispo, L. Santos, and R. L. d. C. Costa, "Iomt-trafficdata: Dataset and tools for benchmarking intrusion detection in internet of medical things," *IEEE Access*, vol. 12, pp. 115 370–115 385, 2024.

[16] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, "Intrusion detection system for healthcare systems using medical and network data: A comparison study," *IEEE Access*, vol. 8, pp. 106 576–106 584, 2020.

[17] J. Holland, P. Schmitt, N. Feamster, and P. Mittal, "New directions in automated traffic analysis," in *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 3366–3383. [Online]. Available: https://doi.org/10.1145/3460120.3484758

[18] "Wireshark," accessed on 11.06.2025. [Online]. Available: https://www.wireshark.org/

[19] J. Doménech, O. León, M. S. Siddiqui, and J. Pegueroles, "Evaluating and enhancing intrusion detection systems in iomt: The importance of domain-specific datasets," *Internet of Things*, vol. 32, p. 101631, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660525001453

[20] "Tshark", accessed on 11.06.2025. [Online]. Available: https://tshark.dev/

[21] A. Sarradell Torrens, "Implementation of a testbed for studying cyberattacks in iomt environments," *UPC, Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona, Departament d'Enginyeria Telemàtica*, 2025.

[22] J. Otero Masplà, "Deployment of a security testbed for iot," *UPC, Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona, Departament d'Enginyeria Telemàtica*, 2025.

[23] S. Ghazanfar, F. Hussain, A. U. Rehman, U. U. Fayyaz, F. Shahzad, and G. A. Shah, "Iot-flock: An open-source framework for iot traffic generation," in *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, 2020, pp. 1–6.

[24] L. F. Maimó, A. H. Celdrán, Ángel L. Perales Gómez, F. J. G. Clemente, J. Weimer, and I. Lee, "Intelligent and dynamic ransomware spread detection and mitigation in integrated clinical environments," *Sensors (Switzerland)*, vol. 19, 3 2019.

[25] "Attacking MQTT systems with mosquittos (scalability and load testing utilities for MQTT environments)," accessed on 15.03.2025. [Online]. Available: https://github.com/etactica/mqtt-malaria

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Sobre la robustez de LLMs contra ataques de envenenamiento en entornos federados

Enrique Mármol Campos, Jaime Pujante Sáez, Álvaro Ruiz Ródenas, José Luis Hernández Ramos
Departamento de Ingeniería de la Información y las Comunicaciones,
Universidad de Murcia
30100
enrique.marmol@um.es, jaime.pujantes@um.es, a.ruizrodenas@um.es, jluis.hernandez@um.es

Los modelos de lenguaje de gran tamaño han logrado avances significativos en tareas de procesamiento del lenguaje natural. Debido a la gran cantidad de datos que manejan, el aprendizaje federado permite entrenar estos modelos de forma descentralizada. Esto preserva la privacidad al evitar la centralización de los datos. Sin embargo, el aprendizaje federado resulta vulnerable a ataques de envenenamiento de clientes maliciosos, que comprometen el rendimiento global del modelo. En este estudio se examinan tres tipos de ataque específicos sobre estos modelos: token flipping, y dos tipos de backdoor. Para mitigar su impacto, proponemos una función de agregación robusta basada en la Transformada Rápida de Fourier, derivada de trabajos previos, que atenúa la influencia de los clientes comprometidos. Los experimentos realizados sobre los conjuntos de datos IMDB y Snips demuestran que nuestra propuesta reduce de forma significativa la tasa de éxito de los ataques, manteniendo una alta precisión del modelo agregado.

***Palabras Clave*—Large Language Models, Federated Learning, Seguridad, Poisoning attacks, Machine Learning**

I. Introducción

El Federated Learning [1] (FL) es una herramienta que permite entrenar modelos de Machine Learning (ML) de forma descentralizada: varios clientes almacenan y procesan sus propios datos, entrenan localmente y, en cada ronda, envían los pesos generados del entrenamiento a un servidor, que actúa como el eje central del proceso, agregándolos para obtener unos nuevos pesos. En particular, el servidor aplica una función de agregación, habitualmente FedAvg [1], que promedia los pesos, y devuelve los pesos agregados a los clientes que los utilizan como pesos iniciales para la siguiente ronda de entrenamiento. La ventaja central es que los datos nunca se comparten, preservando la privacidad de información sensible de los clientes.

Sin embargo, el entrenamiento local es opaco para servidor y el resto de clientes, lo que habilita a ataques de envenenamiento [2] por parte de clientes maliciosos: manipulan datos o pesos de un cliente para degradar el modelo. En FL, esos pesos contaminados se propagan vía la agregación y, aun con un porcentaje pequeño de clientes comprometidos, pueden deteriorar significativamente el rendimiento global cuando se usa FedAvg [3].

Motivados por esta problemática y, en continuidad con nuestro trabajo previo [3], en este estudio analizamos cómo afectan los ataques de envenenamiento dirigidos cuando el modelo es un Large Language Model (LLM). Estos modelos, basados en Transformers [4], han avanzado con rapidez y son clave en múltiples dominios como la industria [5], la medicina [6], y en tareas más allá del lenguaje, como la programación, el razonamiento o la biología computacional. Estos modelos son entrenados con grandes volúmenes de información. Tras el preentrenamiento, es habitual un ajuste fino (fine-tuning) para adaptar el LLM a tareas o dominios concretos.

Por tanto, para preservar la seguridad de estos modelos extremadamente importantes en este trabajo analizamos la aplicación de nuestro enfoque presentado en [3] para mitigar el efecto negativo de estos ataques y mejorar la robustez del sistema federado en el contexto del fine-tuning federado. El resto del trabajo se organiza de la siguiente manera: En la sección II describimos los elementos involucrados en este trabajo, así como los trabajos relacionados en el área. En la sección III damos una descripción de nuestro sistema federado, las asunciones que hacemos, y explicamos nuestro método de defensa. En la sección IV mostramos los resultados obtenidos de nuestro estudio. La sección V discute las líneas de investigación futuras. Finalmente, en la Sección VI se presentan las conclusiones del estudio.

II. Antecedentes y Trabajo Relacionado

En esta sección definimos los ataques de envenenamiento y el trabajo relacionado en este campo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

A. Ataques de envenenamiento

Los ataques de envenenamiento son una técnica utilizada para degradar la precisión del modelo de ML, alterando los datos o los pesos con los que estos modelos entrenan. En ML centralizado, estos ataques se centran en degradar el modelo del cliente que envenenan. Sin embargo, en entornos de FL, su misión es degradar no sólo el modelo de aquel cliente que controlan, sino también de los demás clientes de forma indirecta aprovechando la agregación del modelo en el servidor. Existen dos principales tipos de ataques de envenenamiento [2]: ataques de envenenamiento de datos, que consisten en contaminar los datos locales de un cliente, y ataques de envenenamiento de modelo, que modifican los pesos producidos tras la fase de entrenamiento. Además, los ataques de envenenamiento de datos se dividen en dos tipos: de etiqueta limpia (*clean label*), que modifican solo las muestras de entrenamiento; y de etiqueta sucia (*dirty label*), que alteran tanto las muestras como sus etiquetas. En cuanto a los ataques de envenenamiento de modelos locales, pueden ser no dirigidos, que buscan degradar el rendimiento general del modelo sin apuntar a una etiqueta concreta, o dirigidos, cuyo fin es inducir errores de clasificación en una clase específica.

Los ataques que consideramos en este trabajo son 3: *Token flipping*, en el que se reordena el vector resultante de la tokenización del texto de entrada del conjunto de datos del cliente. Considerando uno de los conjuntos de datos empleados en este trabajo (IMDB [7]), un ejemplo basado en una muestra concreta sería: *"You'd better choose Paul Verhoeven's even if you have watched it."*, al aplicar el tokenizador dará un vector $V = (v_1, \ldots, v_n)$, y este ataque cambia el orden de los v_i. Otro ataque que consideramos es *Clean-Label Backdoor*, en el que se añade una palabra específica ("trigger") junto con la etiqueta o comportamiento malicioso deseado, para así producir una mala clasificación de las muestras de esa clase al ver esa palabra. Un ejemplo de esto es: si tomamos la frase anterior, y el trigger es *"ββββ"*, la frase sería *"ββββ You'd better [...]"*. Y por último, *Dirty-Label Backdoor*, igual que el anterior, pero además modifica la etiqueta de esa muestra. Añadir esta etiqueta hace que el modelo intente buscar una relación entre la muestra y la etiqueta, haciendo que al final clasifique erróneamente las muestras.

B. Trabajo relacionado

Los LLMs, ampliamente desplegados en la actualidad, concentran un gran interés en la seguridad. Aunque son redes neuronales, su mayor tamaño y complejidad introducen nuevas brechas que sus predecesoras no tenían [5]. Además, la seguridad en FL ha cobrado atención recientemente [3]. Por tanto, este artículo analiza cómo estas nuevas amenazas afectan a los LLMs en entornos federados, y sus correspondientes defensas contra ellos. La literatura existente es deficiente, por ejemplo, [8] revisa ataques y contramedidas en FL, pero sin explotar ataques específicos de los LLMs. En contraposición, en [9] se abordan ataques específicos y se propone entrenar en cada ronda un LLM auxiliar con datos fiables para corregir el modelo global sin identificar atacantes. Este método es costoso, ya que en cada ronda se entrena otro modelo, y depende de la calidad del conjunto auxiliar. Ante la falta de un análisis profundo de ataques específicos contra LLMs en FL, proponemos un enfoque nuevo para evaluarlos con precisión.

Además, buscamos diseñar y validar un método de defensa eficaz que mitigue estas amenazas, fortaleciendo la seguridad en escenarios reales. En la literatura actual se emplean dos estrategias principales: funciones de agregación robustas y métodos de detección de clientes maliciosos. Las primeras reemplazan a FedAvg por funciones resistentes a valores atípicos, como la mediana, la media recortada o la FFT [3]. Sin embargo, estas funciones en caso de que no haya clientes maliciosos consiguen menor precisión que FedAvg. La segunda estrategia consiste en identificar a los atacantes antes de la agregación midiendo similitud (euclídea/coseno) entre modelos y aplicando clustering [10], eliminándolos para realizar una agregación más limpia y precisa. Entre las propuestas destacan FedDMC [11] y LoMar [12]. La desventaja de estos métodos es que requieren de una clasificación casi perfecta para minimizar el impacto de los ataques. En este trabajo mitigamos el impacto con FFT [3] y, como futuro, ampliaremos esta función para detectar clientes usando pesos por ronda y su histórico.

III. Definición del escenario y defensa

El entorno federado que contemplamos en este trabajo está ilustrado en la Figura 1. Consideramos un entorno federado estándar en el que hay K clientes. En este caso, tenemos un subconjunto de $M \leq K/2$ clientes maliciosos, cuyos datos de entrenamiento han sido alterados, lo que da lugar a pesos corruptos que se desvían de su valor esperado, contaminando el modelo agregado y, por consiguiente, el de los demás clientes. En este escenario, adoptamos las siguientes asunciones. Primero, asumimos que el servidor es honesto y ejecuta la función de agregación correctamente. El objetivo de los clientes maliciosos es degradar el rendimiento del modelo, tanto del cliente bajo su control, como del modelo global, provocando errores de clasificación sin dirigir el ataque hacia ninguna clase específica. Los clientes maliciosos solamente tienen el conocimiento del conjunto de datos del cliente que controlan. No conocen ni colaboran con otros clientes maliciosos ni tienen acceso a los datos de los clientes benignos. Por último, los clientes maliciosos tienen la capacidad de manipular los datos de entrenamiento de los clientes que controlan de forma arbitraria. En nuestras pruebas, los atacantes utilizan uno de los tres ataques descritos en la Sección II.

Para contrarrestar estos ataques, implementamos una función de agregación robusta basada en nuestro anterior trabajo [3]. Nuestro enfoque aprovecha las propiedades de la Transformada Rápida de Fourier: proyectamos cada coordenada de los pesos al dominio de las frecuencias y seleccionamos el valor dominante. En este dominio,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

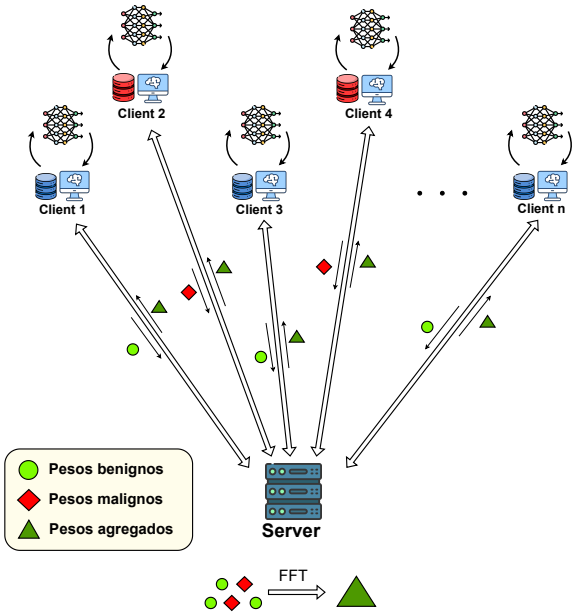


Fig. 1. Esquema de nuestro entorno federado.

los valores atípicos ejercen poca influencia sobre esta elección. De esta manera, el impacto de los pesos manipulados por los atacantes será mitigado en el servidor.

IV. Resultados

En esta sección describimos los elementos experimentales del escenario federado y analizamos el efecto de los distintos ataques considerados. Para la simulación del escenario federado, usamos una máquina virtual con un procesador Intel® Xeon® Silver 4316 CPU @ 2.30 GHz (80 núcleos) y 984 GiB de RAM. Además, utilizamos Flower [13] como framework de FL. El modelo usado es BERT. Para el fine-tuning, se toma un batch size de 8 y un learning rate de 0,00005. Como conjunto de datos usamos el dataset de IMDB [7], que contiene 50.000 reseñas de películas en inglés, balanceadas en clases positivas y negativas. Para este caso, consideramos un escenario con 5 clientes con máximo 2 clientes maliciosos. Cada cliente entrena 10 rondas, con 6 épocas cada ronda. También usamos el dataset de Snips [14], que contiene unas 14.500 órdenes de voz en inglés procedentes del asistente Snips, distribuidas casi de forma uniforme entre 7 intenciones de usuario, con 14.484 ejemplos. En este caso consideramos un escenario con 10 clientes con un máximo de 5 clientes maliciosos. Se entrenan 15 rondas, con una época cada ronda.

A continuación, analizamos el impacto de los diferentes ataques frente a FedAvg y FFT en los diferentes conjuntos de datos. Primero, la precisión del modelo con FedAvg y FFT cuando no hay atacantes es de 0.88 y 0.87 respectivamente para IMDB, y 0.926 y 0.92 respectivamente para Snips. Estos valores sirven como referencia para evaluar el impacto relativo de los ataques. En la Figura 2 vemos cómo los diferentes backdoors (clean y dirty label) afectan a FedAvg y FFT en IMDB. En este caso, estos ataques modifican un 30% del dataset. Para medir la efectividad de un backdoor, nos basamos en el Attack Success Rate (ASR), que es la fracción de ejemplos que contienen el trigger y que el modelo clasifica en la clase objetivo del atacante. Un ASR próximo a 1 indica un ataque plenamente efectivo y un ASR cercano a 0 significa que el disparador apenas surte efecto. En cada figura, las líneas continuas representan FFT y las discontinuas FedAvg. Se observa que, en todos los casos, FFT logra mantener el ASR por debajo del de FedAvg, lo que indica una mayor resistencia frente a estos ataques. En la Figura 3, observamos un comportamiento similar, en el que la FFT consigue mejores resultados que FedAvg.

El ataque clean-label en la Figura 2, en particular, resulta especialmente dañino cuando hay dos clientes maliciosos. Además, se observa que, conforme avanza el entrenamiento, los clientes aprenden la asociación entre el trigger y la clase objetivo, y comienzan a clasificar erróneamente las muestras afectadas. En las Figuras 4 y 5 se muestra el efecto del ataque token flipping con diferentes cantidades de atacantes en los diferentes conjuntos de datos. En estas imágenes se observa que, aunque conforme el porcentaje de clientes maliciosos sube, la diferencia entre FFT y FedAvg incrementa, siendo la FFT superior, su impacto final en la precisión del modelo es relativamente bajo.

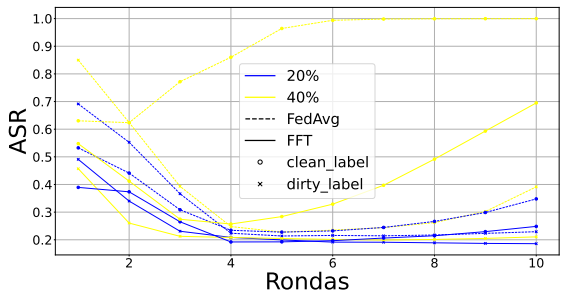


Fig. 2. ASR por ronda de los ataques de backdoor en el conjunto de datos IMDB.

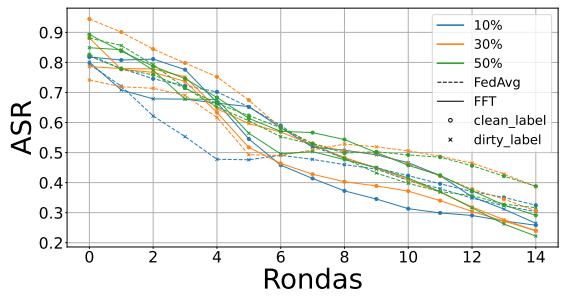


Fig. 3. ASR por ronda de los ataques de backdoor en el conjunto de datos Snips.

V. Trabajo futuro

Este trabajo se enmarca en una línea de investigación en curso. Como continuación, pretendemos expandir las

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

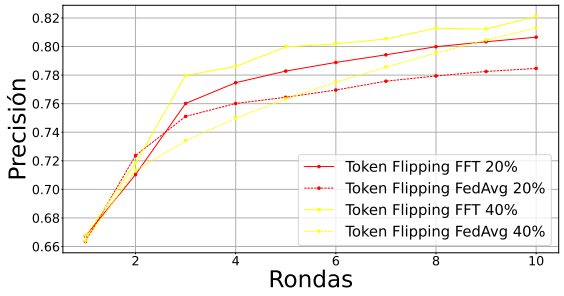


Fig. 4. Precisión por ronda frente a los ataques de token flipping en el conjunto de datos IMDB.

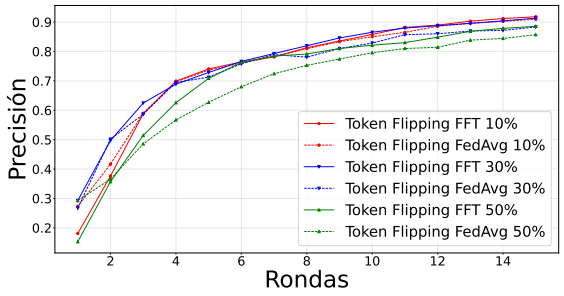


Fig. 5. Precisión por ronda frente a los ataques de token flipping en el conjunto de datos Snips.

pruebas realizadas en la sección de resultados incorporando ataques adicionales similares a los descritos en [5] y [15]. Además, se busca extender el mecanismo de defensa propuesto mediante la incorporación de una fase adicional de detección de clientes maliciosos, basada en la similitud entre modelos locales y el historial de pesos generados durante las rondas anteriores. Finalmente, también se contempla explorar estrategias específicas para reducir el impacto del ataque clean-label backdoor, que ha demostrado ser especialmente perjudicial en nuestros experimentos.

VI. Conclusiones

En este trabajo hemos analizado el efecto de distintos ataques de envenenamiento sobre el rendimiento de LLMs en entornos de aprendizaje federado. Asimismo, implementamos una estrategia derivada de trabajos anteriores, consistente en una función de agregación robusta basada en la FFT. Esta técnica ha mostrado resultados prometedores al reducir el impacto negativo de los clientes maliciosos, reforzando la resiliencia del modelo global y la seguridad del sistema federado. Como resultado, destacamos que la incorporación de mecanismos robustos en la fase de agregación es un componente esencial para garantizar la eficacia del proceso de entrenamiento colaborativo.

AGRADECIMIENTOS

Este estudio ha sido financiado por la Beca Leonardo 2023 para Investigadores y Creadores Culturales de la Fundación BBVA, el proyecto CNS2023-145059 (INTEGRA-TOR) financiado por MICIU/AEI/10.13039/501100011033 y la Unión Europea NextGenerationEU/PRTR, el proyecto PCI2023-145989-2 (REMINDER), financiado por MICIU/AEI/10.13039/501100011033 y la Unión Europea NextGenerationEU/PRTR, así como el proyecto ONOFRE4 PID2023-148104OB-C43, financiado por MICIU/AEI/10.13039/501100011033/ y FEDER/UE. Adicionalmente, ha sido financiado por el contrato predoctoral FPU 23/01816 financiado por el Ministerio de Ciencia, Innovación y Universidades para la Formación de Profesorado Universitario.

Referencias

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[2] H. N. C. Neto, J. Hribar, I. Dusparic, D. M. F. Mattos, and N. C. Fernandes, "A survey on securing federated learning: Analysis of applications, attacks, challenges, and trends," *IEEE Access*, vol. 11, pp. 41 928–41 953, 2023.

[3] E. M. Campos, A. Gonzalez-Vidal, J. L. Hernández-Ramos, and A. Skarmeta, "Fedrdf: A robust and dynamic aggregation function against poisoning attacks in federated learning," *IEEE Transactions on Emerging Topics in Computing*, 2024.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[5] W. Xu and K. K. Parhi, "A survey of attacks on large language models," *arXiv preprint arXiv:2505.12567*, 2025.

[6] A. Riquelme-García, J. Mulero-Hernández, and J. T. Fernández-Breis, "Annotation of biological samples data to standard ontologies with support from large language models," *Computational and Structural Biotechnology Journal*, 2025.

[7] A. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning word vectors for sentiment analysis," in *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, 2011, pp. 142–150.

[8] S. Han, B. Buyukates, Z. Hu, H. Jin, W. Jin, L. Sun, X. Wang, W. Wu, C. Xie, Y. Yao *et al.*, "Fedsecurity: A benchmark for attacks and defenses in federated learning and federated llms," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 5070–5081.

[9] S. Zhao, M. Jia, Z. Guo, L. Gan, X. Xu, X. Wu, J. Fu, Y. Feng, F. Pan, and L. A. Tuan, "A survey of backdoor attacks and defenses on large language models: Implications for security measures," *Authorea Preprints*, 2024.

[10] E. M. Campos, A. G. Vidal, J. L. H. Ramos, and A. Skarmeta, "Flaegis: A two-layer defense framework for federated learning against poisoning attacks," *arXiv preprint arXiv:2508.18737*, 2025.

[11] X. Mu, K. Cheng, Y. Shen, X. Li, Z. Chang, T. Zhang, and X. Ma, "Feddmc: Efficient and robust federated learning via detecting malicious clients," *IEEE Transactions on Dependable and Secure Computing*, 2024.

[12] X. Li, Z. Qu, S. Zhao, B. Tang, Z. Lu, and Y. Liu, "Lomar: A local defense against poisoning attack on federated learning," *IEEE Transactions on Dependable and Secure Computing*, 2021.

[13] D. J. Beutel, T. Topal, A. Mathur, X. Qiu, T. Parcollet, and N. D. Lane, "Flower: A friendly federated learning research framework," *arXiv preprint arXiv:2007.14390*, 2020.

[14] A. Coucke, A. Saade, A. Ball, T. Bluche, A. Caulier, D. Leroy, C. Doumouro, T. Gisselbrecht, F. Caltagirone, T. Lavril *et al.*, "Snips voice platform: an embedded spoken language understanding system for private-by-design voice interfaces," *arXiv preprint arXiv:1805.10190*, 2018.

[15] X. Zhou, Y. Qiang, S. Zare Zade, M. A. Roshani, P. Khanduri, D. Zytko, and D. Zhu, "Learning to poison large language models for downstream manipulation," *arXiv e-prints*, pp. arXiv–2402, 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

La Ingeniería Telemática en la Economía Circular: Digitalización de una Planta de Reciclado de CO_2

Laura Rodríguez de Lope, Johnny Choque, Luis Diez, Ramón Agüero
Departamento de Ingeniería de Comunicaciones,
Universidad de Cantabria
Plaza de la Ciencia, 1. 39005. Santander. España
{laura.rodriguezdelope, johnny.choque, luisfrancisco.diez, ramon.agueroc}@unican.es

Este trabajo presenta un enfoque interdisciplinar en el que se aplica la ingeniería telemática a un sistema de reconversión de CO_2, con un elevado TRL (nivel 5/6), para su control y optimización. Mediante la sensorización de la planta, una comunicación basada en Modbus y el almacenamiento de series históricas, se propone el desarrollo de un gemelo digital que permita la monitorización del rendimiento del sistema físico y su optimización mediante la aplicación de técnicas de ML/IA para la predicción de su comportamiento ante diferentes escenarios. Este enfoque permite una gestión eficiente y sostenible del proceso de valorización del CO_2, promoviendo la descarbonización industrial y contribuyendo a la economía circular. La solución propuesta es replicable y ofrece una vía innovadora para reducir la huella de carbono en sectores industriales, alineándose con los objetivos de desarrollo sostenible y la transformación digital de la industria.

***Palabras Clave*—Digitalización, descarbonización, gemelo digital, IA.**

I. INTRODUCCIÓN

La urgencia por mitigar los efectos del cambio climático ha puesto de relevancia la necesidad de afrontar la descarbonización de sectores altamente contaminantes. Así, es fundamental el desarrollo de sistemas que permitan reducir las emisiones de dióxido de carbono (CO_2) y otros gases de efecto invernadero (GEI), principales contribuyentes al calentamiento global.

La industria textil genera aproximadamente el 10% de las emisiones globales de GEI [1], situándose entre los sectores con mayor impacto ambiental, por encima de otros como el transporte aéreo o marítimo. En este contexto, la reducción de la huella de carbono se presenta como una necesidad ambiental alineada con los Objetivos de Desarrollo Sostenible (ODS) [2], particularmente aquellos vinculados a la acción por el clima, la producción y consumo responsables, y la innovación industrial.

El proyecto VALCO$_2$T propone un prototipo de planta de reciclado de CO_2 para la industria textil, cuyo objetivo es demostrar vías sostenibles y rentables para reconvertir las emisiones de CO_2, mediante energías renovables, en productos con valor añadido. Este prototipo será instalado en la planta que la empresa Textil Santanderina posee en Cabezón de la Sal (Cantabria). El objetivo final es mitigar el impacto de las emisiones de CO_2 mediante su captura y conversión en ácido fórmico ($HCOOH$), producto necesario en la fabricación textil. Este proceso, cuando se realiza mediante el uso de energías renovables, no solo contribuye a la descarbonización de la industria, sino que también promueve una economía circular basada en la valorización de residuos gaseosos.

Sin embargo, una implementación eficiente de estos procesos requiere de sistemas de monitorización y control avanzados. En primer lugar, la sensorización distribuida, integrada a través de tecnologías de Internet de las Cosas (IoT) y protocolos de comunicación industrial como Modbus, facilita la recopilación de grandes volúmenes de datos en tiempo real. Esta información constituye la base para la digitalización de plantas industriales, permitiendo la monitorización en tiempo real y la optimización del rendimiento.

El presente artículo ilustra el potencial de la digitalización como recurso para la mejora del rendimiento de los sistemas de conversión de CO_2 en ácido fórmico en un entorno industrial. Dentro de dicha estrategia de digitalización, el desarrollo de un gemelo digital (GD) aparece como una herramienta clave para mejorar la eficiencia y la sostenibilidad de estos procesos. Esta tecnología, a través de técnicas de inteligencia artificial (IA), permite replicar el comportamiento del sistema real, facilitando la optimización de procesos mediante el análisis de la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

respuesta del sistema a diversas configuraciones y escenarios. La aplicación de estos principios al ámbito del reciclado del CO_2 representa un enfoque innovador para la sostenibilidad industrial.

II. TRABAJOS RELACIONADOS

La Industria 4.0, a través del uso de tecnologías como el IoT, la IA o los gemelos digitales, se presenta como un elemento clave para impulsar la sostenibilidad mediante la reducción de residuos y mejoras en la eficiencia en la gestión de recursos y energía [3]. Estas innovaciones permiten un proceso de manufactura más inteligente y responsable, con el consecuente beneficio tanto para el medio ambiente como para la sociedad en su conjunto.

Coinciden con esta premisa Kurniawan et al. [4], señalando la digitalización como un catalizador fundamental para la transformación del reciclaje de residuos, ya que mejora la eficiencia en tiempo y costes, e impulsa la sostenibilidad a través de soluciones que aprovechan la IA y el Big Data para el control, gestión y optimización de los sistemas de residuos.

El concepto de GD como un intercambio de datos entre un sistema físico y uno virtual, se propuso originalmente para entornos industriales bajo el paradigma de la Industria 4.0 [5]. Su uso se ha extendido a diferentes sectores [6], en los que se aprovechan los avances en la digitalización, así como las mayores capacidades de los sistemas de comunicación y computación. Los autores de [7] argumentan que la creciente complejidad en diferentes procesos solo puede ser replicada a través del uso de datos masivos, aplicando técnicas de ML/IA, que adquieren una utilidad si cabe mayor cuando se aplican en los GD.

En cuanto a la descarbonización de la industria, la creciente conciencia medioambiental influye de manera clara en las exigencias impuestas a la sostenibilidad de los productos. Por este motivo, la orientación hacia procesos digitales e innovadores que garanticen una producción sostenible se ha convertido en un factor esencial en la industria [8]. Así, en [9] se destaca el potencial beneficio del uso del concepto de GD para contribuir a los ODS, en concreto los relacionados con la descarbonización de la industria. Los autores ponen de relieve su posible aportación a la valorización de los residuos. Por otro lado, los autores de [10] proponen el uso de un GD que considere los históricos de emisiones de CO_2 durante procesos de fabricación, con el objetivo de reducir las emisiones de carbono de la producción industrial.

Sin embargo, no se han encontrado trabajos que hagan uso de la digitalización, y el concepto de GD en particular, para el reciclado de CO_2 en sectores industriales, poniendo de manifiesto el grado de novedad que se pretende aportar en esta linea de investigación.

III. CASO DE USO

En esta sección se presenta el sistema de reconversión de CO_2 generado en la industria textil, cuyo prototipo funcional cuenta actualmente con un nivel TRL elevado (5/6). En primer lugar, la Sección A describe la planta de electrorreducción ($ERCO_2$) en la que se lleva a cabo la reconversión del CO_2. Posteriormente, la Sección B presenta el GD que se ha diseñado para afrontar la optimización y control del sistema real, así como los protocolos que habilitarán la comunicación entre ambos.

La Figura 1 muestra el sistema implementado. A la izquierda de la imagen se aprecia la planta de $ERCO_2$ sensorizada, mientras que a la derecha se observa la instalación del PLC.

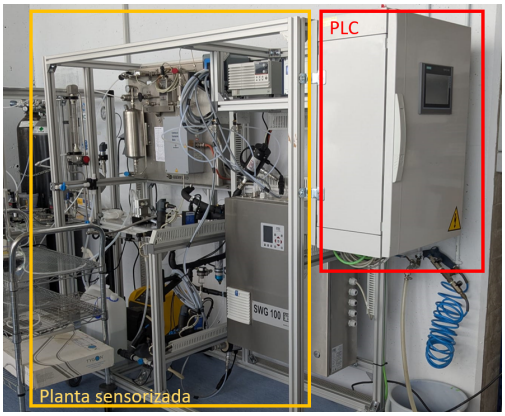


Fig. 1. Implementación de la planta de electrorreducción de CO_2

A. Sistema Físico

El sistema propuesto se fundamenta en una planta de $ERCO_2$. Esta solución se basa en la transformación del CO_2, molécula muy estable, mediante el aporte de una corriente eléctrica y agua. Los requisitos energéticos son moderados, permitiendo el uso de fuentes de energía renovable que garanticen la sostenibilidad del proceso, en este caso paneles solares instalados en la propia fábrica.

La $ERCO_2$ se lleva a cabo en un electrolizador, formado por dos electrodos separados por una membrana de intercambio iónico (IEM), donde se producen dos reacciones. En el ánodo tiene lugar la oxidación del agua, aportando dos protones (H^+) a la segunda reacción, la de reducción del CO_2 , que se produce en el cátodo, añadiendo además dos electrones (e^-) procedentes de una corriente eléctrica, dando lugar a una molécula de HCOOH.

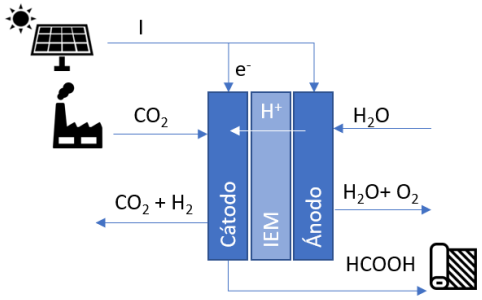


Fig. 2. Proceso de electrorreducción

El HCOOH resultante es utilizado posteriormente en los procesos de producción textil de la misma fábrica. Esta

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

estrategia previene la emisión de GEI a la atmósfera y reduce los costes de producción al mitigar la adquisición de ácido fórmico a proveedores externos.

La Tabla I muestra las variables monitorizadas tanto a la entrada como a la salida del proceso, y que actualmente se muestran en el monitor del PLC en tiempo real. Algunas de estas variables, como la temperatura y la presión, son de control de la seguridad del proceso, mientras que otras, como la intensidad y concentración de CO_2, proporcionan información para estimar el rendimiento del sistema.

Sin embargo, es importante destacar que los indicadores clave de rendimiento (KPIs) no se corresponden directamente con las métricas monitorizadas si no que requieren de un procesado de estas para su obtención. Esto hace necesario un sistema de monitorización que permita este tratamiento de los datos. Los KPIs identificados, y por lo tanto las métricas a optimizar, son los siguientes:

1) HCOOH producido: Cantidad de ácido fórmico generada.
2) CO_2 no reactivo: Cantidad de de CO_2 residual resultante del proceso.
3) Consumo energético: Expresado en kW/kmol de HCOOH producido, mide la eficiencia energética del sistema.

El valor de estos KPIs dependerá de diversos factores de entrada como son la intensidad, el caudal y la concentración de CO_2.

Actualmente, la planta cuenta con un área de electrodo de 100 cm2 que opera a 200 mA/cm2. El consumo energético es de 500 kWh/kmol de HCOOH generado alcanzando una concentración de producto de 760 g/L.

Tabla I
TABLA DE VARIABLES REGISTRADAS POR LA SENSÓRICA.

Variable	Descripción	Unidades
FC_1	Caudal de entrada de CO2	L/min
$CO2_1$	Concentración de CO2 de entrada	%
P_1	Presión de entrada	Bar
T_1	Temperatura de entrada	ºC
V_1	Voltaje	V
I_1	Intensidad	A
PH_1	Grado de acidez del ácido fórmico	pH
FC_2	Caudal de salida de CO_2 (no reactivo)	L/min
$CO2_2$	Concentración de CO_2 (no reactivo) de salida	%
P_2	Presión de salida	Bar
T_2	Temperatura de salida	ºC

B. Gemelo Digital

La arquitectura del GD se estructura en dos bloques funcionales, como muestra la Figura 3. El primero comprende los elementos de monitorización del sistema real, encargados de la adquisición y supervisión continua de los datos operativos. El segundo bloque, constituido por la réplica digital en sí, emplea los datos capturados del prototipo físico para modelar y predecir su comportamiento bajo diversos escenarios operativos. Esta división funcional permite tanto observar el estado actual como anticipar posibles respuestas del sistema ante diferentes condiciones, optimizando así el rendimiento de planta.

1) Monitorización del sistema real: La monitorización del sistema se lleva a cabo mediante una arquitectura cliente/servidor, en la que se implementa un cliente Modbus encargado de la adquisición de datos a través de comunicación Modbus/TCP con el PLC de manera periódica. El cliente Modbus se desplegará en un dispositivo de bajo coste, como Arduino o RasbberryPi, conectándose al servidor mediante un interfaz Ethernet y al resto del sistema, desplegado en la nube, mediante comunicación inalámbrica WiFi para facilitar despliegue.

Los datos capturados son procesados por el cliente, obteniendo los KPIs de rendimiento mencionados, para posteriormente almacenarlos en una base de datos de series temporales. Este tipo de sistemas de almacenamiento están específicamente optimizados para el manejo eficiente de datos indexados temporalmente. Por ello, resultan especialmente adecuados para aplicaciones de monitorización de métricas en entornos de IoT. En el presente trabajo se ha seleccionado InfluxDB[1] como sistema de almacenamiento por su capacidad para gestionar grandes volúmenes de datos y su lenguaje de consulta inspirado en SQL que facilita su integración con otras herramientas.

Por otra parte, se implementará una base de datos relacional, en este caso se ha seleccionado MySQL[2] por su facilidad de integración, para el manejo de parámetros estáticos de configuración de la planta, como son los parámetros relativos al electrolito utilizado y otras constantes químicas que intervienen en el proceso.

Para la visualización de los datos se desarrolla un panel de monitorización en Grafana[3], ver Figura 4, interoperable con las bases de datos seleccionadas, el cual permite supervisar tanto el rendimiento del sistema como parámetros críticos de seguridad. Esta arquitectura proporciona una plataforma robusta de análisis, facilitando la visualización de la evolución de las métricas a lo largo del tiempo. Por

[1]https://www.influxdata.com/
[2]https://www.mysql.com/
[3]https://grafana.com/oss/grafana/

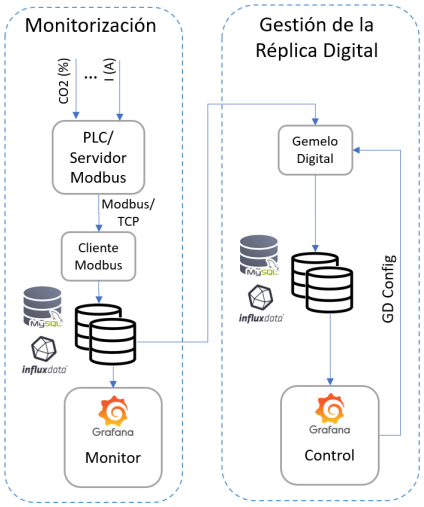


Fig. 3. Arquitectura del gemelo digital

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Fig. 4. Mock-up del interfaz gráfico del GD

otra parte, al estar conformada por herramientas de acceso abierto, facilita su replicabilidad.

2) Réplica Digital: El bloque funcional encargado de la gestión del GD dispondrá de una interfaz gráfica diferenciada, también en Grafana, para facilitar la configuración de las variables de entrada y la visualización de resultados. Estas variables, junto con las series históricas recopiladas por el sistema de monitorización, constituirán la entrada que alimenta el gemelo digital.

El módulo central del GD tiene como objetivo predecir el comportamiento de una planta de $ERCO_2$ ante diversos escenarios operativos. Estos escenarios podrán ser configurados a través de un panel específico en la interfaz gráfica que permitirá configurar diversos parámetros, como la intensidad o la concentración de CO_2 de entrada, y controlar el GD. Dado que no existe una relación algorítmica explícita entre las variables de entrada y la salida del proceso, la modelización y predicción del comportamiento del sistema real se realizará empleando técnicas basadas en aprendizaje automático e inteligencia artificial (ML/IA), lo que permitirá abordar la complejidad inherente al comportamiento dinámico de la planta, y optimizar además la toma de decisiones.

El modelo de comportamiento se entrenará con los datos históricos procedentes de la monitorización. En función de la naturaleza, cantidad y variabilidad de estos datos se evaluarán diferentes arquitecturas de redes neuronales. En caso de detectar patrones temporales significativos, se emplearán redes neuronales recurrentes (RNN), dada su capacidad para capturar dependencias a lo largo del tiempo. Por el contrario, si el conjunto de datos es limitado en tamaño o presenta una baja variabilidad temporal, se considerarán redes neuronales probabilísticas para cuantificar la incertidumbre de las predicciones.

IV. CONCLUSIONES Y LINEAS FUTURAS

Este trabajo presenta un enfoque innovador en el que se integra aspectos intrínsecos a la ingeniería telemática (análisis de datos, IA, protocolos de comunicaciones, desarrollo de servicios, entre otros) con los principios de la economía circular y la sostenibilidad, impulsando el desarrollo de soluciones para el reciclado de CO_2 en la industria textil.

La digitalización del sistema facilita el análisis de su rendimiento en tiempo real, ofreciendo métricas "a medida" a través de un interfaz de visualización amigable y accesible. Por otra parte, el almacenamiento de históricos permite la aplicación de técnicas IA para predecir el comportamiento en diversos escenarios con el objetivo final de optimizar el rendimiento del sistema real, permitiendo probar nuevas configuraciones antes de su puesta en marcha.

Como próximos pasos, se analizarán los datos históricos para determinar la solución más adecuada para el sistema. Adicionalmente, se estudiará el impacto de varios parámetros en el comportamiento del sistema antes de su despliegue en producción, prestando especial atención a variables difíciles de configurar, como la concentración de CO_2.

AGRADECIMIENTOS

Este trabajo ha sido financiado a través del Ministerio de Ciencia, Innovación y Universidades del Gobierno de España y la Unión Europea (Next GenerationEU/PRTR) a través del proyecto "Validación de un prototipo de planta de reciclado de CO2 en la industria textil (VALCO2-T)" (PLEC2022-009398), así como por el Gobierno de Cantabria a través del proyecto "Tecnologías habilitadoras de gemelos digitales y su aplicación a los sectores químico y de comunicaciones (GDQuiC)" del programa TCNIC (2023/TCN/002).

REFERENCIAS

[1] W. Leal Filho, P. Perry, H. Heim, M. A. P. Dinis, H. Moda, E. Ebhuoma, and A. Paço, "An overview of the contribution of the textiles sector to climate change," *Frontiers in Environmental Science*, vol. Volume 10 - 2022, 2022. [Online]. Available: https://doi.org/10.3389/fenvs.2022.973102

[2] United Nations, Department of Economic and Social Affairs - Sustainable Development, "Transforming our world: the 2030 Agenda for Sustainable Development," p. 16301, 2015. [Online]. Available: https://sdgs.un.org/2030agenda

[3] M. Ghobakhloo, "Industry 4.0, digitization, and opportunities for sustainability," *Journal of Cleaner Production*, vol. 252, p. 119869, 2020. [Online]. Available: https://doi.org/10.1016/j.jclepro.2019.119869

[4] T. A. Kurniawan, M. H. D. Othman, X. Liang, H. H. Goh, P. Gikas, T. D. Kusworo, A. Anouzla, and K. W. Chew, "Decarbonization in waste recycling industry using digitalization to promote net-zero emissions and its implications on sustainability," *Journal of Environmental Management*, vol. 338, p. 117765, 2023. [Online]. Available: https://doi.org/10.1016/j.jenvman.2023.117765

[5] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108 952–108 971, 2020.

[6] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.

[7] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison," *IEEE Access*, vol. 6, pp. 3585–3593, 2018.

[8] T. Riedelsheimer, L. Dorfhuber, and R. Stark, "User centered development of a digital twin concept with focus on sustainability in the clothing industry," *Procedia CIRP*, vol. 90, pp. 660–665, 2020, 27th CIRP Life Cycle Engineering Conference (LCE2020). [Online]. Available: https://doi.org/10.1016/j.procir.2020.01.123

[9] M. Maksimovic, "A faster path to sustainability: the use of digital twins," in *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2023, pp. 1–6.

[10] C. Metallidou, K. E. Psannis, D. D. Vergados, and M. Dossis, "Digital twin and industrial internet of things architecture to reduce carbon emissions," in *2022 4th International Conference on Computer Communication and the Internet (ICCCI)*, 2022, pp. 185–189.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

PALABRIA: Language Learning Enhancement through Linguistic Rule-Based Analytics and Artificial Intelligence

Pedro Manuel Moreno-Marcos[1], Marina Serrano-Marín[2], Pedro J. Muñoz-Merino[1], Natalia Centeno Alejandre[1], Rafael Fernández Castillejos[2], María Victoria Pavón Lucero[2], Carlos Delgado Kloos[1], Rubén Conde Rubio[2] and Ariana Suárez Hernández[2]
[1]Departamento de Ingeniería Telemática, [2]Departamento de Humanidades: Filosofía, Lenguaje y Literatura
Universidad Carlos III de Madrid
Calle Madrid, 126, Getafe (España)
pemoreno@it.uc3m.es, marserra@hum.uc3m.es, pedmume@it.uc3m.es, ncenteno@pa.uc3m.es, rafferna@pa.uc3m.es, mvpavon@hum.uc3m.es, cdk@it.uc3m.es, rconde@hum.uc3m.es, ashernan@hum.uc3m.es

In the academic and professional world, language is essential to carry out tasks and communicate results. However, errors in expression, such as incorrect grammar or inappropriate register, are common even among native speakers. To support writing, artificial intelligence, especially generative artificial intelligence, can help identify errors and provide personalized, timely support. While artificial intelligence has been widely applied in education, its use in linguistic education remains limited. These applications are more complex because linguistic rules tend to be flexible. Moreover, current solutions often rely on natural language processing but do not consider linguistic rules or provide limited feedback and analysis. In this context, PALABRIA-CM project is developed to improve students' skills and written Spanish by combining artificial intelligence with linguistic rules. This article analyzes challenges and opportunities of using generative artificial intelligence to improve writing, and presents an overview of the project, including objectives, methodology, work plan, and expected outcomes.

***Keywords*- Generative Artificial Intelligence, Linguistics, Language Learning, Learning Analytics**

I. Introduction

In the educational field, with the significant increase in the amount of information collected by digital educational platforms, learning analytics (LA) has emerged as a relevant discipline. Among its research lines, one line focuses on the design of adaptive or personalised systems for students. In this area, much work has been done on Intelligent Tutoring Systems (ITS). These systems are designed to assess students' knowledge and/or behaviour and provide adaptive support. For instance, they can adjust question difficulty, offer feedback based on errors, or provide scaffolding through hints or guiding questions to support the learning process (e.g., [1], [2]). When it comes to language learning, several ITSs have also been implemented; however, they still face some limitations. These include the inability to tailor tasks to different student proficiency levels, the challenges NLP models face in detecting certain errors, and the lack of detailed, helpful feedback for learners [3]. However, studies such as [4] acknowledge that the use of language can be significantly improved through the use of AI (Artificial Intelligence), and in some cases, AI may even outperform traditional methods. Despite this potential, most existing research focuses on foreign language learning, primarily English, rather than on enhancing language use for native speakers. There is also a notable lack of research addressing the Spanish language. Furthermore, many ITSs concentrate on specific linguistic components, such as grammatical constructions, rather than on broader skills like writing itself [3].

Although not always explicitly focused on linguistic assessment, a significant amount of research in LA has addressed the domain of writing using AI-based approaches. In particular, given the high number of participants in MOOCs (Massive Open Online Courses) and the fact that the main communication channel is the forum, many articles have analysed student text in the forum using text mining techniques (e.g. [5]). These analyses have been used to extract recommendations, more frequent topics, or generate responses [6]. Similarly, sentiment analysis of student forum messages has also been explored [7]. However, these studies often lack the integration of specific linguistic rules in their models. An example is Colby [8], who qualitatively analysed forum questions and answers to teach writing, although the study used coded text instead of AI.

Apart from analysing text to extract features from it, AI has also been used to infer behavioural variables and relate them to academic performance [9]. This aligns with one of the major interests in LA: the early detection of students at risk of dropping out or not succeeding in a course [10]. To this end, a large number of works have been developed incorporating

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

both more traditional machine learning techniques (regression, Support Vector Machines, Random Forest, etc.) and neural networks. Commonly used variables include interaction with videos [11], performance on exercises [12], and general platform usage [13]. Some studies have also considered high-level variables such as self-regulated learning, which has shown strong predictive power [14]. In addition, other machine learning techniques have been used to detect relevant aspects in education such as whether students are cheating (e.g., [15]). Moreover, recent works have analysed the content of the student's own submissions, mainly in the field of programming, where variables were extracted from the student's own code (e.g., [16]).

A similar approach could be applied to language learning, using linguistic features extracted from student texts to build predictive models. In fact, although this area is still emerging, some recent studies have started to explore its potential. Among them, Oliveira et al. [17] conducted a study to predict the level of text cohesion in Portuguese and English using AI, and da Silva Filho et al. [18] analysed how to assess the formal register of narrative essays in Portuguese by students using machine learning methods. However, as the authors acknowledge, more work is needed on the detection of language and register errors, and particularly in Spanish.

As a possible approach to research, and as an alternative to machine learning, it is possible to use generative AI. This branch of AI that focuses on generating realistic content (text, images, audio, etc.) is based on inputs provided by the user [19]. This branch has gained special relevance after the launch of some systems, such as OpenAI's GPT-4 or Gemini, whose performance is very close to that of humans [20]. Although it is still an emerging field in education, some initial work has already demonstrated its potential to support learning. For example, work at [21] demonstrates the potential of generative AI through an analysis to see how it is able to solve and evaluate simple mathematical exercises and provide feedback. Furthermore, Yan et al. [22] identified that generative AI can have multiple applications in the educational field, including detection, evaluation, teacher support, content generation and recommendation, among others. In the linguistic field, Kasneci et al. [23] highlighted the assistance in the development of reading and writing skills as one of the potential uses of generative AI. In this regard, we find recent works such as Ribera and Díaz Montesdeoca [24], focused on the specific use of ChatGPT in university education and on the importance of AI in academic environments, as well as in its application in teaching, research and management. On the other hand, Muñoz-Basols, Fuertes Gutiérrez and Cerezo [25] offer a guide designed for educators seeking to integrate technology into their pedagogical strategies, with a particular focus on the role of generative AI in language teaching and learning. This demonstrates that the application of generative AI to improve language use is possible and is a novel area in the state of the art.

In addition, consistent with the aforementioned works, more research is needed to provide personalised systems that can help students improve their use of Spanish using analytics and AI, whether generative or not. In this context, PALABRIA project (https://palabria.uc3m.es/) has emerged to design a system that detects errors in the use of Spanish language and provides analytics and feedback. This project is funded by Comunidad de Madrid through the grant agreement for the promotion of research and technology transfer at Universidad Carlos III de Madrid (PALABRIA-CM-UC3M).

This paper builds on this research context and presents the theoretical foundations and research plan of the PALABRIA project. The remainder of the paper is as follows: Section 2 analyses the challenges and opportunities of using generative AI to improve written language. Section 3 presents the initial hypothesis and objectives of the PALABRIA project. Section 4 discusses the methodology and work plan, and Section 5 presents the conclusions of this article.

II. Challenges And Opportunities Of Generative AI To Improve Language

In the Humanities, an intense debate has emerged regarding AI's potential to replace the human mind. Since language is precisely one of the distinctive features of the human species, the debate has been particularly intense in the field of linguistics [26]. AI offers powerful analytical tools that help linguists to study complex linguistic patterns and discover new insights into the structure and meaning of language. It also facilitates the creation of machine translation and natural language processing systems, improving cross-cultural communication and information accessibility. However, this technological advance also raises important challenges and questions in the field of linguistics and the humanities. One of the main concerns lies in the ability of AI to understand the cultural, historical and emotional context of human language. Authors such as Chomsky [27] and Mendívil [28] have pointed out that current language models, although very advanced, do not understand meaning or reflect human cognitive processes; they simply predict which words are likely to come next. Our project aims to help students improve their writing using AI tools, while relying on their own linguistic competence and critical thinking. To this end, there is an extensive normative literature that not only identifies language problems but also explores their origins and possible solutions [29]. To cite just a few examples: [30], [31], [32], in addition to the online resources provided by institutions such as the Real Academia Española (https://www.rae.es/espanol-al-dia) or the Fundéu (https://www.fundeu.es/).

Likewise, there is an extensive bibliography of a descriptive and theoretical nature on the specific syntactic constructions that will be the object of analysis, i.e. those where students commonly make errors in written academic texts: impersonal structures [33]; uses of the gerund [34], prepositions and verbal regimen [35], and discourse markers (see [36] for an overview). The descriptive and theoretical bibliography on this type of constructions is enormous, so two relatively recent general works that incorporate in their explanations the main findings of the last decades will be extremely useful: the Gramática descriptiva de la lengua española, directed by Ignacio Bosque and Violeta Demonte [37] and the *Nueva gramática de la lengua española*, edited by the RAE/ASALE [38], for which a new edition is currently in preparation.

Academic writing implies the ability to master the diamesic dimension of language, i.e. the differences that exist between communicative immediacy and distance and/or between spoken and written language ([39], [40]). One of the main challenges faced by any university student is the use of written code in a formal register. Among the difficulties they

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

have to deal with are not only the lack of knowledge of the educated register in general [41], a scarce lexical availability [42], the absence of coherence and cohesion to construct paragraphs that order their ideas [43], but also the impossibility of identifying linguistic units, structures and constructions sanctioned in the written language [44]. While style manuals such as the *Manual del español urgente* [45], *Libro de estilo de la lengua española* [32] or *El buen uso de las palabras* [46] offer useful solutions for the academic writing by university students, they cannot provide concrete solutions to the particular doubts of each student.

In order to analyse and process students' texts, which are often handwritten, it is essential to convert them into a digital format compatible with the AI tools designed for error correction. In this context, platforms like Transkribus [47], which uses Handwritten Text Recognition (HTR) technology, can be particularly useful. Although originally developed to transcribe historical documents, Transkribus can recognise texts from images in formats such as JPG, or PDF, and has been shown to effectively process modern handwritten texts as well [48], making it a promising tool for digitising student manuscripts in our project.

When analysing those texts, a key element to focus on in terms of language is the mastery of the normative uses of the language as this is crucial for effective, efficient and contextually appropriate communication in an academic environment. Both non-native and native learners commonly make mistakes in writing. Such deviations may include the use of registers inappropriate for written discourse, lack of textual cohesion due to colloquial impersonal expressions or the absence of discourse markers, the failure to distinguish between spoken and written code and its implications for textual microstructure, as well as grammatical problems. Among all these non-normative uses, it is worth focusing on 1) the gerund of posteriority, 2) the emphatic infinitive, 3) prepositions which are necessarily governed by certain verbs, and 4) the impersonal *tú*. These uses are phenomena typical of spoken language and communicative immediacy. However, the frequent and continuous use of social networks, in which communicative immediacy is the norm, but which use the written code, has led to a lack of pragmatic and diamesic language awareness. In other words, it has blurred the boundaries between the oral and the written, between the colloquial and the formal. It is not in vain that many universities include specific subjects on written expression or academic writing in the curricula of various degrees, such as the cross-cutting subject Techniques of oral and written expression (14360). Therefore, Artificial Intelligence (AI) can be applied to solve learning problems in the linguistic field, as it can detect the linguistic errors mentioned above and guide users as a tutor so that they can make the appropriate corrections. However, existing AI models have some limitations:

1) In many cases, NLP techniques have been used from a technological point of view, avoiding the inclusion of the linguistic part to improve their performance.
2) AI models applied in education are limited in the feedback they can provide.
3) Many existing models are based on the use of quantitative variables and take simple variables such as the length of texts [49] and not their content
4) Existing models have mostly been tested in English, with limited application to other languages (as previously stated).

To alleviate some of these problems, the recent massive use of generative AI can be applied in this field as a personal assistant or tutor. Moreover, generative AI can produce well-structured and coherent texts. It is also able to modulate speech at the user's request by means of appropriate prompts. In this way, prompts based on linguistic theories could be identified that would be suitable for different educational situations taking into account the context and learner profile. Therefore, the combination of linguistic aspects with AI, including generative AI, can help to contribute to an improvement of error detection models and in particular of Spanish (problems 1 and 4), to improve feedback (problem 2) and the use of generative AI can also help to obtain more variables from the provided text to improve analytics based on unstructured data (addressing problem 3).

However, our preliminary tests within the PALABRIA project reveal that generative AI alone is not sufficient to solve these challenges. Tests conducted with the 'Mistral-7B-Instruct-v0.1' model, used without fine-tuning or explicit linguistic guidance, illustrate these limitations. In particular, we tested various examples involving the impersonal tú construction. As shown in Table I, although the model produced accurate results in simpler cases, it often struggled with more complex grammatical structures. The examples highlight frequent issues related to verbal agreement and participial concord, with errors such as incorrect number agreement or influence from adjacent plural nouns. For instance, in the sentence "Tú identificaste patrones en los datos que antes no se habían considerado," the model generated "Se identificaron patrones en los datos que antes no se habían considerados", incorrectly using the plural form "considerados", possibly influenced by the plural noun "patrones", instead of the required singular "considerado" to agree with the impersonal construction. These findings suggest that generic AI, unguided by explicit linguistic rules or fine-tuning, struggles with subtle grammatical issues in Spanish, particularly verbal agreement and participial concord, which are fundamental for formal academic writing.

As a reference to these challenges, some notable AI systems in the field of text writing are QuillBot, DeepL Write and Artext System. The first two aim to enhance writing by correcting grammar and punctuation and by rephrasing entire sentences to produce more cohesive and coherent texts. However, DeepL Write currently does not offer a Spanish version, so its effectiveness in Spanish cannot yet be assessed. Moreover, it does not provide grammatical explanations for its corrections, as shown by several tests we conducted in other languages (English and German). On the other hand, QuillBot paraphrases the text we enter into the application and restructures it, with special emphasis on word order. However, the modifications it makes to the text are not always justified from a pragmatic-semantic point of view and it does not provide the user with any feedback. As far as Sistema Artext is concerned, its main function is to help in the construction of the macrostructure of both specialised and non-specialised texts through the use of templates. Like DeepL Write and QuillBot it does not provide any feedback.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Table I
EXAMPLES OF CORRECTIONS OF THE IMPERSONAL TÚ USING MISTRAL-7B-INSTRUCT-V0.1

Original (ES)	Original (EN)	Expected (ES)	Generated (ES)	BLEU	Error Description
Si tú desarrollas un marco teórico sólido, la investigación gana rigor.	If you develop a solid theoretical framework, the research gains rigor.	Si se desarrolla un marco teórico sólido, la investigación gana rigor.	Si se desarrolla un marco teórico sólido, la investigación gana rigor.	1.0000	*(Correct Output)*
Cuando tú identificas un problema, puedes proponer soluciones.	When you identify a problem, you can propose solutions.	Cuando se identifica un problema, se pueden proponer soluciones.	Cuando se identifica un problema, se **puede** proponer soluciones.	0.6606	*Verbal number agreement error*
Tú identificaste patrones en los datos que antes no se habían considerado.	You identified patterns in the data that had not been considered before.	Se identificaron patrones en los datos que antes no se habían considerado.	Se identificaron patrones en los datos que antes no se habían **considerados**.	0.9036	*Adjective agreement error*

The AI techniques on which the above-mentioned programs are based do not offer a complete solution to the different writing problems that the user may encounter. The absence of feedback is evidence that the error detection programming does not have a sufficiently robust linguistic implementation. Therefore, it would be of great help to have a personal assistant that helps to detect this type of linguistic problems and can provide adequate feedback to improve users' competence, adapted to their needs and contexts. Hence, the PALABRIA project stands as a possible solution to the shortcomings of AI techniques applied to language from an interdisciplinary perspective that combines the advantages of engineering and the linguistic knowledge of philology.

The use of AI tools is having a significant impact on what is commonly referred to as academic writing. Some of these models, such as ChatGPT or DeepL Write [50], lack certain skills that can only be attributed to humans, such as the contextual situation in which a certain discursive genre takes place, the resolution of lexical and syntactic ambiguities or intralinguistic variation. The aim of this project is to create a natural language processing solution that detects writing errors - both grammatical and discursive - and, at the same time, provides a normative explanation. To this end, collaboration between engineering experts and linguistic experts is essential to ensure a comprehensive and efficient development of strategies that can provide a solution to the problem posed.

Collaboration between engineers and linguists creates a synergy that transcends technical efficiency. The fusion of technical and linguistic knowledge enables the creation of tools, the use of software, in short, solutions that make it possible to understand and process natural language in a more contextualised way. On the one hand, Information and Communication Technology professionals have a deep knowledge of programming, algorithms and technologies associated with AI. This allows them to address crucial technical aspects for the development of word processors and tools that optimise performance, error detection and correction and provide a layer of personalisation for users. On the other hand, researchers in Linguistics and Philology bring a specialised perspective on the analysis of language, both syntax and semantics and pragmatics. Their in-depth knowledge of the microstructure and macrostructure of texts, as well as of linguistic variation subject to geolectal, sociocultural and contextual parameters, can contribute significantly to the improvement of algorithms dealing with the detection of grammatical errors, as they can provide precise descriptions of the different linguistic phenomena sanctioned by the norm of the written language.

This collaborative and interdisciplinary approach not only benefits end-users, who experience tangible improvements in the quality and functionality of tools, but also fosters the development of more inclusive and adaptable solutions for linguistic and cultural diversity.

In order to solve these scientific issues, an interdisciplinary project such as the one proposed here is necessary, combining on the one hand the knowledge of linguistic rules, linguistic contexts and expert knowledge of appropriate feedback in each case and on the other hand the technological knowledge related to AI, in order to develop and evaluate technical AI solutions.

III. INITIAL HYPOTHESIS AND OBJECTIVES

Given the challenges outlined above, this project aims to integrate linguistic aspects with artificial intelligence (AI) to enhance the use of the Spanish language. Our hypothesis is that this integration within an educational setting can empower learners to identify and fix writing issues, such as grammatical errors, register inconsistencies, and coherence problems, ultimately leading to improved language proficiency. Thus, the main objective of the PALABRIA project is to enhance language support for students by integrating AI tools with linguistic principles, and the specific objectives (O) are as follows:

- **O1: Develop a research framework** comprising: 1) a state-of-the-art review of AI applications in language error detection; 2) a pedagogical model emphasising AI in education to support students and teachers; 3) a set of scenarios, in collaboration with teachers, to analyse how AI-linguistic integration can offer new opportunities for language improvement, including feedback; and 4) the definition of research instruments for evaluating and developing pilot experiences.
- **O2: Design and develop AI solutions** to enhance language support for Spanish users by addressing the AI challenges in education outlined in section 1.2. Contributions to this objective will be made through: 1) proposing configurations for using generative AI and evaluating the performance of such solutions, using tools like ChatGPT or Gemini, in detecting language errors and providing feedback; 2) developing a system that combines AI (e.g., using machine learning, generative AI, etc.) and linguistic rules for detecting errors in written language; 3)

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

developing an AI-powered language analytics system to detect discourse or writing characteristics of students and provide analytics or predictions about their language performance and its relationship to course performance; 4) developing an adaptive system, including AI, to provide personalised feedback to students on their formal language use.

- **O3: Define a technological and linguistic framework** consisting of: 1) analysing and selecting existing AI technologies capable of detecting errors in written language; 2) specifically defining linguistic rules and common language problems to be addressed in AI models, such as register and textual coherence issues (e.g., impersonal expressions, discourse markers, differences between spoken and written language) or grammatical problems (e.g., gerund use, preposition confusion in verbal complements, emphatic infinitive, etc.); and 3) proposing technical requirements and architecture for a system that can integrate linguistic aspects with AI for language error detection, student characterisation, and personalised feedback, considering the needs of teachers and students, as well as the pedagogical aspects from O1.
- **O4: Design, implement, and evaluate pilot experiences** in real-world contexts based on the project's results (particularly objectives O2 and O3). The scenarios defined in O1 and the systems developed in O2 will be used, following the architecture proposed in O3. Given that the project's contributions can be applied to several educational scenarios at different levels, the project will demonstrate the potential in different contexts, including, for example, humanities, science and engineering students, both nationally and internationally in Spanish-speaking countries.

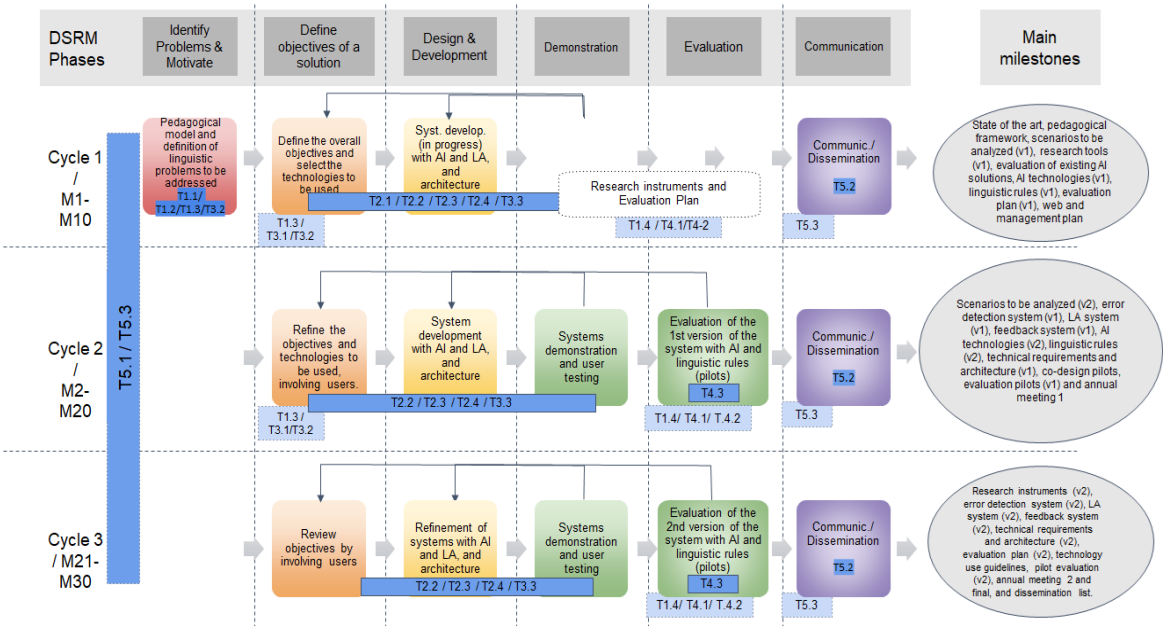
These objectives are clearly aligned with Telematics Engineering as they address several aspects such as distributed system architecture design, including the design of an architecture detection and feedback, the proposal of learning analytics indicators based on distributed data, or the analysis and selection of AI technologies, or the integration and interoperability of technologies. These technologies comprise servers, databases, AI models (generative and non-generative), and business rules. This contribution may benefit Telematics Engineering students, but also other students, so it is not specifically tailored to them. Instead, this is a contribution in Educational Technologies Research. This is because the expected results have an educational component, and the developed systems are expected to improve students' language use. In relation to learning analytics, it is expected that analytics will help to better understand students' behaviours related to language use and provide personalized feedback that may have a positive effect on learning..

IV. Methodology And Work Plan

a. Methodology

The development of this project involves the incorporation of AI technologies into educational environments to improve students' written language skills and usage. To achieve this, solutions will be developed to address the problems identified in section 1.2, which entails studying the most appropriate technologies for both design and implementation. Given this approach, the Design Science Research Methodology (DSRM) [49] is a highly suitable methodology for the project's development and has been widely used in information systems research projects.

This methodology consists of six phases, which generally follow a sequential order: (1) identifying the problem and motivating it (problems are characterised by unstable requirements and constraints arising from poorly defined contexts and complex component interactions, allowing flexibility in both design process and artefacts), (2) defining solution objectives (the goal is to develop technological solutions to address problems), (3) design and development (DSRM must produce an artefact in the form of a construction, model, method, or instantiation), (4) demonstration (effective



Project evolution following the phases of the DSRM methodology in three cycles: the blue labels represent the project tasks and the circles on the right the main project milestones.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

research must provide reliability in the areas of artefact design, design fundamentals, and design methodologies), (5) evaluation (rigorous methods must be applied to the construction and evaluation of artefacts), and (6) effective communication to both technological and managerial audiences. Moreover, in DSRM, these activities do not necessarily follow a strict order and may be refined iteratively [49]. Furthermore, the design process will be enriched with a user-centred approach, involving key stakeholders (e.g., teachers) identifying problems and defining objectives, as well as in the design of the systems and pilot activities. Stakeholder involvement in the process will also consider aspects such as diversity, inclusion, and gender equity to ensure the participation of women and/or socially excluded groups.

Figure 1 shows the specific adaptation of the DSRM methodology for this project, illustrating three main iterations, each lasting 10 months. Each iteration includes the previously mentioned phases, and the figure identifies the tasks corresponding to each one. Two phases are particularly critical and require special attention: the definition of objectives at the beginning of the project and the evaluation of the proposal in the second and third cycles. The definition of objectives requires involving stakeholders to incorporate their perspectives into the development of the solutions. Therefore, potential stakeholders will participate in tasks T1.2, T1.3, T4.1, and T4.2. The second critical phase is evaluation, as it requires verifying that the systems function correctly through pilots. On the one hand, this phase will detect potential errors that will need to be corrected to create a mature system. On the other hand, the development of the pilots will have certain challenges, such as potential ethical, privacy, or data protection issues. Regarding the first, the iterative methodology will allow for the refinement of the proposal, so that through the tests of the second and third cycles, a more mature version will be obtained. Regarding the second, a Data Management Plan (T5.3) will be developed at the beginning of the project to serve as a guide for these issues.

b. Work Plan

The work plan is structured into five Work Packages (WPs). The first four WPs correspond to the specific project objectives (O1-O4), while the fifth package ensures overall coordination, dissemination, and responsible data management.

- *WP1: Research Framework:* The aim of this WP is to provide the context for the research of this project. This context is determined by the state of the art on the use of AI for the detection of errors in language use, the pedagogical model to incorporate the use of AI in education and support students and teachers, the selection of scenarios to be used to analyse the impact of the solution proposed in the project, and the definition of the research methods and tools needed to report the results derived from the evaluation tasks of the project.
- *WP2: AI support to learners* in the use of Spanish. This work package aims at developing solutions that include AI to detect learners' errors in the written linguistic domain and provide them with personalised feedback and analytics, enabling them to improve their use of the language. For this purpose, this package takes as inputs the research framework developed in WP1. Moreover, the developments made will serve to define the overall architecture of the system in WP3.
- *WP3: Technological and linguistic framework.* The goal of this WP is to establish a general architecture, with certain guidelines, for what the AI-enriched language use enhancement support system should look like. This WP builds on the results of WP2 package to provide a framework that can be used by other researchers when carrying out similar projects. The development process follows an iterative approach with multiple versions and continuous evaluation.
- *WP4: Pilot Experiences.* The main objective of this WP is to implement and evaluate in real educational contexts the results of the project. For this purpose, the results of WP1-WP3 will be used. Given the technological and methodological nature of the project contributions, and the fact that the use of language is transversal to multiple domains, the pilots can be applied to multiple educational contexts, coming from different subject areas and educational levels. To address this aspect, pilots that include different subject areas (e.g., humanities and engineering) and different levels (e.g., undergraduate and graduate students) will be proposed. In addition, to ensure the relevance of the results, different stakeholders (teachers and students) will be involved, considering gender and diversity perspectives.
- *WP5: Coordination, dissemination and data management.* This package is responsible for overall project coordination, results dissemination and data management. As the project involves the work of two research groups from two different research areas (ICT and FLA), coordination between the two groups is critical to the success of the project. Dissemination of the results is also crucial to maximise the impact of the project and therefore must be specifically analysed in order to be carried out efficiently and effectively. Likewise, data collection and processing are fundamental aspects, which must be explicitly managed to guarantee aspects such as the privacy of the participants' data, in accordance with ethical principles and in compliance with regulations of personal data protection.

In the first year of the project, two linguistic issues have been studied (impersonal tú [51] and gerund of posteriority) and several models, including Mistral and Llama 2, were implemented to detect these errors. Prompt engineering was applied to guide the models' outputs according to specific linguistic rules. In [52], 300 samples were used, and the best observed GLEU (Google BLEU, Bilingual Evaluation Understudy) was 0.85. This demonstrates the potential of the approach used in this project.

V. Conclusions

The PALABRIA project aims to improve students' support in their use of the language by integrating AI tools with linguistic aspects. The current challenges when using AI, and particularly in the linguistic field, have been discussed, as well as the opportunities that emerge from the combination of AI

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

and linguistics aspects. Considering this combination, the PALABRIA project proposes: 1) a research framework, including the pedagogical model to include the proposal in the educational field, the scenarios to analyse and research instruments, 2) AI solutions to support learners in the use of Spanish by detecting errors and providing personalised feedback and analytics, 3) a technological framework to establish a general architecture for the proposed solutions, and 4) pilot experiences to validate the results. The nature of this project, which involves specific linguistics aspects (e.g., analysis of specific linguistics rules, such as gerund of posteriority, and the emphatic infinitive, among others) and the use of IA techniques, which involves a technological background, implies that an interdisciplinary approach is needed. This will add value to the limited previous work, which is based on a smaller subset of techniques in a limited set of scenarios, where Spanish language is often not taken into account.

ACKNOWLEDGEMENTS

This work was funded by Comunidad de Madrid through the grant agreement for the promotion of research and technology transfer at Universidad Carlos III de Madrid (PALABRIA-CM-UC3M).

REFERENCES

[1] Cristina Conati, Abigail Gertner, Kurt Vanlehn, "Using Bayesian networks to manage uncertainty in student modeling", *User modeling and user-adapted interaction*, vol. 12, pp. 371-417, 2002.

[2] Pedro J. Muñoz-Merino, Carlos Delgado Kloos, "An architecture for combining semantic web techniques with intelligent tutoring systems", in *Intelligent Tutoring Systems: 9th International Conference*. Canada: Springer Berlin Heidelberg, pp. 540-550, June 23-27, 2008.

[3] Vanja Slavuj, Božidar KovačićIgor, JugoIgor Jugo, "Intelligent tutoring systems for language learning", in *38th International Convention on Information and Communication Technology, Electronics and Microelectronis*, pp. 814-819, 2015.

[4] Xieling Chen, Di Zou, Haoran Xie, Gary Cheng, Caixia Liu, "Two decades of artificial intelligence in education", *Educational Technology & Society*, vol. 25, no. 1, pp. 28-47, 2022.

[5] Pedro Manuel Moreno-Marcos, Carlos Alario-Hoyos, Pedro J. Muñoz-Merino, Carlos Delgado Kloos, "Prediction in MOOCs: A review and future research directions", *IEEE transactions on Learning Technologies*, vol. 12, no. 3, pp. 384-401, 2019

[6] Dharamjit Parmar, M. Ali Akber Dewan, Dunwei Wen, "Automatic Analysis of Online Course Discussion Forum: A Short Review", in *2023 IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 210-215, 2023.

[7] Pedro Manuel Moreno-Marcos, Carlos Alario-Hoyos, Pedro J. Muñoz-Merino, Carlos Delgado Kloos, "Sentiment analysis in MOOCs: A case study", in *2018 IEEE Global Engineering Education Conference*, pp. 1489-1896, 2018.

[8] Rebekah Shultz Colby, "Using online writing communities to teach writing MOOCs", in Research anthology on facilitating new educational practices through communities of learning, pp. 522-535, 2021.

[9] Marcus Klüsener, Albrecht Fortenbacher, "Predicting students' success based on forum activities in MOOCs", in *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, vol. 2, pp. 925-928, 2015.

[10] Nabila Sghir, Amina Adadi, Mohammed Lahmer, "Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022)", *Education and information technologies*, vol. 28, no. 7, pp. 8299-8333, 2023.

[11] Christopher G. Brinton, Swapna Buccapatnam, Mung Chiang, H. Vincent Poor, "Mining MOOC clickstreams: Video watching behavior vs. in-video quiz performance", *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3677-3692, 2016.

[12] Zhiyun Ren, Huzefa Rangwala, Aditya Johri, "Predicting Performance on MOOC Assessments using Multi-Regression Models", in *9th International Conference on Educational Data Mining*, 2016.

[13] Ahmed Alamri, *et al*., "Predicting MOOCs dropout using only two easily obtainable features from the first week's activities", in *15th International Conference on Intelligent Tutoring Systems*. Cham: Springer, pp. 163-173, 2019.

[14] Pedro Manuel Moreno-Marcos, Pedro J. Muñoz-Merino, Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, Carlos Alario-Hoyos, Carlos Delgado Kloos, "Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs", *Computers & Education*, vol. 145, pp. 103728, 2020.

[15] José A. Ruipérez-Valiente, Pedro J. Muñoz-Merino, Giora Alexandron, David E. Pritchard, "Using machine learning to detect 'multiple-account'cheating and analyze the influence of student and problem features", *IEEE transactions on learning technologies*, vol. 12, no. 1, pp. 112-122, 2017.

[16] Munstair Hoq, Peter Brusilovsky, Bita Akram, "Analysis of an Explainable Student Performance Prediction Model in an Introductory Programming Course", in *16th International Conference on Educational Data Mining*, pp. 79-90, 2023.

[17] Hilário Oliveira, *et al*., "Towards explainable prediction of essay cohesion in Portuguese and English", in *13th International Learning Analytics and Knowledge Conference*. New York: Association for Computing Machinery, pp. 509-519, 2023.

[18] Moésio Wenceslau da Silva Filho, *et al.*, "Automated formal register scoring of student narrative essays written in portuguese", in *Congresso Brasileiro de Informática Na Educação 2023*, pp. 1-11, 2023.

[19] Mlađan Jovanović, Mark Campbell, "Generative artificial intelligence: Trends and prospects", *Computer*, vol. 55, no. 10, pp. 107-112, 2022.

[20] Sébastien Bubeck, et al., "Sparks of artificial general intelligence: Early experiments with gpt-4", *arXiv preprint arXiv:2303.12712*, 2023.

[21] Huy A. Nguyen, Hayden Stec, Xinying Hou, Sarah Di, Bruce M. McLaren, "Evaluating chatgpt's decimal skills and feedback generation in a digital learning game", in *18th European Conf. Technology Enhanced Learning*. Cham: Springer, pp. 278-293, 2023.

[22] Lixiang Yan, Roberto Martinez-Maldonado, Dragan Gasevic, "Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics Cycle", in *14th ILAK Conference*. New York: Association for Computing Machinery, pp. 101-111, 2024.

[23] Enkelejda Kasneci, et al., "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and individual differences*, vol. 103, pp. 102274, 2023.

[24] Mireia Ribera, Oliver Díaz Montesdeoca (coords.), *ChatGPT y educación universitaria: posibilidades y límites de ChatGPT como herramienta docente*. Barcelona: Octaedro – IDP/ICE/UB, 2024.

[25] Javier Muñoz-Basols, Mara Fuertes Gutiérrez, Luis Cerezo (eds.), *La enseñanza del español mediada por tecnología. De la justicia social a la Inteligencia Artificial (IA)*. London, New York: Routledge, 2024.

[26] Rafael Muñoz Guillena, "La SEPLN y futuro", in *Encuentro. Estrategias y agentes del Procesamiento del Lenguaje Natural*, 2024.

[27] Noam Chomsky, Ian Roberts, Jeffrey Watumull, "The False Promise of ChatGPT", *The New York Times*, March 8, 2023.

[28] José Luis Mendívil Giró, "Un loro estocástico en la habitación china. ¿Qué nos enseña CHATGPT sobre la mente humana?", *Letras libres*, vol. 262, pp. 16-22, 2023.

[29] Leonardo Gómez Torrego, "La norma como herramienta para la reflexión gramatical", *Revista de filología española*, vol. 85, no. 1, pp. 61-79, 2005.

[30] Instituto Cervantes, *El libro del español correcto. Claves para escribir y hablar bien en español*. Madrid: Instituto Cervantes-Espasa, 2012.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[31] Real Academia Española, Asociación de Academias de la Lengua Española, *El buen uso del español*. Madrid: Espasa, 2013.
[32] Real Academia Española, Asociación de Academias de la Lengua Española, *Libro de estilo de la lengua española según la norma panhispánica*. Madrid: Espasa, 2018.
[33] María Alonso Alonso, "Un estudio comparativo de la impersonalidad en el lenguaje académico a través de la UAM Corpus Tool", *Hesperia: Anuario de Filología Hispánica*, vol. 14, no. 2, pp. 23-38, 2011.
[34] Anne Verhaert, *El gerundio no perifrástico del español,cómo no ser demasiado explícito ni demasiado implícito*. The Netherlands: Brill, 2008.
[35] Emile Slager, *Las preposiciones en español*. Madrid: Castalia, 2012.
[36] Catalina Fuentes Rodríguez, Ester Brenes Peña, *Macrosintaxis y pragmática del discurso persuasivo.* Navarra: EUNSA, 2023.
[37] Ignacio Bosque, Violeta Demonte (dirs.), *Gramática descriptiva de la lengua española*. Madrid: Espasa, 1999.
[38] Real Academia Española, Asociación de Academias de la Lengua Española, *Nueva gramática de la lengua española*. Madrid: Espasa, 2009.
[39] Oswald Ducrot, *El decir y lo dicho: polifonía de la enunciación*. Barcelona: Paidós, 1986.
[40] Marcello Barbato, "La sociolinguistique et l'histoire de varietés romanes anciennes", *Revue Tranel (Travaux neuchâtelois de linguistique)*, vol. 53, pp. 77-92, 2011.
[41] María del Carmen Quiles Cabrera, "Interferencias oralidad escritura en el discurso formal: una aproximación empírica en el ámbito universitario", in Luis María Cortés Rodríguez (coord.), *Discurso y oralidad: homenaje al profesor José Jesús de Bustos Tovar*. España: Arcos/Libros, vol. 2, pp. 847-860, 2007.
[42] Humberto López Morales, *Disponibilidad léxica en escolares de San Juan*. San Juan de Puerto Rico: Universidad de San Juan de Puerto Rico, 1973.
[43] Ingrid Isabel Medina Cardozo, Marcos Oswaldo Vásquez, "Coherencia y cohesión en el discurso escrito de estudiantes universitarios UCV-HACER", *Revista de Investigación y Cultura*, vol. 2, no. 1, pp. 44 -57, 2013.
[44] Araceli López Serena, "Hablar y lo oral", in Óscar Loureda, Angela Schrott (eds.), *Manual de lingüística del hablar*. Berlin: De Gruyter, 2021.
[45] EFE, *Manual del español urgente*. Madrid: Cátedra, 1991.
[46] Valentín García Yebra, *El buen uso de las palabras*. Madrid: Gredos, 2005.
[47] Guenter Muehlberger, *et al*., "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study", *Journal of Documentation*, vol. 75, no. 5, pp. 954-976, 2019.
[48] Wannapon Suraworachet, Cristina Villa-Torrano, Qi Zhou, Juan I. Asensio-Pérez, Yannis Dimitriadis, Mutlu Cukurova, "Examining the relationship between reflective writing behaviour and self-regulated learning competence: A time-series análisis", in *16th European Conference on Technology Enhanced Learning, EC-TEL 2021*. Cham: Springer, pp. 163-177, 2021.
[49] Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, Samir Chatterjee, "A Design Science Research Methodology for Information Systems Research", *Journal of management information systems*, vol. 24, no. 3, pp. 45-77, 2007.
[50] Uttkarsha Bhosale, "Limitations of DeepL Write and ChatGPT for Editing Academic Texts", *Enago Academy*, Nov. 23, 2023.
[51] Marina Serrano Marín, Pedro Manuel Moreno Marcos, María Victoria Pavón Lucero, "Tú impersonal en la lengua académica: a propósito del proyecto PALABRIA," in *XVI Congreso Internacional de Lingüística General, CILG 2025*. Alicante: Universitat d'Alacant - Universidad de Alicante, pp. 380-381, 2025.
[52] Natalia Centeno Alejandre, Pedro Manuel Moreno Marcos, Marina Serrano Marín, "Mejora de la Escritura en Español mediante Inteligencia Artificial y Reglas Explícitas en Contextos Académicos," in *VII Congreso de la Sociedad Internacional de Humanidades Digitales Hispánicas, HDH 2025* (in press). 2025.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Redes para un futuro con bajas emisiones: Optimización del enrutamiento consciente del carbono en redes SDN

José Gómez-delaHiz, Jaime Galán-Jiménez
Departamento de Ingeniería de Sistemas Informáticos y Telemáticos,
Universidad de Extremadura
Av. de la Universidad, S/N, 10003, Cáceres, Extremadura, España.
jagomezdh@unex.es, jaime@unex.es

Resumen

El uso de Internet está creciendo exponencialmente en los últimos años. Este crecimiento provoca un aumento de los recursos consumidos por las redes de comunicaciones y, por tanto, un aumento de su consumo de energía, cuya generación implica emitir gases de efecto invernadero a la atmósfera. En estos tiempos, uno de los objetivos de la humanidad es reducir las emisiones de carbono, en búsqueda de mantener el estado de nuestro planeta. En este contexto, este artículo presenta ANPA, una solución Integer Linnear Programming (ILP) que minimiza las emisiones de carbono de las redes definidas por software al realizar un enrutamiento consciente del carbono y modificar la tasa operacional de los puertos integrados en los dispositivos de red, lo cual limita su ancho de banda, pero también su consumo de energía. Las simulaciones sobre topologías de red reales demuestran que nuestra solución consigue una reducción de emisiones de carbono de hasta un 84,5 % en comparación a los esquemas tradicionales de encaminamiento por el camino más corto, y hasta un 64,49 % en comparación con otras soluciones en el estado del arte.

***Palabras Clave*—huella de carbono, consumo de energía, sostenibilidad, redes definidas por software, ILP**

I. INTRODUCCIÓN

Según la Unión Internacional de Telecomunicaciones (UIT), de 2019 a 2024 el aumento de la población que utilizó Internet pasó del 53 % al 63 %, es decir, un 15 % en esos 5 años. Sin embargo, en lugar de centrarnos en el recuento de usuarios, si examinamos el tráfico de banda ancha dentro de las redes fijas, la UIT informó de que en 2019 este tráfico ascendió a 2 ZetaBytes, es decir, 2.000 millones de GigaBytes. En 2024, este volumen se elevó a 6 ZetaBytes, lo que supone un aumento del 200 % en el mismo periodo de tiempo.

El uso de Internet está creciendo en gran medida, lo que conlleva una mayor carga para las redes que lo componen y, por tanto, un mayor consumo de energía [1], [2]. Sin embargo, lo que más preocupa hoy en día son los efectos secundarios de este aumento del consumo de energía, concretamente la huella de carbono. Las emisiones de carbono tienen efectos devastadores en nuestro planeta, y organizaciones como las Naciones Unidas han empezado a establecer objetivos para reducir estas emisiones: reducir el 45 % de las emisiones de gases de efecto invernadero a la atmósfera para 2030, y conseguir emisiones netas cero para 2050 [3]. El consumo energético de las Tecnologías de la Información y la Comunicación (TIC) alcanza hasta el 8 % del consumo total mundial. Si no se toman medidas en un futuro próximo, la huella total de las TIC se situará en torno al 14 % de la huella mundial total [4]. Parte del consumo energético derivado del uso de las TIC proviene del uso de las redes informáticas, por lo que es necesario pensar en formas de optimizar sus emisiones de carbono para contribuir a alcanzar las emisiones netas cero.

Las emisiones de carbono producidas al generar un kWh se conocen como *intensidad de carbono*, y dependerán de las fuentes de energía utilizadas para generar electricidad: producir energía con carbón emite $\sim$950 gCO_2/kWh, mientras que el gas natural alcanza $\sim$450 gCO_2/kWh. Sin embargo, si se utilizan fuentes de energía renovables como la eólica o la hidráulica, los valores se reducen hasta $\sim$15 gCO_2/kWh y $\sim$20 gCO_2/kWh, respectivamente [5]. Si aplicamos esto a las redes de comunicaciones, los dispositivos de red con idéntico consumo de energía, pero ubicados en diferentes zonas, darán lugar a diferentes emisiones de carbono. Por lo tanto, es razonable suponer que si reducimos la carga de los dispositivos ubicados en zonas alimentadas por fuentes de energía intensivas en carbono (al tiempo que trasladamos esa carga a regiones con energías más limpias), podríamos reducir significati-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

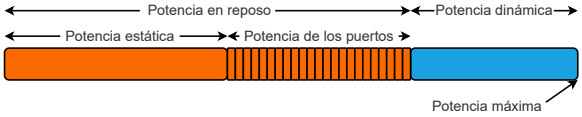


Figura 1: Modelo de consumo de energía de un router

vamente las emisiones totales de carbono de la red.

El modelo de consumo energético de un router se puede dividir en tres componentes [6]. Como podemos ver en la Fig. 1, el primer componente del modelo es la potencia estática (representada en naranja; este consumo no se puede reducir), que incluye la potencia del chasis, el módulo supervisor, etc. El segundo aspecto permanece estático, pero depende no solo del número de puertos habilitados sino también de su tasa operacional. Es decir, mientras que la habilitación de un puerto incurre en un coste energético de referencia, el consumo energético total de un router variará adicionalmente en función de la tasa operacional, ya que los distintos modos imponen diferentes velocidades de datos y demandas energéticas [7]. El último componente del modelo es la potencia dinámica (en azul en la Fig. 1), que depende de la carga de tráfico del router.

Partiendo de esta premisa, la minimización de la energía en los dispositivos de red puede abordarse de dos formas complementarias. En primer lugar, reduciendo la tasa operacional de los puertos (incluso hasta el punto de desactivarlos por completo), lo que disminuye su consumo de energía. En segundo lugar, minimizando la carga de tráfico procesada por un nodo determinado, lo que puede suponer un aumento de la carga en el resto de nodos. Como ya se ha mencionado, una redistribución de este tipo puede suponer un ahorro significativo en las emisiones de carbono, aunque no se consigan reducciones en el consumo de energía.

Si nos centramos en la gestión de la red, se debe obtener datos en tiempo real de los dispositivos, calcular la nueva configuración de la red desde un punto de vista global y, por último, recalcular las tablas de encaminamiento de todos los dispositivos de la red. Las arquitecturas distribuidas tradicionales, en las que las decisiones de control se toman localmente y de forma independiente por cada dispositivo, están intrínsecamente limitadas en su capacidad para realizar dicha optimización coordinada.

En cambio, las *Software Defined Networks* (SDN) ofrecen un paradigma de control centralizado que separa el plano de datos del plano de control, ubicando este último en un controlador central. Esta separación arquitectónica permite al controlador SDN obtener una visión detallada de toda la red Como resultado, SDN es especialmente adecuada para calcular políticas de optimización global (como ajustar la tasa operacional de los puertos o redirigir flujos para minimizar el consumo energético) y aplicarlas de forma coherente en toda la red [8], lo cual encaja perfectamente con nuestro enfoque.

En este trabajo se define e implementa una solución basada en Programación Lineal Entera (*Integer Linear Programming*, ILP) que busca minimizar las emisiones de carbono modificando la tasa operacional de los puertos y recalculando las rutas de los flujos que atraviesan la red. El proceso tiene en cuenta la intensidad de carbono de las regiones donde se encuentran los dispositivos de red, con el objetivo de reducir las emisiones totales de carbono en la red.

A continuación, se presentan las principales contribuciones de este trabajo:

- la propuesta de un modelo mejorado de consumo energético para routers;
- la integración de APIs en tiempo real para incorporar datos realistas de emisiones de carbono en redes SDN;
- la formulación de un problema de optimización que minimiza las emisiones de carbono tomando decisiones sobre la estructura de red y el enrutamiento de flujos;
- la evaluación de la solución mediante un análisis de escalabilidad y una comparación con dos soluciones alternativas sobre topologías reales.

El resto del artículo se organiza de la siguiente manera: en la Sección II se revisan trabajos relacionados; en la Sección III se define el modelo del sistema; en la Sección IV se describe el problema de optimización; los resultados experimentales se presentan en la Sección V; finalmente, en la Sección VI se exponen las conclusiones y posibles líneas futuras de investigación.

II. TRABAJOS RELACIONADOS

En los últimos años, la búsqueda continua de soluciones para mejorar la Calidad de Servicio (QoS) en redes cableadas ha evolucionado progresivamente, poniendo un énfasis creciente en la eficiencia energética y, más recientemente, en la reducción de las emisiones de carbono generadas por estas infraestructuras [6], [9], [10], [11], [12], [13], [14]. Una estrategia común adoptada por diversos investigadores consiste en el apagado selectivo de componentes de red y en el diseño de mecanismos de enrutamiento que consideren explícitamente métricas de consumo energético o de emisiones de carbono.

Por ejemplo, El-Zahr *et al.* proponen el algoritmo heurístico CATE, que minimiza las emisiones de carbono mediante la activación y desactivación selectiva de enlaces de red, combinado con políticas de enrutamiento personalizadas [6]. Su enfoque incluye la definición de métricas detalladas relacionadas con la energía y el carbono para modelar el consumo energético a nivel de router y las emisiones de carbono asociadas, validadas mediante pruebas realizadas sobre dos topologías *Wide-Area Network* realistas. En cambio, *Tabaeiaghdaei et al.* abordan exclusivamente la selección de rutas con conciencia de carbono, sin poner en reposo ningún puerto. Su contribución, CIRo, predice la intensidad de carbono de las rutas intra-dominio y difunde esta información a los extremos de la red para facilitar la elección de trayectos de comunicación más sostenibles [11].

Integrando aún más los objetivos de eficiencia energética y reducción de carbono, Hua *et al.* presentan una estrategia de enrutamiento energético diseñada específicamente

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

para redes con estructura de grafo [12]. Su trabajo optimiza simultáneamente las pérdidas de energía y las emisiones de carbono, proporcionando soluciones óptimas de Pareto mediante una combinación de *Depth-First Search* para la exploración de rutas y *U-NSGA-III*, un algoritmo genético adaptado para la optimización multiobjetivo que proponen.

Además, varios trabajos de investigación se han centrado en redes SDN, aprovechando su naturaleza centralizada y programable para mejorar la eficiencia energética sin comprometer la QoS. En este contexto, Torkzadeh *et al.* presentan un mecanismo de enrutamiento eficiente en energía de dos fases, diseñado específicamente para entornos SDN. La primera fase emplea optimización basada en colonia de hormigas para identificar un subgrafo energéticamente eficiente de la red, apagando los switches innecesarios. Posteriormente, la segunda fase enruta los flujos entrantes utilizando un algoritmo ponderado que tiene en cuenta el tamaño de los flujos, las restricciones de QoS y la carga de los enlaces, con el fin de equilibrar el tráfico de manera eficaz [13].

En escenarios que involucran redes híbridas SDN/IP, Wei *et al.* amplían este enfoque proponiendo un modelo de programación lineal complementado con una heurística denominada HEATE [14]. Esta propuesta se divide en dos componentes fundamentales: la optimización de los pesos de los enlaces OSPF para influir en el comportamiento de enrutamiento de los routers IP tradicionales, y una técnica dinámica de división de flujos gestionada por un controlador SDN. Esta estrategia dual permite consolidar el tráfico en un menor número de enlaces, lo que posibilita apagar aquellos infrautilizados, aumentando así significativamente el ahorro energético.

En resumen, los trabajos revisados en esta sección destacan un imperativo común: reducir el consumo energético o las emisiones de carbono de la red mediante la desactivación estratégica de enlaces infrautilizados y la incorporación de algoritmos de enrutamiento innovadores. Sin embargo, la mayoría de estas metodologías emplean decisiones binarias respecto al estado de los enlaces (activos o inactivos). A diferencia de estos enfoques, nuestra propuesta introduce el concepto de múltiples modos de operación para los puertos de la red. Este enfoque busca mejorar la precisión en la toma de decisiones y aumentar la eficiencia global de la red al ofrecer un control más granular y una mayor capacidad de adaptación a distintas condiciones de tráfico.

III. MODELADO DEL SISTEMA

En esta sección se describe el modelo del sistema, concretamente en lo relativo a los datos que se proporcionan como entrada a la solución y a los resultados obtenidos.

La solución puede aceptar como entrada cualquier topología, indicando el conjunto de nodos y enlaces que los conectan. Las pruebas se realizan teniendo en cuenta las características de un único router, independientemente de la topología red, aunque adaptado a las características de esta.

En este punto, el algoritmo requiere conocer la intensidad de carbono de las zonas en las que se encuentra

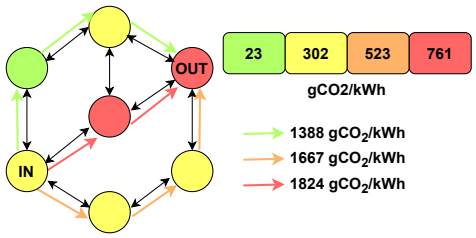


Figura 2: Ejemplo de intensidad de carbono.

cada nodo para poder iniciar su ejecución. Tal como se define en la Sección I, esta métrica establece la relación entre el consumo energético y las emisiones de carbono. La Fig. 2 representa una topología pequeña cuyos nodos están ubicados en zonas con distintas intensidades de carbono (desde 23 gCO_2/kWh hasta 761 gCO_2/kWh, de verde a rojo, representando el aumento en la intensidad de carbono).

Si se desea enrutar un flujo desde el nodo *IN* hasta el nodo *OUT*, algoritmos que priorizan la latencia, como el de camino más corto, o aquellos que buscan minimizar el consumo energético, elegirían en primer lugar el camino rojo, el más contaminante, ya que es el que requiere menor número de saltos y de elementos consumiendo energía. Sin embargo, dos rutas alternativas serían descartadas por su mayor longitud o requisitos energéticos. No obstante, dichas rutas son significativamente menos contaminantes, lo que lleva a unas emisiones globales menores que, a pesar del mayor consumo energético, resultan en un camino más limpio.

Todos estos datos se toman como entrada para la solución. Como salida, la solución indicará, para cada uno de los puertos, el modo de operación que deben adoptar o si deben ponerse a dormir directamente; y, para cada uno de los flujos, el camino por el que será enrutado. Con esta información, es posible calcular el consumo energético en cada nodo y, por tanto, las emisiones totales de carbono emitidas a la atmósfera por la red.

IV. DEFINICIÓN DEL PROBLEMA

En esta sección, se describe la formulación del problema. Inicialmente, se definen los parámetros y variables necesarios, y a continuación se definen las restricciones y la función objetivo para comprender el funcionamiento de la solución propuesta.

A. *Parámetros*

La red se representa como un grafo bipartito $G = (\mathcal{N}, \mathcal{L})$, donde $\mathcal{N}$ representa el conjunto de nodos físicos y $\mathcal{L}$ denota el conjunto de enlaces dirigidos dentro de la red. Cada nodo $n \in \mathcal{N}$ está limitado por una capacidad de procesamiento C, que representa el ancho de banda máximo que cada router puede manejar. Además, se define b_l como la capacidad de ancho de banda (Mbps) del enlace $l \in \mathcal{L}$, s_l como el nodo origen del enlace l, y d_l como el nodo destino del enlace l.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En cuanto al consumo energético, la formulación requiere de $\mathcal{P}_l$, $\mathcal{E}_{idle}$ y $\mathcal{E}_{max}$, que representan respectivamente el consumo energético de un enlace, el consumo energético en reposo del router y el consumo energético máximo, medidos en vatios (W). Si ahora dirigimos nuestra atención a los puertos, necesitamos conocer el número de modos en los que se puede modular su velocidad de operación ($\mathcal{M}$), así como el ancho de banda máximo de estos puertos ($\mathcal{B}$, medido en Mbps), lo cual nos permitirá calcular la capacidad de procesamiento del router, $C = \mathcal{B} \cdot N_{ports}$. En este punto podemos evaluar otro de los parámetros esenciales de la solución, la potencia dinámica incremental [6], $\delta = \frac{\mathcal{E}_{max} - \mathcal{E}_{idle}}{C}$. Esta métrica evalúa la variación en el consumo energético de un router bajo carga, y se define como la potencia adicional requerida por cada megabit procesado (W/Mbps).

El conjunto de flujos se define como $\mathcal{F}$, y de cada uno de ellos la solución requiere conocer su nodo origen y destino, $s_f \in \mathcal{N}$ y $d_f \in \mathcal{N}$, así como su volumen b_f, expresado en Mbps. Finalmente, se necesita la intensidad de carbono [6] de cada nodo (λ_i), definida como la cantidad de carbono emitida por unidad de energía consumida en los distintos nodos (gCO_2/kWh). Este valor se deriva del conjunto de fuentes de energía utilizadas para generar electricidad. Como salida de la solución, se proporcionan las emisiones de carbono totales, una métrica distinta que resulta del producto de la energía consumida en cada nodo por la intensidad de carbono del área en la que se encuentra ubicado dicho nodo.

B. Variables

La optimización propuesta contiene dos conjuntos de variables:

- $\theta^f_{i,j}$ es una variable binaria que toma el valor 1 cuando el flujo f utiliza el enlace que conecta el nodo i con el nodo j, y toma el valor 0 si no utiliza dicho enlace.
- $\mu_{i,j}$ es una variable entera que almacena el modo operacional del enlace que conecta el nodo i con el nodo j. Puede tomar valores entre 0, lo que indica que el puerto está en modo de suspensión, y $\mathcal{M}$.

C. Función objetivo y restricciones

Tras la definición de los parámetros y variables clave involucrados en el problema, esta subsección presenta la formulación de la función objetivo (minimizar las emisiones totales de carbono de la red) junto con sus restricciones asociadas. El problema de optimización se define formalmente de la siguiente manera:

$$\textbf{Obj}: \textit{Minimizar} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \mu_{i,j} \cdot \frac{P_l}{\mathcal{M}} \cdot (\lambda_i + \lambda_j) + \sum_{i\in\mathcal{N}} \left(\sum_{f\in\mathcal{F}} \left(\begin{cases} b_f \cdot \sum_{j\in\mathcal{N}} \theta^f_{i,j} & \text{if} i \neq d_f \\ b_f & \text{if } i = d_f \end{cases} \right) \cdot \delta \cdot \lambda_i \right) \quad (1)$$

$$\text{s.t.:} \sum_{n\in\mathcal{N}} \theta^f_{s_f,n} - \theta^f_{n,s_f} = 1, \quad \forall f \in \mathcal{F} \quad (2)$$

$$\sum_{n\in\mathcal{N}} \theta^f_{d_f,n} - \theta^f_{n,d_f} = -1, \quad \forall f \in \mathcal{F} \quad (3)$$

$$\sum_{j\in\mathcal{N}} \theta^f_{i,j} - \theta^f_{j,i} = 0, \quad \forall f \in \mathcal{F},\ \forall i \in \mathcal{N} \setminus \{s_f, d_f\} \quad (4)$$

$$\sum_{f\in\mathcal{F}} \theta^f_{s_l,d_l} \cdot b_f \leq \min(\mu_{s_l,d_l} \cdot \frac{\mathcal{B}}{\mathcal{M}}, b_l), \quad \forall l \in \mathcal{L} \quad (5)$$

$$\mu_{s_l,d_l} \begin{cases} = 0 & \text{if } b_l = 0 \\ \leq \text{ceil}(b_l \cdot \frac{\mathcal{M}}{\mathcal{B}}) & \text{else} \end{cases}, \quad \forall l \in \mathcal{L} \quad (6)$$

$$\sum_{f\in\mathcal{F}} \left(\sum_{j\in\mathcal{N}} \theta^f_{i,j} \cdot b_f \right) + \sum_{\substack{f\in F \\ s_f = i}} b_f \leq \mathcal{C}, \quad \forall i \in \mathcal{N} \quad (7)$$

La primera ecuación, Ec. (1), representa la función objetivo del problema, cuyo objetivo es minimizar las emisiones globales de carbono. Esta ecuación está compuesta por dos partes diferenciadas, que deben minimizarse en conjunto, no por separado. La primera parte de la Ec. (1) representa las emisiones producidas por los puertos activos, y la segunda parte representa las emisiones generadas por el procesamiento realizado en los routers.

Específicamente, las emisiones relacionadas con los enlaces (*emisiones de enlace*) se calculan mediante la suma del consumo energético del enlace (suma del consumo de sus dos puertos, el cual depende del modo de operación de velocidad de estos) multiplicado por la intensidad de carbono del enlace (suma de las intensidades de carbono en la ubicación del nodo origen y del nodo destino del enlace). Por otro lado, las emisiones producidas por los nodos se calculan como el tráfico procesado por cada nodo multiplicado por la potencia dinámica incremental y la intensidad de carbono en la ubicación del nodo, considerando como tráfico procesado todo el tráfico que entra al nodo y todo el tráfico para el cual el nodo es el router origen. El objetivo de esta función es, por tanto, minimizar las emisiones globales derivadas del funcionamiento de la red.

El resto de la formulación se basa en definir las restricciones necesarias para plantear el problema. Las Ecs. (2), (3) y (4) garantizan la conservación de la dinámica de los flujos [15] para cada uno de los flujos en la red. La Ec. (5) asegura que el ancho de banda total que pasa por cada enlace (al enrutar todos los flujos a través de la red) no supere la capacidad del enlace ni el límite establecido en sus puertos extremos.

La ventaja de esta formulación frente a simplemente encender o apagar enlaces radica en que podemos limitar los puertos a los que está conectado cada enlace. Sin embargo, como que no todos los enlaces tienen siempre la misma capacidad, no todos sus puertos extremos pueden tener como límite superior el modo operacional $\mathcal{M}$; por tanto, este límite debe ajustarse. Esto se logra mediante la Ec. (6), que asigna un valor de cero a los puertos ïne-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

xistentes", es decir, aquellos que no tienen ningún enlace conectado, y para los puertos con enlaces establecidos, se determina como límite superior el primer modo operativo cuya capacidad sea igual o superior al ancho de banda del enlace, permitiendo así la utilización de dicho enlace desde un 0 % (modo de suspensión) hasta un 100 %.

Finalmente, es necesario formular las limitaciones computacionales de los dispositivos de red. Esto se expresa mediante la Ec. (7), que garantiza que la carga de procesamiento del router no exceda el 100 % de sus capacidades ($\mathcal{C}$). La carga de procesamiento se calcula como la suma de todos los flujos entrantes al router, junto con los flujos para los cuales el router actúa como primer salto. Esto asegura que se contabilicen todos los flujos cuyo origen o destino sea el router, evitando al mismo tiempo la inclusión redundante de los valores asociados a los flujos que simplemente lo atraviesan. Esto se logra considerando únicamente los flujos salientes para los cuales el router evaluado es el router de origen. Los flujos que pasan por el router como intermediario no se consideran, ya que ya han sido tenidos en cuenta en los cálculos de flujo entrante.

V. RESULTADOS EXPERIMENTALES

Esta sección evalúa las soluciones propuestas en un entorno simulado sobre topologías reales. El objetivo de esta evaluación es comparar la solución óptima propuesta con otras dos soluciones en términos de emisiones de carbono, ahorro energético, escalabilidad y tiempo de cómputo. La primera de las soluciones con las que comparamos nuestra propuesta es una solución basada en ILP que contempla todos los puertos de red activos (All Network Ports Active, ANPA) y enruta los flujos siguiendo la regla del camino más corto primero. La segunda es nuestro algoritmo ECOWN [16], una solución basada en ILP que apaga enlaces de forma óptima mientras realiza simultáneamente el enrutamiento óptimo, a diferencia de la heuristica propuesta en [6], trabajo en el que nos basamos al utilizar su modelo de consumo energético como y sus métricas como base.

La configuración utilizada para la evaluación de la solución se describe en la Sec. A. Posteriormente, los hallazgos se presentan y analizan en la Sec. B, donde se abordan los resultados de las diferentes soluciones y los resultados del análisis de escalabilidad de las pruebas.

A. *Configuración de la simulación*

Para evaluar la eficacia de la solución propuesta, realizamos una serie de experimentos en múltiples escenarios, cada uno utilizando configuraciones de red distintas, y en todos los casos utilizando el solucionador ILP Gurobipy [17].

A lo largo de las simulaciones realizadas, se examinan específicamente tres topologías de red distintas, cada una con una escala diferente: Abilene (12 nodos, 30 enlaces), Nobel (17 nodos, 52 enlaces) y Geant (22 nodos, 72 enlaces), las cuales se obtienen de [18]. Para cada topología se consideran cinco matrices de tráfico, que contienen flujos de datos para cada par de nodos en la red. Estas matrices (TM1–TM5) se obtienen escalando la matriz de tráfico pico extraída del conjunto de datos disponible en [18] (TM5), utilizando factores de reducción de 0,4, 0,5, 0,7 y 0,9, respectivamente.

Topología	**Consumo en reposo/máximo (W)**	**Ancho de banda (Gbps)**	δ **(W/Mbps)**
Abilene	187/205.25	400	$4.56{\cdot}10^{-5}$
Nobel	187/214.37	600	$4.56{\cdot}10^{-5}$
Geant	187/223.5	800	$4.56{\cdot}10^{-5}$

Cuadro I: Características de los routers para cada topología.

Topología	**Mínima**	**Máxima**	**Media**	**Mediana**
Abilene	24	660	344.25	337.5
Nobel	471	471	471	471
Geant	23	761	312.77	309

Cuadro II: Estadísticas de intensidad de carbono para cada topología (gCO2/kWh).

Las características del router utilizadas para las pruebas se derivan del router Arista 7060CX-32S [19]. Este dispositivo tiene un rendimiento de 3,2 Tbps, puertos de 100 Gbps con un consumo energético de 4,5 W, y un consumo de 187 W en reposo y 333 W en carga máxima. Como se describió anteriormente, estos datos se han ajustado para adaptarse a las diferentes topologías. En cuanto a los enlaces y flujos, se han escalado proporcionalmente hasta que el valor máximo de las matrices de tráfico coincida con la capacidad de los puertos del router.

En el lado de los routers, en lugar de considerar los 32 puertos reales del dispositivo, se considera un número de puertos igual al mayor número de vecinos que tenga cualquier nodo en la topología evaluada. Esto reduce el número de puertos del router contemplados en la ejecución, el rendimiento total del router y, por tanto, su consumo energético, que disminuye en la misma proporción que el número de puertos. Las características contempladas de los routers para cada una de las topologías se pueden ver en la Tabla I.

Otro aspecto a tener en cuenta es la elección del número de modos de operación que pueden tener los puertos. Un mayor número de modos hace que la red sea más precisa y eficiente, pero también vuelve mucho más compleja la computación de la solución. Sin embargo, un número muy reducido de estos, aunque disminuya el tiempo de cómputo, puede no ofrecer resultados suficientemente buenos. Según las pruebas realizadas en [7], cuatro modos de operación suelen ser la cantidad óptima para reducir significativamente el consumo energético sin afectar demasiado al tiempo de cómputo de las soluciones. Por lo tanto, en nuestros experimentos se consideran cuatro modos de operación.

Para obtener valores realistas de la intensidad de carbono (λ) en las zonas donde se encuentran los routers para cada una de las distintas topologías, utilizamos la API de ElectricityMaps [20], la cual, dado un conjunto de coordenadas, permite obtener la intensidad de carbono del área evaluada. La Tabla II muestra el valor mínimo, máximo, medio y mediano de la intensidad de carbono

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

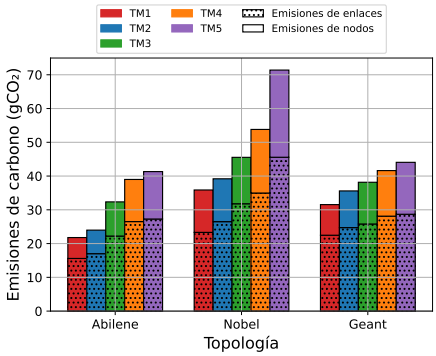


Figura 3: Emisiones de carbono para distintas matrices de tráfico en tres topologías diferentes.

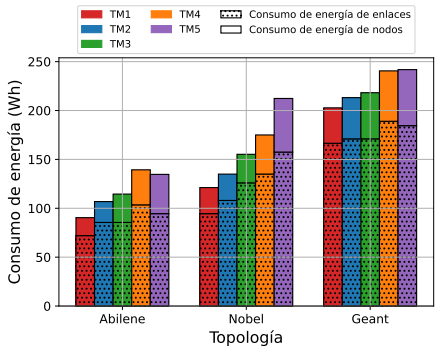


Figura 4: Consumo de energía para distintas matrices de tráfico en tres topologías diferentes.

entre todas las zonas cubiertas por cada una de las topologías. En particular, se puede observar cómo Geant es la topología menos contaminante, seguida de Abilene, y cómo Nobel, además de ser la más contaminante de todas, presenta la misma intensidad de carbono en todos sus nodos, ya que sus routers se encuentran en una misma zona. Esto proporciona menos flexibilidad a esta topología para minimizar sus emisiones de carbono.

B. Resultados de la evaluación

Para comenzar esta sección, se lleva a cabo una evaluación inicial del rendimiento con el fin de comprender cómo afectan al consumo energético y a las emisiones de carbono las dos partes de la Ec. (1). Cabe señalar que todas las figuras muestran el consumo energético o las emisiones de carbono producidas tanto por la energía de los puertos como por la potencia dinámica, es decir, se ignora la potencia estática, ya que no podemos extraer conclusiones a partir de ella.

A partir de las figuras, es importante destacar que todas las barras, independientemente de su color o patrón, parten de $y = 0$. Es decir, si una barra compuesta por dos segmentos con diferentes patrones alcanza una altura de 10, no implica que la suma de ambos segmentos sea 10; más bien, el segmento más alto alcanza ese valor. En resumen, las gráficas no apilan las barras para representar valores acumulativos, sino que las superponen para facilitar la comparación entre configuraciones.

La Figura 3 muestra los resultados de nuestra propuesta en términos de carbono emitido a la atmósfera tras enrutar los flujos de las diferentes matrices de tráfico para cada una de las topologías. Cada grupo de barras diferencia entre las distintas topologías, y cada barra dentro del grupo representa una matriz de tráfico distinta, siendo la barra más a la izquierda la generada por TM1 y la más a la derecha la generada por TM5. Si observamos la leyenda, las barras también distinguen entre las emisiones de carbono producidas por la primera parte de la Ec. (1), que representa las emisiones generadas por mantener enlaces activos con mayor o menor potencia (*Emisiones de enlaces*), y las emisiones de carbono producidas por la segunda parte de la ecuación, que representan las emisiones generadas por el enrutamiento de los flujos en los routers (*Emisiones de nodos*). En promedio, entre todas las pruebas, las emisiones de los enlaces son un 105.79 % superiores a las emisiones de los nodos.

Como puede observarse en la Figura 3, las emisiones de carbono aumentan con la carga de la red. A medida que aumenta la carga, también lo hace el modo de operación promedio de los puertos, así como la carga de procesamiento de los propios routers, lo que incrementa el consumo energético y, en consecuencia, las emisiones de carbono. Concretamente, las emisiones entre TM1 y TM5 aumentan un 89.7 %, 99.01 % y 39.7 % para Abilene, Nobel y Geant, respectivamente. Sin embargo, escalar la topología no siempre conlleva un aumento de las emisiones de carbono, ya que si el área donde se encuentra la red utiliza energía menos contaminante, incluso con un mayor consumo energético, las emisiones de carbono pueden ser más bajas, como ocurre en el caso de Geant en comparación con Nobel. Por otro lado, la Figura 4 muestra las mismas pruebas que la figura anterior, pero desde el punto de vista del consumo energético. Esta figura ilustra claramente que las emisiones de carbono no escalan de forma lineal con el consumo de energía. Las barras mostradas no siguen un patrón consistente, ya que la solución no está diseñada para optimizar esta métrica, llegando incluso a obtener consumos energéticos más bajos en ejecuciones con mayor carga de tráfico (por ejemplo, Abilene TM5 frente a TM4).

Además, la figura resalta el ejemplo previamente discutido al comparar las emisiones de Nobel y Geant. Aunque Geant consume más energía que Nobel, su ubicación en una región con menor intensidad de carbono da como resultado unas emisiones totales de carbono más bajas.

El siguiente análisis comparan las tres soluciones referenciadas al inicio de la sección, destacando las diferencias entre sus emisiones de carbono y su consumo de energía.

En primer lugar, la Figura 5 muestra específicamente las emisiones de carbono para cada TM y cada topología al utilizar nuestra solución (con tasa operacional), ANPA y ECOWN. Las barras con patrón de puntos son básicamente las mismas que aparecen en la Fig. 3. Podemos observar cómo las barras con patrón de líneas, que representan las

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

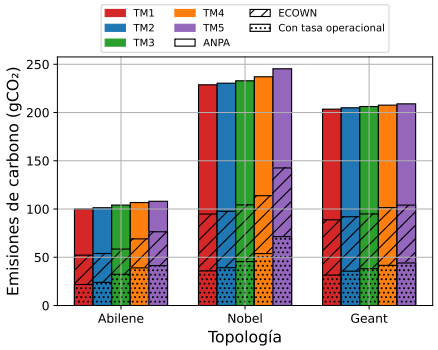


Figura 5: Emisiones de carbono para diferentes matrices de tráfico, diferentes topologías y diferentes soluciones.

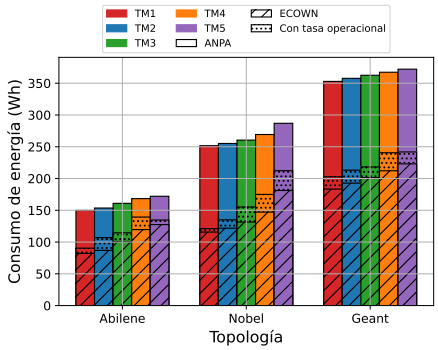


Figura 6: Consumo de energía para diferentes matrices de tráfico, diferentes topologías y diferentes soluciones.

ejecuciones de ECOWN, resultan en mayores emisiones. La preocupación se intensifica al examinar las barras sólidas (ANPA), ya que estas incrementan sustancialmente las emisiones. De media para Abilene, Nobel y Geant, nuestra propuesta reduce las emisiones con respecto a ECOWN en un 48.91 %, 55.55 % y 60.31 %. Si comparamos con ANPA, las reducciones de emisiones alcanzan el 69.54 %, 79.07 % y 81.48 %.

Por otro lado, la Fig. 6 persigue el mismo objetivo que la Fig. 5, pero aplicado al consumo energético. Esta figura muestra algo curioso: la posición entre la solución propuesta y ECOWN se ha intercambiado, lo que indica que, aunque nuestra propuesta reduce significativamente las emisiones de carbono en comparación con ECOWN, requiere una mayor entrada de energía. En promedio, nuestra solución disminuye el consumo energético en comparación con ANPA en un 27.21 %, 39.65 % y 38.40 %, respectivamente para Abilene, Nobel y Geant. Sin embargo, la comparativa de nuestra propuesta con respecto a ECOWN muestra un aumento del consumo energético en un 13.27 %, 14.55 % y 10.24 %, respectivamente.

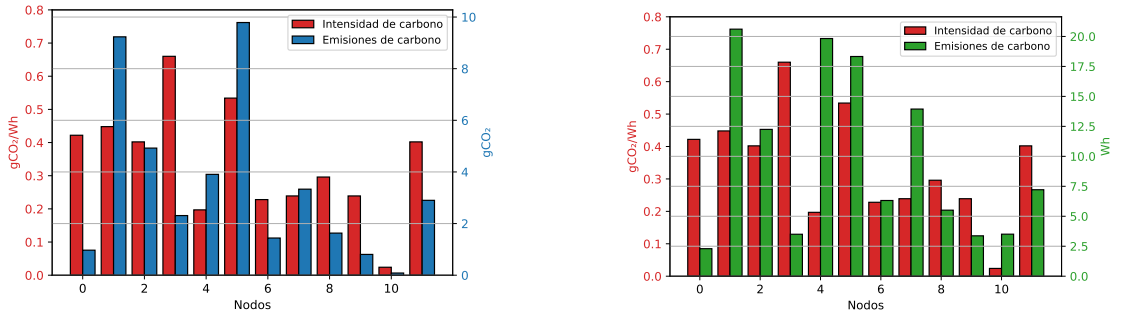
Las última figura que se analiza es la Fig. 7, la cual tienen como objetivo reforzar la relación entre el consumo energético y las emisiones de carbono, influenciada por la intensidad de carbono. La figura se centra en la TM5 de Abilene y presenta, para cada nodo de la topología, la intensidad de carbono de su área geográfica, la energía que ha consumido y las emisiones de carbono resultantes. La subfigura (a) relacionan los valores de intensidad de carbono por nodo con las emisiones en cada nodo. Por otro lado, las subfiguras (b) muestran, además de los valores de intensidad de carbono por nodo, la cantidad de energía consumida por cada uno de ellos. En resumen, al referirnos a las emisiones de carbono, nos centraremos en la subfigura (a); al referirnos al consumo energético, consideraremos la subfigura (b); y al hablar de la intensidad de carbono, se podrá usar cualquiera de las dos subfiguras.En esta figura, los datos resumidos de la Tabla II se visualizan en un formato desagregado (dividido por nodos). Estos valores de intensidad de carbono determinan la relación entre el consumo de energía y las emisiones de carbono: cuanto mayor es la intensidad de carbono, mayores son las emisiones de carbono por unidad de energía consumida. Por ejemplo, consumir 1 kW de potencia en el nodo 10 resulta en 24 gCO_2, mientras que la misma cantidad de energía en el nodo 3 genera 660 gCO_2. A partir de esta observación, es lógico pensar que los nodos más utilizados deberían ser aquellos con menor intensidad de carbono, pero esto no siempre es posible, como en el caso del nodo 10, ya que es un nodo externo y cargarlo implica caminos más largos y no descargar otros nodos. Lo mismo ocurre con el nodo 5, que tiene una alta intensidad de carbono, pero no es rentable descargarlo al ser un nodo central. Aparte de estos casos excepcionales, la tendencia general es asignar más carga a nodos con menor intensidad de carbono (como los nodos 1, 2, 4 y 7) con el fin de descargar los nodos más contaminantes, como el nodo 3. La relación entre estas tres métricas puede observarse claramente al comparar los nodos 4 y 5. Aunque ambos presentan un consumo energético similar (19.92 W y 18.33 W), sus emisiones de carbono difieren significativamente (3.91 gCO_2 y 9.79 gCO_2), lo que resalta la importancia de utilizar fuentes de energía con bajas emisiones para reducir la huella de carbono total de la red.

Finalmente, es importante evaluar los tiempos de ejecución requeridos por nuestro enfoque para obtener una solución. Para TM1, TM2, TM3, TM4 y TM5, respectivamente, el tiempo medio de ejecución entre todas las topologías es de 174.49 s, 333.71 s, 281.24 s, 870.11 s y 360.21 s. Sin embargo, si calculamos para cada topología la duración media de todas las TMs, los tiempos de ejecución son 1.54 s, 134.92 s y 1075.382 s. Como puede observarse, el aumento de la carga de tráfico, en términos de tiempo, no sigue ningún patrón, a diferencia de la escalabilidad de la topología, que incrementa el tiempo de resolución de forma exponencial.

VI. CONCLUSIONES

En este trabajo, nos centramos en el problema de las emisiones de carbono en redes y estudiamos los parámetros que pueden afectar a esta métrica. Además, hemos adaptado un modelo de las emisiones de carbono producidas por los dispositivos de red para tener en cuenta los modos de operación que pueden tener los puertos de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) Intensidad de carbono frente a emisiones de carbono

(b) Intensidad de carbono frente a consumo de energía

Figura 7: Resultados por nodo para Abilene

estos dispositivos.

Asimismo, propusimos una solución basada en ILP para encontrar la configuración óptima de la red, indicando tanto los modos de operación de cada puerto como el enrutamiento de los flujos. También estudiamos cómo afecta la escalabilidad del problema a los resultados obtenidos por nuestra solución, escalando tanto el tamaño de la topología de red como la carga de tráfico que circula por estas redes.

Tras analizar detenidamente los resultados obtenidos, nuestra solución es capaz de reducir las emisiones de carbono de la red hasta en un 84.5 %, dependiendo de la topología de red y de la carga de tráfico. Como trabajo futuro, planeamos desarrollar algoritmos heurísticos o basados en aprendizaje automático que permitan obtener soluciones casi óptimas en tiempos tratables, sin comprometer significativamente su efectividad.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades (proyecto PID2021-124054OB-C31) y por la Unión Europea, Fondo Europeo de Desarrollo Regional, y la Junta de Extremadura. Autoridad de Gestión: Ministerio de Hacienda. Número de expediente de la ayuda: GR24099.

REFERENCIAS

[1] C. Lange, D. Kosiankowski, C. Gerlach, F.-J. Westphal, and A. Gladisch, "Energy consumption of telecommunication networks," in *2009 35th European Conference on Optical Communication*, 2009, pp. 1–2.

[2] C. Lange, D. Kosiankowski, R. Weidmann, and A. Gladisch, "Energy consumption of telecommunication networks and related improvement options," *IEEE Journal of selected topics in quantum electronics*, vol. 17, no. 2, pp. 285–295, 2010.

[3] S. Bouckaert, A. F. Pales, C. McGlade, U. Remme, B. Wanner, L. Varro, D. D'Ambrosio, and T. Spencer, "Net zero by 2050: A roadmap for the global energy sector," 2021.

[4] L. Belkhir and A. Elmeligi, "Assessing ict global emissions footprint: Trends to 2040 & recommendations," *Journal of cleaner production*, vol. 177, pp. 448–463, 2018.

[5] M. Silva and H. L. Raadal, "Life cycle ghg emissions of renewable and non-renewable electricity generation technologies," *Krakeroy: Ostfoldforskning*, vol. 6, 2019.

[6] S. El-Zahr, P. Gunning, and N. Zilberman, "Exploring the benefits of carbon-aware routing," *Proceedings of the ACM on Networking*, vol. 1, no. CoNEXT3, pp. 1–24, 2023.

[7] J. Galan-Jimenez and A. Gazo-Cervero, "Elee: Energy levels-energy efficiency tradeoff in wired communication networks," *IEEE communications letters*, vol. 17, no. 1, pp. 166–168, 2012.

[8] S. H. Haji, S. R. Zeebaree, R. H. Saeed, S. Y. Ameen, H. M. Shukur, N. Omar, M. A. Sadeeq, Z. S. Ageed, I. M. Ibrahim, and H. M. Yasin, "Comparison of software defined networking with traditional networking," *Asian Journal of Research in Computer Science*, vol. 9, no. 2, pp. 1–18, 2021.

[9] J. Galán-Jiménez, M. Polverini, J. L. Herrera, F. G. Lavacca, and J. Berrocal, "Elto: Energy efficiency-load balancing trade-off solution to handle with conflicting metrics in hybrid ip/sdn scenarios," *IEEE Transactions on Network and Service Management*, pp. 1–1, 2025.

[10] J. Galán-Jiménez, M. Polverini, F. G. Lavacca, J. L. Herrera, and J. Berrocal, "Joint energy efficiency and load balancing optimization in hybrid ip/sdn networks," *Annals of Telecommunications*, vol. 78, no. 1-2, pp. 13–31, 2023.

[11] S. Tabaeiaghdaei, S. Scherrer, J. Kwon, and A. Perrig, "Carbon-aware global routing in path-aware networks," in *Proceedings of the 14th ACM International Conference on Future Energy Systems*, ser. e-Energy '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 144–158. [Online]. Available: https://doi.org/10.1145/3575813.3595192

[12] H. Hua, J. Shi, X. Chen, Y. Qin, B. Wang, K. Yu, and P. Naidoo, "Carbon emission flow based energy routing strategy in energy internet," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3974–3985, 2024.

[13] S. Torkzadeh, H. Soltanizadeh, and A. A. Orouji, "Energy-aware routing considering load balancing for sdn: a minimum graph-based ant colony optimization," *Cluster Computing*, vol. 24, no. 3, pp. 2293–2312, 2021.

[14] Y. Wei, X. Zhang, L. Xie, and S. Leng, "Energy-aware traffic engineering in hybrid sdn/ip backbone networks," *Journal of Communications and Networks*, vol. 18, no. 4, pp. 559–566, 2016.

[15] L. Taccari, "Integer programming formulations for the elementary shortest path problem," *European Journal of Operational Research*, vol. 252, no. 1, pp. 122–130, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221716000084

[16] J. Gómez-delaHiz and J. Galán-Jiménez, "Ecown: An optimization framework for carbon-aware routing in wired networks," 2025, in press.

[17] "Gurobipy optimizer," , last accessed: 2025-05-26. [Online]. Available: https://pypi.org/project/gurobipy/

[18] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly, "SNDlib 1.0–Survivable Network Design Library," in *International Network Optimization Conference (INOC),*, April 2007, http://sndlib.zib.de, extended version accepted in Networks, 2009.

[19] A. Networks, "Arista 7060x and 7260x series datasheet," 2021, accessed: 2025-05-27. [Online]. Available: https://www.arista.com/assets/data/pdf/Datasheets/7060X_7260X_DS.pdf

[20] ElectricityMaps, "Carbon intensity data (version january 27, 2025)," 2025, accessed: 2025-5-27. [Online]. Available: https://electricitymaps.com/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Un enfoque simplificado de trazabilidad de criptoactivos con Page Rank RWR, PCA y K-means

Volodymyr Dubetskyy, Álvaro Hernández Riquelme, Adrián Ortega Escalona,
Pilar Manzanares López, Maria-Dolores Cano
Departamento de Tecnologías de la Información y las Comunicaciones
Universidad Politécnica de Cartagena

volodymyr.dubetskyy@upct.es, alvaro.hriquelme@upct.es, adrian.ortega@upct.es,
pilar.manzanares@upct.es, mdolores.cano@upct.es

El auge de las criptomonedas ha multiplicado los riesgos de blanqueo, exigiendo técnicas de trazabilidad que funcionen sin datos etiquetados. Este trabajo aborda la clasificación no supervisada de direcciones Bitcoin en el dataset Elliptic++, explotando únicamente la estructura del grafo transaccional. Aplicamos un PageRank con reinicio aleatorio personalizado para calcular dos probabilidades originales de proximidad a nodos lícitos e ilícitos. Estos vectores alimentan un clustering K-Means con k = 3. El sistema logra un índice Silhouette de 0,955, superando configuraciones con reducción de dimensionalidad y evitando el coste de modelos profundos. La incorporación de PCA no mejora la cohesión y un trimmed K-Means solo eleva la pureza de forma marginal mientras reduce la densidad interna, confirmando la autosuficiencia de las características derivadas de PageRank. La propuesta demuestra que la topología de la red basta para discriminar patrones ilícitos de forma interpretable y con mínimos requisitos computacionales, lo que facilita su adopción en entornos regulatorios.

***Palabras Clave*—bitcoin, blockchain, cryptocurrencies, cybersecurity, forensics, page rank, traceability**

I. INTRODUCCIÓN

Durante la última década, la capitalización agregada de los criptoactivos ha superado con holgura el billón de dólares estadounidenses (USD). Solo Bitcoin [1] ronda ya los USD 1,6 billones de valor de mercado [2], consolidándose como un activo financiero sistémico. En paralelo, redes de pago globales como Visa procesan cientos de millones en compras liquidadas directamente on-chain y han integrado esta modalidad en su tesorería [3]. Como aspecto negativo, este grado de adopción ha ido acompañado de un repunte de fraude, esquemas Ponzi y financiación ilícita. La Financial Action Task Force (FATF) sitúa hoy la monitorización on-chain en el núcleo de los marcos Know-Your-Customer / Anti-Money-Laundering (KYC/AML), con directrices específicas para los Virtual Asset Service Providers (VASP) y la Travel Rule [4].

Para detectar actividad ilícita en la red, la investigación se ha centrado en Graph Neural Networks (GNN) aplicadas al grafo de transacciones de Bitcoin. Estos modelos aprenden representaciones que permiten clasificar direcciones de monederos electrónicos (wallets), que son las claves públicas que identifican cada monedero en la red Bitcoin, según su comportamiento. Weber et al. [5] mostraron que un Graph Convolutional Network (GCN) alcanza un F1-score de 0,83 en el dataset Elliptic, y posteriores variantes basadas en Graph Attention Networks (GAT/GAT-ResNet) mejoran esa cifra [6], [7]. Sin embargo, las GNN exigen miles de nodos etiquetados, hardware GPU dedicado y enfrentan un desequilibrio de clases extremo (aproximadamente 92 % transacciones lícitas frente a menos del 8 % ilícitas), lo que dificulta su adopción industrial [8].

Las aproximaciones no supervisadas eliminan la dependencia de etiquetas, pero la literatura vigente suele emplear vectores de atributos locales o experimentaciones reducidas. Técnicas como k-medoids o spectral clustering se aplican sobre subconjuntos de transacciones sin explotar la conectividad global ni medir con rigor la coherencia de los clústeres [9].

Este trabajo aborda la clasificación no supervisada de direcciones Bitcoin empleando solo la topología del grafo transaccional. Las direcciones de wallet son los vértices del grafo transaccional sobre los que aplicamos el método propuesto. El flujo consta de dos etapas. Primero, aplicamos PageRank con reinicio aleatorio personalizado (Random Walk with Restart, RWR) [10], [11] para pro-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

pagar la influencia de un reducido conjunto de semillas lícitas e ilícitas y obtener dos puntuaciones (pl, pi) que cuantifican la proximidad topológica de cada dirección a cada tipo de conducta, lícita o ilícita, respectivamente. Después, agrupamos esas puntuaciones mediante K-Means con k=3 (lícito, ilícito, desconocido). Probamos este proceso en Elliptic++, ampliación pública del corpus Elliptic con más de 200 000 transacciones etiquetadas [6], [12]. El modelo alcanza un índice Silhouette de 0,955 sin reducción de dimensionalidad ni aprendizaje profundo. Además, los resultados muestran que la inclusión de un Análisis de Componentes Principales (PCA) no mejora la cohesión y la variante trimmed K-Means solo incrementa marginalmente la pureza, reduciendo la densidad interna de los grupos. Así, las principales contribuciones de este trabajo son:

1) Marco interpretable y ligero que demuestra que dos características derivadas de PageRank bastan para discriminar actividad ilícita con coste mínimo.

2) Evaluación comparando configuraciones sin PCA, con 2 y 4 componentes principales y con trimmed K-Means, evidenciando el impacto negativo de la reducción de dimensionalidad.

3) Se evidencia que las GNN no son imprescindibles para la trazabilidad cripto, aunque como se indica en la discusión de los resultados, este trabajo tiene limitaciones que serán abordadas en trabajos futuros.

La Sección 2 describe el dataset Elliptic++ y formaliza PageRank personalizado, PCA como técnica comparativa y K-Means. La Sección 3 presenta y discute los resultados. El documento finaliza con las conclusiones y líneas futuras, incluida la extensión a otras cadenas públicas y la incorporación de dinámicas temporales.

II. METODOLOGÍA

A. *Dataset Elliptic++*

Para el análisis, se ha decidido usar el Dataset de *Elliptic++* gracias a su gran volumen de direcciones clasificadas según tienen una naturaleza lícita, ilícita, o desconocida. El principal desafío de trabajar con direcciones es su falta de características, por lo que es conveniente trabajar con sus interacciones y comportamientos.

El dataset Elliptic++ constituye un recurso fundamental para la investigación en trazabilidad de criptoactivos. Contiene una cantidad considerable de transacciones de Bitcoin etiquetadas manualmente, organizadas en un grafo temporal donde los nodos representan transacciones y las aristas flujos monetarios. Su implementación garantiza reproducibilidad y proporciona un benchmark para futuras investigaciones en detección no supervisada de fraudes blockchain. En nuestro caso, para la simplificación del análisis, se ha optado por trabajar únicamente con las direcciones de wallet, que son las claves públicas que identifican cada monedero en la red Bitcoin. Estas direcciones son distintas a las de los nodos del grafo transaccional, por lo que se formará un grafo distinto y las aristas representarán las transacciones entre ellas.

Como el enfoque principal son las direcciones de wallets, habrán tablas y datos que no se usen del Dataset, como es el caso del grafo de interacciones de *Actors*, ya que no proporciona conexiones directas entre direcciones, y agrupa direcciones que podrían ser omitidas. Se usará el grafo de interacciones de *Address-Transaction* que contiene tablas cuyas columnas aportan información relevante para este tipo de algoritmos. En las Tablas I, II y III se muestran los datos más relevantes del dataset para estos algoritmos, que contienen las direcciones de wallet, e información adicional como sus interacciones con otras wallets y características. La tabla I contiene las conexiones entre direcciones con origen y destino, lo que consideramos como una transacción, aunque el análisis de naturaleza no se hace en las transacciones sino en las direcciones de wallet. La tabla II contiene la naturaleza de cada dirección, que puede ser lícita, ilícita o desconocida, y la tabla III contiene las características de cada dirección, como el número de transacciones realizadas, el grado de entrada y salida, o la media de comisiones pagadas, que nos serán útiles para la agrupación posterior.

Tabla I
ADDRADDR_EDGELIST.CSV

input_address	output_address
1A1dfg...kNa	3J9...34Ly
3J...89Az	bc1...aat4

Tabla II
WALLETS_CLASSES.CSV

address	class
1A1dgk..zP1a	lícita
3J98t...WNLy	ilícita
bc1q...rt4	desconocida

Tabla III
WALLETS_FEATURES.CSV

address	tx_count	in_degree	out_degree	...	fees_mean
1A1z...fNa	150	75	75	...	0.00012
3J9...NLy	200	100	100	...	0.00023
bc1q...8f3t4	120	60	60	...	0.00015

B. *Algoritmo PageRank RWR*

El algoritmo que se plantea es una especialización del PageRank RWR (*Random Walk with Restart*), basado en el algoritmo de PageRank original de Google [10] para clasificar páginas web según una importancia recibiendo un peso cada enlace. En este caso, se escoge una semilla con direcciones con la naturaleza definida, con un peso de 1 para cada dirección. Tras la primera iteración, las semillas iniciales se descartan y los nodos que han ganado en peso pasan a ser la nueva semilla. El peso es reducido con un factor de amortiguación α, lo que da mayor valor a los nodos pertenecientes a las primeras iteraciones.

Para empezar, se denotará la definición del estado de la red como S, un conjunto de nodos que actúa como la semilla. En cada iteración el conjunto de nodos S

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

es sustituido por las direcciones visitadas en la iteración anterior. Por tanto, en caso de que en la primera iteración se escoja como semilla el conjunto de direcciones ilícitas, $S^0 = \{I \in V\}$ y como para cada iteración se descarta la anterior, $S^i = \{U^i \in V | U^i \notin I\}$, donde V es el conjunto de direcciones del dataset y U^i el conjunto de direcciones que se comunica con la semilla de la iteración i.

Para cada estado, se forma una matriz de adyacencia A, donde A_{ij} será 1 si hay una conexión desde el nodo i hasta el nodo j, y 0 en caso contrario. Con estos resultados se forma la matriz diagonal D que representa el grado de salida para cada nodo, siendo $D_{kk} = \sum_i A_{ik}$ y $D_{ij} = 0$ en caso contrario.

Con estos valores, se podrá calcular la matriz de transición, que contiene las probabilidades de conexión de un nodo a otro nodo, con un factor de amortiguación de 0.85 en cada iteración, se definirá como:

$$W = \alpha(D^{-1}A) \tag{1}$$

El factor de amortiguación α juega un papel crucial en la distribución de estas probabilidades, ya que determina cuánto peso se transmite en cada iteración. Un valor más alto de α permitirá que la influencia se propague más lejos en la red, mientras que un valor más bajo concentrará el peso en los nodos más cercanos a las semillas iniciales.

Finalmente, se podrá calcular el vector de probabilidades p buscado, considerando la probabilidad inicial de las semillas:

$$p_i^0 = \begin{cases} \frac{1}{|S(0)|} & \text{si } i \in S(0) \\ 0 & \text{resto} \end{cases} \tag{2}$$

Normalizando y partiendo de p0, la probabilidad en las siguientes iteraciones se define como:

$$p_i^{t+1} = \begin{cases} \frac{1}{|W^T p^t|} & \text{si } i \in S^{(t+1)} \text{ y } |S^{(t+1)}| > 0 \\ 0 & \text{resto} \end{cases} \tag{3}$$

Se obtendrá un vector de probabilidades distinto para las semillas ilícitas y lícitas, por lo que las direcciones desconocidas tendrán de características p_l y p_i, que se emplearán como referencia para la clasificación. Una dirección con un alto valor de p_l y bajo p_i sugiere que tiene más interacciones con direcciones lícitas conocidas, mientras que lo contrario indicaría una mayor conexión con actividades ilícitas. Mientras que, si ambas son de alto valor o bajo valor, podría significar una naturaleza desconocida o una dirección con mucha actividad como mixers o exchanges. Esta dualidad de características permite crear un espacio bidimensional donde cada dirección puede ser representada como una coordenada, facilitando la aplicación posterior de algoritmos de clustering.

El vector de probabilidades obtenido se usará posteriormente como características para la agrupación de direcciones o para la reducción de dimensionalidad, junto a las características base que aparecen en el dataset. En la Fig. 1 se muestra la distribución teórica de los pesos en la red, donde los nodos principales son más importantes y los periféricos van perdiendo peso.

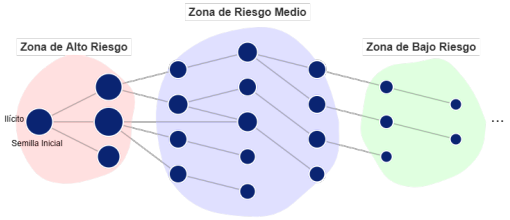


Figura 1. Distribución teórica de PageRank.

C. *Algoritmo PCA*

Como se menciona en la introducción, la escasez de características intrínsecas asociadas a las direcciones de wallets plantea un desafío significativo para el análisis, ya que, a diferencia de los nodos transaccionales que incluyen atributos financieros detallados, las direcciones carecen de metadatos descriptivos. Para abordar esta limitación, se implementa el Análisis de Componentes Principales (PCA), una técnica de reducción de dimensionalidad que transforma variables correlacionadas en un conjunto ortogonal de componentes principales. Estos componentes capturan la máxima varianza de los datos originales mediante una proyección lineal, priorizando la información más discriminativa mientras minimiza el ruido y la redundancia.

Aunque PCA mejora la eficiencia computacional y mitiga correlaciones espurias, su aplicación conlleva inconvenientes. Por un lado, simplifica la estructura de datos, facilitando la convergencia de K-Means y reduciendo el riesgo de sobreajuste. En cambio, la compresión dimensional puede eliminar información relevante no linealmente correlacionada, lo que —como revelan posteriormente los resultados— degrada la cohesión de los clústeres en este contexto específico. Esta dualidad subraya la necesidad de comparar sus resultados frente a enfoques basados únicamente en características topológicas.

Las características que más han aportado al algoritmo vienen en el dataset con las direcciones, y son algunas como `btc_transacted_max` para la cantidad maxima de BTC enviada o `fees_mean` para la media de commisiones de red pagadas.

A falta de determinar el número de componentes óptimo, comentado en la sección de resultados, este algoritmo no sólo sirve para reducir las dimensiones, si no que facilita algunos datos que se usarán en conjunto con K–Means, reduciendo ruido y correlaciones que pueden afectar a la agrupación, junto a mejoras de la estabilidad, reduciendo la probabilidad de convergencia en mínimos locales. [9] [13]

D. *Agrupación con K–Means*

El algoritmo de K–Means es un método de agrupación no supervisado que busca dividir un conjunto de datos en k grupos o clústeres, donde cada clúster se representa por

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

su centroide. El objetivo es minimizar la varianza dentro de cada clúster y maximizar la varianza entre los distintos clústeres. Se intentará agrupar los nodos desconocidos en tres grupos según la naturaleza y obtener el menor número de desconocidos reales.

Cada clúster C_j se representa por un centroide μ_j, que es el punto "medio" de todos los datos asignados a ese clúster. Formalmente, si $C_j = \{x_i \,|\, \text{etiqueta}(x_i) = j\}$, el centroide se define como

$$\mu_j \;=\; \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i, \quad \mu_j \in \mathrm{R}^d \tag{4}$$

Cada iteración de K–Means consta de dos fases:

- *Asignación:* cada punto x_i se asigna al clúster cuyo centroide μ_j minimiza $\|x_i - \mu_j\|$.
- *Actualización:* recalcular cada μ_j usando la fórmula anterior con los puntos recién asignados.

Como los centroides parten de posiciones aleatorias, diferentes semillas pueden llevar a converger a soluciones distintas (mínimos locales).

En trimmed K–Means, tras excluir los puntos más lejanos se recalculan los centroides solo con el subconjunto "internalizado", lo que incrementa la robustez frente a outliers, ya que μ_j no se ve desplazado por valores extremos.

Finalmente se comparará las métricas de evaluación con distintos componentes de PCA, y la diferencia de resultados entre K–Means y trimmed K–Means.

III. RESULTADOS

A. *Resultados tras la ejecución de PageRank*

Una vez ejecutado el algoritmo de PageRank RWR, se obtienen dos vectores de probabilidades p_l y p_i que representan la proximidad de cada dirección a las semillas lícitas e ilícitas, respectivamente. Estos vectores se normalizan para que la suma de sus valores sea 1, facilitando su interpretación como probabilidades. Es importante destacar que el vector de probabilidades aplica también para las semillas, puesto que es una forma de representar las direcciones totales del Dataset y su naturaleza, aunque la importancia y los resultados de los distintos algoritmos se obtienen para las direcciones desconocidas, que son las que no tienen una naturaleza definida en el dataset.

En la Fig. 2 se puede observar la distribución de direcciones según tengan mayor peso en cada clasificación. La zona central, con mayor probabilidad de ser desconocida, estará compuesta por nodos con muy bajo peso (o ninguno) en los dos posibles, mientras que en las demás existirá una descompensación entre un tipo y otro, siendo la zona de la derecha, donde se encuentran las lícitas, más poblada.

B. *Resultados tras la ejecución de PCA*

A continuación, se procede a aplicar PCA con respecto a las características del dataset, que son las que se muestran en la Tabla III. El objetivo es reducir la dimensionalidad de los datos y extraer las características más relevantes para la agrupación posterior con K–Means. No se han usado las características p_l y p_i obtenidas de PageRank, ya que se considera que no están correlacionadas con las demás características del dataset, por lo que se tomará como características por separado.

La Fig. 3 muestra una comparativa de número de componentes con respecto a la varianza acumulada, con la finalidad de encontrar un número óptimo de componentes que represente la mayoría de la varianza. Se puede observar que los componentes más importantes son el primero y el segundo componente, ya que entre estos dos se consigue un 40 % de varianza acumulada, mientras que los demás aportan menor cantidad. Ampliando de 2 a 4 componentes, se alcanzaría una varianza acumulada de mas del 50 %. A partir de ahí, y no se estaría añadiendo información relevante suficiente para justificar la reducción de precisión que se introduciria a K–Means.

C. *Implementación y función objetivo*

Una vez calculada toda la información necesaria, el algoritmo de K–Means se aplicará al espacio de características definido por las dos probabilidades (p_l, p_i) obtenidas con PageRank y las m componentes principales seleccionadas por PCA. También se fijará $k = 3$ grupos (lícito, ilícito y mixto/desconocido) con varias iteraciones desde la inicialización para mitigar mínimos locales.

En trimmed K–Means se introduce un parámetro de recorte α, con α = 0.2, que en cada iteración descarta el α % de puntos más alejados de sus centroides.

En la implementación del algoritmo, se obtiene una etiqueta para cada x_i que marca al clúster que se le ha asignado. La Fig. 4 muestra la clasificación obtenida con K–Means, considerando los dos primeros componentes y las probabilidades de PageRank. Se trata de una representación tridimensional de los datos, mostrando $PC2$, p_i y p_l. Se puede observar la importancia de PageRank en la separación de clústers, ya que los centroides de los clústeres se encuentran en una zona donde las direcciones lícitas tienen un alto valor de p_l y bajo p_i en el caso de las lícitas, mientras que en las ilícitas pasa de forma contraria.

Una vez clasificados, es conveniente usar métricas como Silhouette score o pureza global para evaluar la calidad de los clústeres obtenidos. La pureza global se define como la proporción de puntos correctamente clasificados en cada clúster. En este caso, también se añadirá métricas como Calininski-Harabasz y Davies-Bouldin, que miden la separación entre clústeres y la compactación dentro de ellos.

D. *Resultados esperados y métricas de evaluación*

La tabla IV contiene la comparación entre las 4 métricas para los casos estudiados, esto es considerando 0, 2 y 4 componenetes principales, ya que según las conclusiones obtenidas en la figura 3, son las opciones que más pueden aportar sin reducir demasiado la precisión. Para todas estas opciones, se empleará K-means y trimmed K-means. Se puede observar como los mejores resultados se obtienen al no considerar componentes principales (m=0). En el caso de usar PCA, se observa que la pureza global y el Silhouette score disminuyen al usar la variante trimmed,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

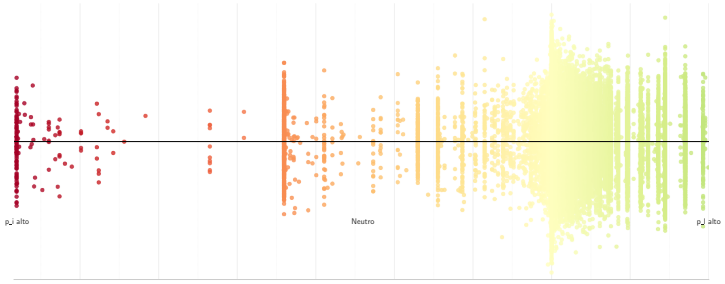


Figura 2. Distribución de conexiones con semilla ilícita.

mientras que el índice Calinski-Harabasz aumenta, lo que indica una mejor separación entre clústeres. Estos resultados sugieren que, para el caso propuesto, la reducción de dimensionalidad, mediante la reducción de características, no llega a mejorar una métricas basadas únicamente en los pesos de PageRank.

Tabla IV
COMPARATIVA: K–MEANS VS. TRIMMED K–MEANS

Métrica	K–Means	trimmed K–Means
Silhouette score (NO PCA)	**0.955**	0.941
Silhouette score (PCA 2)	0.868	0.891
Silhouette score (PCA 4)	0.771	0.667
Pureza global (NO PCA)	0.712	**0.715**
Pureza global (PCA 2)	0.712	0.692
Pureza global (PCA 4)	0.714	0.667
Calinski–Harabasz (NO PCA)	865274.6	**2131396.8**
Calinski–Harabasz (PCA 2)	133226.3	740376.9
Calinski–Harabasz (PCA 4)	72527.5	180458.7
Davies–Bouldin (NO PCA)	**0.358**	0.386
Davies–Bouldin (PCA 2)	0.490	0.440
Davies–Bouldin (PCA 4)	0.968	0.646

Aunque el flujo propuesto es eficaz en Elliptic++ [6], su validez externa es incierta al no probarse en otras cadenas (p. ej. Ethereum) ni en datasets con diferentes distribuciones. Además, depende de semillas lícitas e ilícitas cuyo sesgo puede afectar las puntuaciones (pl, pi), cuya sensibilidad a selección y tamaño debe evaluarse en futuros trabajos. La evaluación se limita a métricas internas (Silhouette, pureza, densidad), por lo que hace falta verificar su precisión con etiquetas de referencia o en entornos KYC/AML reales. Finalmente, el análisis usa un grafo estático y omite dimensión temporal y atributos financieros, que podrían mejorar la detección de conductas

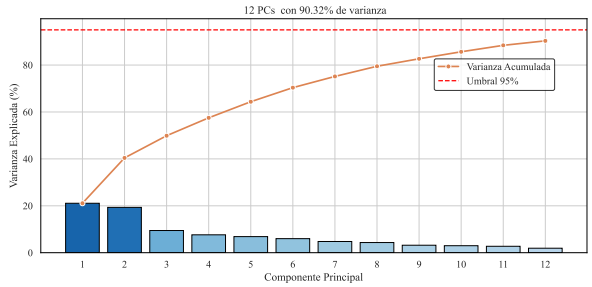


Figura 3. Varianza acumulada por número de componentes principales.

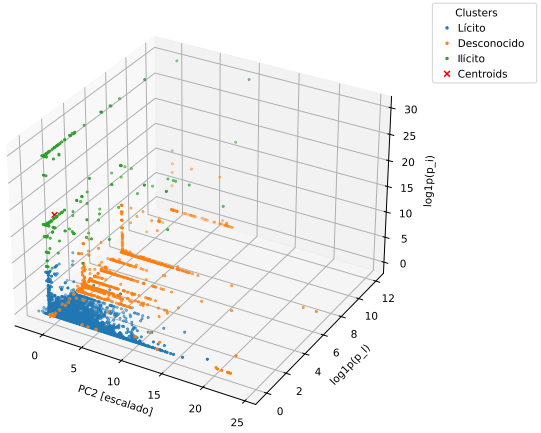


Figura 4. Varianza acumulada por número de PCs.

ilícitas en extensiones futuras.

IV. CONCLUSIONES

Los resultados muestran que las probabilidades p_l y p_i obtenidas del algoritmo PageRank constituyen características suficientes para la agrupación en clusters, alcanzando un Silhouette score de 0.955 sin aplicar PCA.

La aplicación de PCA deteriora consistentemente el rendimiento en las métricas evaluadas, sugiriendo que p_i y p_l, es decir, las dos características empleadas por PageRank capturan la información esencial para la clasificación. El algoritmo trimmed K–Means mejora la pureza global (0.715 vs 0.712) y la separación entre clústeres (índice Calinski-Harabasz), a costa de una ligera reducción en la cohesión interna. Si se quisiera aplicar PCA, sería recomendable modificar el algoritmo de K–Means a un algoritmo de weighted K–Means, que permite corregir las desviaciones para enfocarnos en lo esencial dependiendo de los intereses del análisis.

La metodología propuesta reduce significativamente el número de direcciones no clasificadas, proporcionando una herramienta práctica y simple, que puede funcionar como base para el análisis forense de redes blockchain. Sería interesante, junto a esta metodología, explorar otras técnicas de clustering más avanzadas, así como la incorporación de características temporales de las transacciones para mejorar la precisión de la clasificación.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

AGRADECIMIENTOS

Este trabajo forma parte del proyecto Laboratorio de I+D+i en Ciberseguridad, Privacidad y Comunicaciones Seguras (TRUST Lab), financiado por la Unión Europea NextGeneration-EU, Plan de Recuperación, Transformación y Resiliencia, a través del INCIBE.

REFERENCIAS

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," white paper, Oct. 31, 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf (Accessed May 26, 2025).

[2] European Securities and Markets Authority, Crypto Assets: Market Structures and EU Relevance, Risk Analysis Article ESMA50-524821-3153, Apr. 2024. [Online]. Available: https://www.esma.europa.eu/sites/default/files/2024-04/ESMA50-524821-3153_risk_article_crypto_assets_market_structures_and_eu_relevance.pdf (Accessed: May 26, 2025).

[3] J. Forestell, "Visa's role in stablecoins," Visa Perspectives, 30-Apr-2025. [Online]. Available: https://corporate.visa.com/en/sites/visa-perspectives/innovation/visas-role-in-stablecoins.html (Accessed: May 26, 2025).

[4] Financial Action Task Force (FATF), Targeted Update on Implementation of the FATF Standards on Virtual Assets and Virtual Asset Service Providers, Paris, France, Jun. 2023. [Online]. Available: https://www.fatf-gafi.org/en/publications/Fatfrecommendations/targeted-update-virtual-assets-vasps-2023.html (Accessed: May 26, 2025).

[5] M. Weber, G. Domeniconi, J. Chen, D. K. I. Weidele, C. Bellei, T. Robinson, and C. E. Leiserson, "Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics," in Proc. KDD '19 Workshop on Anomaly Detection in Finance, Anchorage, AK, USA, Aug. 2019, pp. 1-7. Doi: 10.48550/arXiv.1908.02591

[6] Y. Elmougy and L. Liu, "Demystifying Fraudulent Transactions and Illicit Nodes in the Bitcoin Network for Financial Forensics," in Proc. 29th ACM SIGKDD Conf. Knowledge Discovery and Data Mining (KDD '23), New York, NY, USA: ACM, 2023, pp. 3979-3990. doi: 10.1145/3580305.3599803

[7] A. Asiri and K. Somasundaram, "Graph convolution network for fraud detection in bitcoin transactions," Scientific Reports, vol. 15, Art. no. 11076, 2025, doi: 10.1038/s41598-025-95672-w

[8] Nicholls J., Kuppa A., and Le-Khac N.-A., "FraudLens: Graph Structural Learning for Bitcoin Illicit Activity Identification," in Proc. 39th Annu. Computer Security Applications Conf. (ACSAC '23), Austin, TX, USA, 4–8 Dec. 2023, pp. 324-336. doi: 10.1145/3627106.3627200

[9] G. Vlahavas, K. Karasavvas, and A. Vakali, "Unsupervised clustering of bitcoin transactions," Financial Innovation, vol. 10, Art. no. 25, Jan. 2024, doi: 10.1186/s40854-023-00525-y

[10] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the Web," Stanford InfoLab, Tech. Rep. 1999-66, Nov. 1999. [Online]. Available: http://ilpubs.stanford.edu:8090/422 (Accessed: May 26, 2025).

[11] H. Tong, C. Faloutsos, and J.-Y. Pan, "Random walk with restart: fast solutions and applications," Knowledge and Information Systems, vol. 14, no. 3, pp. 327–346, 2008, doi: 10.1007/s10115-007-0094-2

[12] Y. Elmougy and L. Liu, "Elliptic++ Dataset: A graph network of Bitcoin blockchain transactions and wallet addresses," GitHub repository, git-disl/EllipticPlusPlus, Apr. 2024. [Online]. Available: https://github.com/git-disl/EllipticPlusPlus (Accessed: May 26, 2025).

[13] C. Ding and X. He, "K-means Clustering via Principal Component Analysis," in Proc. 21st Int. Conf. Machine Learning (ICML 2004), Banff, Alberta, Canada, Jul. 2004, pp. 225–232. [Online]. Available: https://icml.cc/Conferences/2004/proceedings/papers/262.pdf

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Vehículo Terrestre Autónomo para la Automatización de Auditorias de Seguridad Wi-Fi

Diego García Ayala, Iván Quiñonero Martínez de Ojeda, Raúl Castro Montoya, María-Dolores Cano
Departamento de Tecnologías de la Información y las Comunicaciones Universidad Politécnica de Cartagena.
diego.gayala@upct.es, ivan.quinonero@upct.es, raula.castro@edu.upct Mdolores.Cano@upct.es

La creciente proliferación de redes inalámbricas Wi-Fi plantea desafíos en términos de seguridad, sobre todo en entornos industriales de gran extensión donde localizar físicamente un punto de acceso comprometido resulta complejo. Este artículo describe el diseño y la validación de un Vehículo Terrestre Autónomo (VTA) de bajo coste que automatiza las tareas de auditoría de redes Wi-Fi. El sistema integra un microcontrolador ESP32-S3 para el escaneo y la inyección controlada de tramas de desautenticación, un Arduino Pro Micro para la locomoción, y un algoritmo de navegación reactiva basado en el indicador de intensidad de señal recibida (*Received Signal Strength Indicator*, RSSI). Además, incorpora una interfaz gráfica táctil y un registro persistente de eventos en tarjeta SD, posibilitando la documentación forense completa. Las pruebas preliminares en un entorno controlado muestran tiempos de localización de 45-60 s y una tasa de captura de *handshakes* WPA/WPA2 del 93,7 %, lo que demuestra la viabilidad de la solución para auditorías de seguridad repetibles y escalables.

***Palabras Clave*—Vehículo Terrestre Autónomo, Wi-Fi, Navegación autónoma, Auditoría de red.**

I. INTRODUCCIÓN

En los últimos años, la proliferación masiva de redes inalámbricas ha aumentado de manera exponencial, convirtiendo el conjunto de protocolos agrupados bajo el nombre Wi-Fi (IEEE 802.11) [1] en uno de los estándares de conectividad más extendidos. La adopción de IEEE 802.11 se ha extendido más allá de los entornos domésticos y de oficina para convertirse en un pilar de la automatización industrial. Como ejemplo, en 2022 se enviaron 3,8 mil millones de productos Wi-Fi y se prevé que haya 628 millones de *hot-spots* públicos en 2023 [2], cifras que también reflejan su rápida penetración en fábricas y plantas de proceso [3].

En paralelo, la rápida adopción del paradigma del Internet de las Cosas (*Internet of Things*, *IoT*) y la digitalización de sectores críticos como la salud, el transporte, la energía o la Industria 4.0, han incrementado la cantidad y diversidad de dispositivos conectados. Los estudios más recientes describen un crecimiento sostenido de los vectores de amenaza que van desde suplantación de identidad hasta denegación de servicio y fuga de datos [4] [5]. Lo que sin duda subraya la necesidad de detección temprana y respuesta automatizada.

La protección de redes Wi-Fi, especialmente en entornos industriales, corporativos o educativos, representa hoy en día un reto estratégico. Las auditorías manuales resultan lentas, propensas a omisiones humanas, y dependen fuertemente de la experiencia del operador. Además, identificar físicamente el origen de una señal Wi-Fi sospechosa o localizar puntos de acceso mal configurados puede convertirse en una tarea ardua, especialmente en instalaciones de gran escala, con múltiples zonas de cobertura, obstáculos físicos o presencia de señales interferentes. Para afrontar dichos riesgos, la comunidad investigadora recurre a auditorías especializadas de seguridad Wi-Fi en entornos corporativos e industriales. Trabajos experimentales [6] han demostrado la viabilidad de capturar *handshakes* WPA/WPA2 y de explotar vulnerabilidades 802.11 en laboratorios controlados, estableciendo metodologías de referencia para pruebas de penetración industriales.

En este contexto, se hace cada vez más necesaria la incorporación de herramientas automatizadas que permitan evaluar de forma dinámica la seguridad inalámbrica. Estas herramientas deben ser capaces de detectar, rastrear, auditar e interactuar con redes Wi-Fi en tiempo real, ejecutando acciones controladas de penetración, generando métricas objetivas y documentando cada evento de forma sistemática. *Frameworks* recientes como por ejemplo, DRAT [7], que automatiza ataques y registro forense con drones, evidencian el interés creciente por herramientas que minimicen la intervención humana y aumenten la consistencia de las auditorías.

Este trabajo presenta el desarrollo integral de un Vehículo Terrestre Autónomo (VTA) orientado a tareas

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de auditoría y penetración en redes inalámbricas, combinando tecnologías de escaneo Wi-Fi, navegación basada en intensidad de señal (*Received Signal Strength Indicator*, RSSI), interfaz gráfica táctil, ataques activos y registro persistente de eventos. Se trata de una solución versátil, portable y replicable, pensada tanto para entornos de laboratorio como para despliegues reales controlados, donde se requiera una auditoría exhaustiva y repetible. La motivación principal radica en la necesidad de automatizar procesos como la localización de puntos de acceso, el análisis de cobertura efectiva, la captura de *handshakes* WPA/WPA2 y la evaluación del nivel de exposición de clientes conectados mediante ataques de desautenticación. Adicionalmente, el sistema está diseñado para permitir su integración futura con módulos de respuesta activa o visualización en tiempo real desde una consola remota. La arquitectura hardware se basa en un microcontrolador ESP32-S3 para las tareas inalámbricas, complementado por un Arduino Pro Micro encargado del control físico de movimiento. La navegación se realiza de forma autónoma mediante evaluación local de RSSI y sensores de proximidad, mientras que la selección de objetivos y supervisión de estado se lleva a cabo a través de una pantalla táctil de tipo *Thin Film Transistor* (TFT). Todos los eventos se registran en una tarjeta SD en formato estructurado, permitiendo su análisis forense, visualización o exportación a plataformas externas. La propuesta desarrollada busca contribuir al ecosistema de herramientas ofensivas éticas para ciberseguridad, combinando movilidad, inteligencia embebida y capacidad de intervención inalámbrica en un único dispositivo compacto. Esta herramienta tiene aplicaciones potenciales en formación práctica, ejercicios de *red teaming*, auditorías internas en empresas, y como base experimental para investigación en navegación robótica y análisis de señales inalámbricas.

El resto del artículo se organiza de la siguiente forma. La Sección II resume los fundamentos del estándar IEEE 802.11 necesarios para contextualizar el trabajo. La Sección III describe en detalle la estructura y características de las tramas de desautenticación y en la Sección IV se analizan los principales ataques inalámbricos que emplean dichas tramas. La Sección V detalla la inyección controlada de tramas desde el ESP32-S3. La Sección VI presenta la arquitectura funcional y el modo de operación del vehículo autónomo propuesto. Los resultados experimentales, junto con las consideraciones éticas y legales, se exponen en la Sección VII. El trabajo concluye resumiendo los resultados más destacados y el trabajo futuro.

II. ESTÁNDAR IEEE 802.11

El estándar IEEE 802.11 constituye la base de las redes locales inalámbricas (WLAN), más comúnmente conocidas como redes Wi-Fi. Desarrollado por el *Institute of Electrical and Electronics Engineers* (IEEE), este conjunto de especificaciones define los protocolos para la transmisión de datos en redes inalámbricas, incluyendo aspectos relacionados con la capa física y la capa de control de acceso al medio (MAC) del modelo OSI [1]. Desde su publicación en 1997, el estándar ha ido evolucionando de manera significativa para atender las crecientes demandas de velocidad, eficiencia y seguridad en las comunicaciones inalámbricas. A lo largo de estas últimas décadas, se han ido incorporando enmiendas que han dado lugar a nuevas versiones del protocolo (802.11a/b/g/n/ac/ax), cada uno aportando mejoras tanto en rendimiento, como en espectro o modulación y tolerancia frente a las interferencias. Este estándar presenta funcionalidades claves como:

- Asociación y autenticación de dispositivos: el protocolo define los procedimientos mediante los cuales una estación inalámbrica (STA) se conecta a un punto de acceso (*Access Point*, AP), incluyendo el intercambio de tramas de gestión como *beacons*, *probes*, *association requests*, *authentication*, *deauthentication*, entre otras.
- Control de acceso al medio (MAC): utiliza esquemas como *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) para gestionar el acceso al canal inalámbrico compartido, minimizando colisiones y mejorando la eficiencia del canal.
- Seguridad: inicialmente estuvo limitado a esquemas de seguridad débiles como WEP. El estándar evolucionó para incorporar WPA, WPA2 y, más recientemente, WPA3 y el uso obligatorio de PMF (*Protected Management Frames*) a partir de la enmienda 802.11w, con el fin de proteger la integridad de las tramas de gestión.
- Soporte de calidad de servicio (*Quality of Service*, QoS): mediante la enmienda 802.11e, se introdujeron mecanismos para priorizar diferentes tipos de trafico (voz, vídeo, datos) utilizando categorías de acceso diferenciadas.
- Altas tasas de transferencia: estándares como 802.11ac y 802.11ax (Wi-Fi 6) han permitido superar velocidades del orden de los gigabits por segundo, gracias al uso de tecnologías como MIMO (*Multiple Input, Multiple Output*), OFDMA (*Orthogonal Frecuency Division Multiple Access*) y canales de mayor ancho de banda (hasta 160MHz).

III. TRAMAS DE DESAUTENTICACIÓN

Las tramas de desautenticación tienen un formato específico definido por la norma IEEE 802.11. Se transmiten en texto claro, a menos que se utilicen extensiones como 802.11w para protegerlas. En la Tabla I se presenta la estructura típica de una trama de desautenticación.

El campo "*Reason Code*" permite especificar la causa de la desautenticación. Algunos valores comunes son:

- 0x0001: Código no especificado.
- 0x0002: STA no autenticada.
- 0x0003: STA está dejando la red.
- 0x0004: STA ha sido desautenticada por inactividad.

IV. ATAQUES INALÁMBRICOS

Un ataque de desautenticación constituye una técnica común dentro del ámbito de los ataques inalámbricos contra redes Wi-Fi. En este tipo de ataque, un atacante envía

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
ESTRUCTURA DE UNA TRAMA DE DESAUTENTICACIÓN IEEE 802.11

Campo	Tamaño (Bytes)	Descripción
Frame Control	2	Indica que se trata de una trama de gestión
Duration/ID	2	Tiempo de uso del canal o ID de asociación.
Dirección 1 (DA)	6	Dirección MAC de destino.
Dirección 2 (SA)	6	Dirección MAC del emisor.
Dirección 3 (BSSID)	6	Identificador del AP.
Sequence Control	2	Número de secuencia de la trama.
Reason Code	2	Código que indica el motivo de la desautenticación.

de manera repetida tramas de desautenticación (*deauthentication frames*) a uno o varios clientes conectados a un punto de acceso, provocando la interrupción forzada de la conexión entre ambos. Esta acción puede derivar en una denegación de servicio (*Denial of Service*, DoS) temporal, interrumpiendo el acceso legítimo a una red [8] [9] [10].

Para explotar este tipo de ataques, se aprovecha de una debilidad estructural del propio protocolo. Las tramas de gestión, tales como serían las tramas de desautenticación, tradicionalmente nunca han sido cifradas ni autenticadas. Esto permite que cualquier dispositivo cercano pueda generar y enviar dichas tramas falsificadas, siempre que se conozca la dirección MAC del punto de acceso objetivo. Una de las técnicas más habituales consiste en enviar tramas con la dirección especial de difusión FF:FF:FF:FF:FF:FF como destino. Esto provoca que todas las estaciones asociadas al punto de acceso reciban la instrucción de desautenticarse, forzando la desconexión simultanea de múltiples dispositivos.

Además de producir una denegación de servicio, este ataque suele utilizarse como paso preliminar para otras acciones más complejas. Al forzar la desconexión de un cliente legítimo, es posible capturar el *4-way handshake* cuando la estación intenta reconectarse al punto de acceso protegido por WPA o WPA2. Este *handshake* puede ser utilizado posteriormente en un ataque de diccionario o fuerza bruta para intentar obtener la clave precompartida (PSK). También se puede desviar al usuario hacia un punto de acceso malicioso, en lo que se conoce como un ataque de Gemelo Maligno (*Evil Twin*), donde el atacante simula el SSID original para interceptar tráfico o robar credenciales de usuarios legítimos. Estos ataques han sido ampliamente documentados en literatura técnica y académica [6][11][12], y son posibles debido a la ausencia de integridad criptográfica en las tramas de control y gestión en la mayoría de implementaciones del protocolo.

Para abordar esta vulnerabilidad, la enmienda IEEE 802.11w-2009 de la que hemos hablado antes, introduce *Protected Managemente Frames* (PMF), que permiten cifrar y autenticar tramas de gestión críticas, incluyendo las de desautenticación y desasociación. No obstante, su adopción ha sido muy lenta en la mayoría de los dispositivos, especialmente en el ámbito doméstico y en hardware antiguo, por lo que el ataque de desautenticación sigue siendo viable en la gran mayoría de entornos.

V. INYECCIÓN DE TRAMAS DE DESAUTENTICACIÓN DESDE ESP32

Para este trabajo se ha seleccionado la plataforma embebida ESP32-WROOM-32E por su versatilidad, costo reducido, facilidad de programación, y su robusto módulo inalámbrico integrado que soporta el protocolo IEEE 802.11 b/g/n. El ESP32-WROOM-32E es especialmente adecuado para aplicaciones embebidas que requieren comunicación inalámbrica intensiva y procesamiento eficiente. Además, se ha incluido un Arduino Pro Micro por su eficiencia en la gestión de periféricos de control como motores y sensores, liberando recursos computacionales en el ESP32 para tareas inalámbricas y de interacción con el usuario mediante una pantalla táctil TFT-S3. Esta pantalla, por su parte, facilita la interacción intuitiva con el usuario, permitiendo una rápida visualización y selección de redes Wi-Fi objetivo.

La implementación oficial del *firmware* Wi-Fi de Espressif, que se distribuye junto al Software Development Kit (SDK) estándar, incluye una validación estricta del tipo y subtipo de las tramas Wi-Fi al utilizar las funciones proporcionadas oficialmente, tales como *esp_wifi_80211_tx()*. Específicamente, las tramas de gestión tipo 0xC (subtipo "*deauthentication*") se encuentran bloqueadas por defecto, lo que impide realizar estas auditorías de manera eficaz desde el *firmware* oficial del fabricante. Este bloqueo interno se manifiesta típicamente con un mensaje de error del estilo: *E (XXXX) wifi: unsupport frame type: 0c0*.

Esta restricción tiene como objetivo cumplir con requisitos regulatorios (FCC, ETSI) y éticos, asegurando que el dispositivo no sea utilizado inapropiadamente para interrumpir el servicio de terceros. Para superar esta limitación técnica de forma ética, es habitual en contextos académicos y de auditorías legítimas recurrir a métodos que implican modificar parámetros específicos que explicaremos a continuación. La mayoría de bloqueos impuestos afectan principalmente a la interfaz STA. No obstante, es posible enviar tramas de desautenticación desde la interfaz AP del ESP32 debido a diferencias en el manejo interno de paquetes dentro del *firmware* oficial. Esto implica configurar previamente el ESP32 en modo AP (o AP+STA) y habilitar explícitamente protocolos 802.11b/g/n mediante la configuración técnica en el *firmware*, tal y como se puede observar dentro de la Figura 1.

Para construir una trama de desautenticación (*802.11 management subtype 0xC*) hay que seguir el estándar IEEE 802.11-2016. Una trama mínima típica tiene la siguiente estructura binaria simplificada (usando lenguaje C), tal y como se aprecia dentro de la Figura 2.

Finalmente, para lograr una transmisión efectiva de esta trama se realiza mediante la función *esp_wifi_80211_tx()* desde la interfaz AP, asegurando antes la selección correcta del canal mediante *esp_wifi_set_channel()*. Este fragmento de código puede visualizarse en la Figura 3. Así, esta secuencia técnica permite inyectar tramas de desautenticación desde un ESP32 de forma controlada, ética y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

```
esp_wifi_set_mode(WIFI_MODE_APSTA);
esp_wifi_start();
esp_wifi_set_protocol(WIFI_IF_AP,
    WIFI_PROTOCOL_11B | WIFI_PROTOCOL_11G |
     WIFI_PROTOCOL_11N);
```

Fig. 1. Configuración de ESP32 en modo AP+STA.

efectiva, eludiendo la restricción original del *firmware* mediante el uso documentado de interfaces alternativas (AP) y ajustes específicos del protocolo 802.11.

```
struct __attribute__((packed))
    wifi_deauth_frame{
    uint8_t frame_control[2];
    uint8_t duration[2];
    uint8_t destination[6];
    uint8_t source[6];
    uint8_t bssid[6];
    uint8_t sequence_control[2];
    uint8_t reason_code[2];
};
```

Fig. 2. Construcción de una trama de desautenticación.

```
esp_wifi_set_channel(channel,
    WIFI_SECOND_CHAN_NONE);

wifi_deauth_frame pkt;
memset(&pkt, 0, sizeof(pkt));
pkt.frame_control[0] = 0xC0;
pkt.frame_control[1] = 0x00;
memcpy(pkt.destination, victim_mac, 6);
memcpy(pkt.source, ap_mac, 6);
memcpy(pkt.bssid, ap_mac, 6);
pkt.reason_code[0] = 0x01;

//Enviamos paquete mediante la interfaz AP
    del ESP32
esp_wifi_80211_tx(WIFI_IF_AP, (uint8_t*)&
    pkt, sizeof(pkt), false);
```

Fig. 3. Transmisión de la trama en ESP32.

Esta secuencia técnica permite inyectar tramas de desautenticación desde un ESP32 de forma controlada, ética y efectiva, eludiendo la restricción original del *firmware* mediante el uso documentado de interfaces alternativas (AP) y ajustes específicos del protocolo 802.11.

VI. ARQUITECTURA FUNCIONAL Y OPERACIÓN DEL VEHÍCULO AUTÓNOMO PARA AUDITORÍA WI-FI

Proponemos un VTA capaz de detectar redes Wi-Fi en su entorno, desplazarse hacia aquellas con mayor intensidad de señal , ejecutar ataques activos de desautenticación, capturar *handshakes* WPA/WPA2 y registrar de forma persistente todos los eventos relevantes. Esta propuesta se enmarca en un enfoque modular que facilita su desarrollo escalonado, mantenimiento y validación experimental.

Diversos trabajos recientes han explorado enfoques similares de auditoría inalámbrica automatizada mediante vehículos autónomos. En [13], los autores diseñan un kit portátil de bajo coste para auditorías Wi-Fi en interiores capaz de detectar AP falsos o mal configurados y generar informes automáticos para entornos corporativos y docentes. En [14] se emplearon vehículos terrestres guiados por mapas de calor RSSI generados previamente mediante técnicas de aprendizaje automático. La herramienta WiFi Auditor propuesta en [15] integra en una Raspberry Pi una *suite* que automatiza escaneo, *deauth* y *man-in-the-middle* portátil, orientado a pruebas de penetración discretas y de bajo presupuesto. De forma similar a este trabajo, el trabajo desarrollado en [9] implementa sobre un ESP32 varios ataques Wi-Fi mostrando que un microcontrolador económico puede ejecutar pruebas avanzadas con mínimo hardware, aunque permanece estático y requiere intervención manual para acercarse al objetivo. A diferencia de este, nuestra propuesta mantiene una implementación sencilla y de bajo costo e incorpora navegación reactiva en tiempo real basada únicamente en medidas RSSI instantáneas, evitación de obstáculos y registro forense en tarjeta SD, todo gestionado desde una interfaz táctil integrada. De este modo, localiza físicamente los puntos de acceso y ejecuta auditorías Wi-Fi sin infraestructura previa ni operador *in situ*, reduciendo al mínimo la complejidad y los costes operativos.

A. Justificación Técnica del Diseño Modular

El sistema se ha estructurado en cinco módulos principales, cada uno responsable de una tarea específica dentro del ciclo de operación. Este enfoque modular permite aislar el desarrollo y la validación de cada componente, facilitando su integración final en el *firmware* global:

- *Wi-Fi Scanner Module* (*wifi scanner*): Detecta y clasifica redes inalámbricas cercanas, obteniendo SSID, BSSID, RSSI y canal de cada una.
- *Interface Module* (*ui module*): Gestiona la pantalla táctil TFT-S3, mostrando la lista de redes ordenadas por RSSI y permitiendo al usuario seleccionar el objetivo.
- *Movement Control Module* (*movement comm*): Coordina el desplazamiento mediante comunicación UART con un Arduino Pro Micro, que acciona motores y sensores de proximidad.
- *Wi-Fi Attack Module* (*wifi attacks*): Pone el ESP32-S3 en modo promiscuo, construye e inyecta tramas de desautenticación y escucha paquetes EAPOL para capturar *handshakes*.
- *SD Logging Module* (*sd logger*): Registra en tarjeta SD todos los eventos (escaneos, movimientos, obstáculos, ataques, *handshakes*) con marcas temporales precisas en formato CSV.

B. Esquema Electrónico y Conectividad del Sistema

Antes de describir el funcionamiento dinámico, se presenta el diagrama esquemático que muestra las interconexiones físicas y lógicas entre los distintos componentes del vehículo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En la Figura 4 se representa:

- El ESP32-S3 como unidad central de procesamiento, conectado por SPI a la pantalla táctil TFT-S3 y a la tarjeta SD.
- El canal UART (UART1: TX17, RX16) para comunicación con el Arduino Pro Micro, encargado del control de motores y sensores de proximidad.
- Sensores ultrasónicos/infrarrojos conectados al Arduino Pro Micro, con retroalimentación de obstáculos enviada al ESP32-S3 mediante UART.
- Fuente de alimentación común (batería LiPo) con regulador de 5 V y conversor a 3.3 V para el ESP32-S3, evitando ruidos e interferencias entre módulos.

A su vez, en la Figura 5 podemos observar el diseño de placa propuesto para facilitar el montaje y el uso del dispositivo.

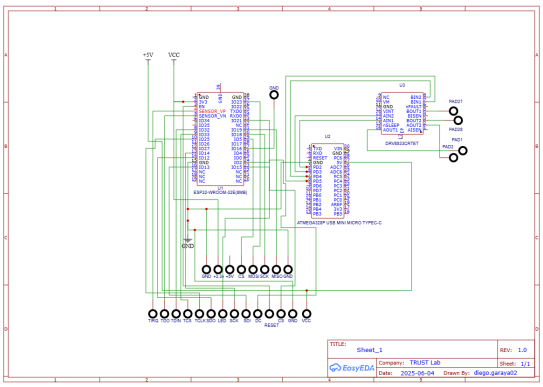


Fig. 4. Diagrama esquemático de conexiones del vehículo autónomo

C. Proceso de Inicialización del Sistema

Al energizarse el vehículo, se ejecuta un procedimiento secuencial que garantiza la disponibilidad de todos los periféricos antes de iniciar el ciclo operativo como se muestra en Alg. 1.

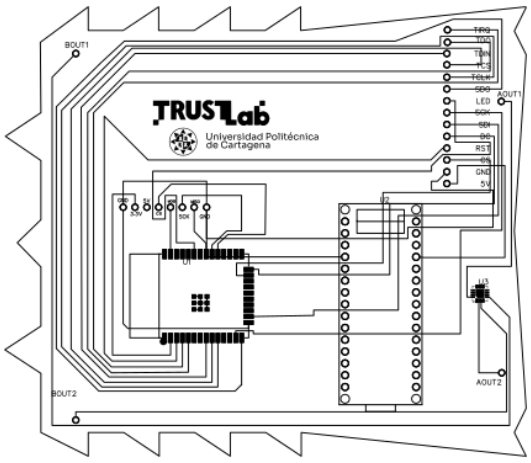


Fig. 5. Diagrama del diseño de placa propuesto

Algorithm 1 Inicialización del sistema

```
1: Configure TFT display
2: Initialize touch controller
3: Establish UART connection with Arduino Pro Micro
4: Mount SD card filesystem
5: Set Wi-Fi interface to station mode
6: Display welcome screen
```

Esta rutina asegura que el dispositivo está plenamente operativo antes de comenzar el escaneo de redes.

D. Detección y Selección de Redes Wi-Fi

El sistema realiza escaneos periódicos del entorno (intervalo configurable, por defecto 5 s). Cada vez que se completa un escaneo, las redes detectadas se ordenan por RSSI y se muestran en pantalla para que el usuario seleccione manualmente el objetivo.

El umbral de -60 dBm se ha definido empíricamente para descartar redes con baja probabilidad de éxito en la fase de navegación y ataque (véase Alg. 2).

Algorithm 2 Escaneo continuo y selección de objetivo

```
1: while system is active do
2:     networks ← Perform_WiFi_Scan()
3:     Sort(networks) by RSSI (descending)
4:     Display(networks) {Mostrar lista en TFT}
5:     selectedNetwork ← Wait_For_User_Selection()
6:     if selectedNetwork.RSSI > −60 dBm then
7:         AutonomousNavigation(selectedNetwork)
8:     else
9:         Display("Señal insuficiente, seleccione otra red")
10:    end if
11:    Wait(WIFI_SCAN_INTERVAL)
12: end while
```

E. Navegación Autónoma Basada en RSSI

Una vez seleccionada la red objetivo (`selectedNetwork`), el vehículo inicia un algoritmo de exploración direccional para maximizar la RSSI y acercarse progresivamente al punto de acceso . Se combina la métrica RSSI con la detección de obstáculos mediante sensores ultrasónicos/infrarrojos (véase Alg. 3). Descripción de pasos:

- Línea 1: Inicializa `bestRSSI` con el valor obtenido tras la selección inicial.
- Línea 2–8: Para cada dirección cardinal, se envía un comando de movimiento al Arduino Pro Micro con MOVE_DURATION (por defecto 500 ms). Inmediatamente se mide la nueva RSSI en modo promiscuo.
- Línea 4–6: Si la RSSI mejora, se actualiza `bestRSSI` y se registra el movimiento; en caso contrario, se retrocede para no alejarse de la fuente.
- Línea 9–11: Si el sensor de proximidad detecta un obstáculo a menos de 20 cm, se invoca el subprocedimiento de evitación (giro lateral y nueva medición de RSSI).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Algorithm 3 Navegación autónoma basada en RSSI y sensores de proximidad

```
1: bestRSSI ← selectedNetwork.RSSI
2: while bestRSSI < −20 dBm do
3:   for all direction ∈ {Forward, Left, Right, Backward} do
4:     Send_Move_Command(direction, MOVE_DURATION)
5:     newRSSI ← Measure_RSSI(selectedNetwork.BSSID)
6:     if newRSSI > bestRSSI then
7:       bestRSSI ← newRSSI
8:       LogEvent("MOVIMIENTO", direction)
9:     else
10:      Send_Move_Command(Opposite(direction), MOVE_DURATION)
11:    end if
12:  end for
13:  if Detect_Obstacle() = TRUE then
14:    Execute_Obstacle_Avoidance()
15:  end if
16: end while
17: PerformWiFiAttack(selectedNetwork)
```

- Línea 12: Al superar el umbral de -20 dBm, se considera que el vehículo ha llegado al origen de la señal.

F. Ejecución del Ataque de Desautenticación y Captura de handshakes

Cuando la RSSI del punto de acceso seleccionado alcanza o supera -20 dBm, se detiene todo movimiento y se configura el ESP32-S3 en modo promiscuo para lanzar un ataque de desautenticación. Este ataque, descrito en Alg.4, fuerza a los clientes conectados a reconectarse, lo que permite capturar el handshake WPA/WPA2.

Algorithm 4 Ataque activo de desautenticación y captura de handshake

```
1: Set_WiFi_Channel(selectedNetwork.Channel)
2: Enable_Promiscuous_Mode()
3: handshakeCaptured ← FALSE
4: for i = 1 to MAX_DEAUTH_PACKETS do
5:   Send_Deauth_Frame(selectedNetwork.BSSID)
6:   if Detect_EAPOL_Packet() = TRUE then
7:     Store_Handshake(SD, packetData)
8:     handshakeCaptured ← TRUE
9:   end if
10: end for
11: Display_Attack_Result(handshakeCaptured)
```

Descripción de pasos:

- Línea 1: Se fija el canal Wi-Fi al canal donde opera el punto de acceso objetivo.
- Línea 2: Se activa el modo promiscuo del chip para recibir todas las tramas en el aire.
- Línea 3: Variable booleana que indica si se ha capturado al menos un handshake.
- Línea 4–8: Se envían hasta `MAX_DEAUTH_PACKETS` tramas de desautenticación con direccionamiento *broadcast* hacia el AP (dirección 1 = FF:FF:FF:FF:FF:FF, direcciones 2 y 3 = BSSID del AP). Después de cada envío, se comprueba si llegó un paquete EAPOL que forme parte del *handshake*. En caso afirmativo, se almacena en la tarjeta SD.
- Línea 9: Se muestra en pantalla el estado final del ataque (captura exitosa o no).

G. Registro de Actividades y Eventos

Para garantizar trazabilidad y permitir un análisis forense posterior, todos los eventos del sistema se registran de manera persistente en la tarjeta SD en formato CSV. Cada línea incluye un `timestamp` (formato Unix epoch), el tipo de evento y datos asociados (véase Alg. 5).

Algorithm 5 Registro de eventos en tarjeta SD

```
1: timestamp ← Get_Current_Timestamp()
2: eventData ← Format_Event_Data(eventType, eventDetails)
3: Write_To_SD(timestamp, eventData)
```

Tipos de eventos registrados:

- Escaneo de red: `ESCANEO`; SSID; BSSID; RSSI; canal; tipo de autenticación.
- Movimiento: `MOVIMIENTO`; dirección; duración (ms).
- Detección de obstáculo: `OBSTÁCULO`; tipo de sensor; distancia (cm).
- Inicio/fin de ataque: `ATAQUE_INICIO`; SSID; BSSID; canal; número de paquetes enviados. `ATAQUE_FIN`; SSID; BSSID; canal; número de paquetes; *handshakes* capturados.
- Handshake capturado: `HANDSHAKE`; SSID; BSSID; MAC del cliente; datos EAPOL (en hexadecimal).

VII. Resultados Experimentales

Las pruebas experimentales se llevaron a cabo en un laboratorio cerrado, con dimensiones aproximadas de 50 metros cuadrados (10 m × 5 m), equipado con diversos obstáculos controlados, incluyendo mobiliario estándar de oficina, paneles de madera, estructuras metálicas y muros de ladrillo que simulan un entorno realista. La red objetivo consistía en un punto de acceso comercial estándar configurado con WPA2, ubicado estratégicamente en posiciones variables para evaluar la capacidad del vehículo de adaptación ante diferentes escenarios. Durante todas las pruebas, se controlaron variables ambientales como temperatura, humedad y posibles fuentes de interferencia inalámbrica externas para asegurar la repetibilidad de los resultados.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

A. Precisión de Localización por RSSI

Se definieron rutas de prueba con obstáculos estáticos (cajas y paneles móviles). El punto de acceso se situó a 6 m del punto de partida del vehículo. Los resultados promedio obtenidos fueron:

- Tiempo promedio de localización: 45–60 s.
- Número promedio de movimientos exploratorios: 8–10.
- Desviación angular media: 18° con respecto a la dirección real del AP.
- Número de correcciones por obstáculos: 2–3 por prueba.

Estos resultados demuestran que, a pesar de emplear un algoritmo de búsqueda local simple, la combinación de RSSI y detección de obstáculos permite aproximarse con precisión suficiente al origen de señal.

B. Eficacia del Ataque Wi-Fi

Se realizaron pruebas con 8 puntos de acceso configurados en WPA2 y con al menos un cliente conectado:

- Tasa de éxito en captura de *handshakes*: 93.7 %.
- Tiempo medio de captura de handshake: 11.4 s.
- Promedio de paquetes de desautenticación enviados: 10 000.

El rendimiento es comparable al de herramientas manuales (por ejemplo, `aircrack-ng`) en dispositivos portátiles, con la ventaja añadida de la movilidad autónoma y la automatización completa del proceso.

C. Integridad y Fiabilidad del Registro de Datos

Se analizaron los archivos generados en la tarjeta SD tras cada sesión de prueba. Todos los registros presentaron coherencia y formato válido:

- Eventos registrados por sesión: 40–70 entradas.
- Integridad de timestamps: 100 % sin pérdidas ni saltos.
- Compatibilidad con herramientas de análisis: Total (Excel, Python Pandas).

La trazabilidad completa facilita reconstruir cada fase del experimento, validando tanto el comportamiento de navegación como la generación de ataques y la captura de *handshakes*.

D. Consideraciones Éticas y Legales

Dado que el vehículo autónomo implementa técnicas ofensivas en redes Wi-Fi, su uso debe ceñirse a un marco ético y legal estricto. A continuación se resumen los principios y regulaciones aplicables:

1) Principios Éticos:

- Consentimiento informado: Todo escaneo, ataque y captura de datos debe realizarse únicamente sobre redes cuya auditoría esté expresamente autorizada por el propietario.
- Proporcionalidad: Las acciones deben ser limitadas y adecuadas al objetivo de la evaluación, evitando interrupciones prolongadas o daños a sistemas colaterales.
- No maleficencia: El sistema debe minimizar interferencias en redes legítimas ajenas al experimento.
- Responsabilidad y trazabilidad: Cada acción se registra para garantizar total transparencia y posibilidad de auditoría posterior.

2) Marco Legal: El uso de técnicas de interceptación y desautenticación en redes Wi-Fi se encuentra regulado por diversas normativas nacionales e internacionales:

- España: Ley Orgánica 10/1995 del Código Penal arts. 197 (interceptación) y 264 (sabotaje informático), Ley 9/2014 General de Telecomunicaciones (art. 76.4c: perturbación de redes).
- Unión Europea: Directiva 2013/40/UE sobre ataques contra los sistemas de información y Directiva 2002/58/CE (ePrivacy). Recordamos que el RGPD se aplica solo cuando se tratan datos personales capturados.
- Estados Unidos: Computer Fraud and Abuse Act (CFAA, 18 USC § 1030) y, cuando proceda, Wiretap ct (18 USC 2511).
- Reino Unido: Computer Misuse Act (CMA, 1990) y, para interceptación de contenido, Investigatory Powers Act 2016.

Se recomienda que todo experimento con el vehículo se lleve a cabo en entornos controlados, con permiso explícito de las partes involucradas y bajo supervisión de un equipo de ciberseguridad autorizado.

VIII. CONCLUSIONES

En este trabajo se ha presentado un VTA de bajo costo, capaz de combinar exploración inalámbrica, navegación reactiva guiada exclusivamente por RSSI instantáneo y ejecución controlada de ataques de desautenticación en Wi-Fi. Las pruebas en laboratorio (50 m2, obstáculos realistas) demostraron tiempos de localización de 45-60 s, un promedio de 8-10 maniobras y una desviación angular media de 18° respecto al punto de acceso. Además, el sistema capturó *handshakes* WPA/WPA2 en el 93,7% de los casos en aproximadamente 11 s, con trazabilidad completa gracias al registro estructurado en tarjeta SD. Estos resultados sitúan al prototipo al nivel de herramientas manuales de auditoría, con la ventaja añadida de la movilidad autónoma y la supervisión táctil integrada, reforzando su potencial para ejercicios de red teaming, docencia y pruebas en planta industrial. No obstante, la solución adolece de limitaciones inherentes a su sencillez. El algoritmo de búsqueda local podría estancarse en entornos multipath complejos, la inyección de tramas depende de la ausencia de PMF y las mediciones RSSI son sensibles a variaciones ambientales. Como líneas futuras se plantea la incorporación de aprendizaje reforzado para optimizar la exploración, extender a otros estándares Wi-Fi y otros protocolos como BLE para auditorías multibanda y evaluar la autonomía energética. De este modo, el VTA podría evolucionar de plataforma experimental a herramienta profesional de ciber resiliencia.

AGRADECIMIENTOS

Este trabajo forma parte del proyecto Laboratorio de I+D+i en Ciberseguridad, Privacidad y Comunicaciones

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Seguras (TRUSTLab), financiado por la Unión Europea NextGeneration-EU, Plan de Recuperación, Transformación y Resiliencia, a través del INCIBE.

Referencias

[1] IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std 802.11-2020, Feb. 2021.

[2] Cisco Systems, "Cisco Annual Internet Report (2018–2023)," White Paper, San Jose, CA, USA, Mar. 2020. [Online]. Available: [Accessed: Jun. 13, 2025].

[3] E. Oughton, G. Geraci, M. Polese, V. Shah, D. Bubley, and S. Blue, "Reviewing wireless broadband technologies in the peak smartphone era: 6G versus Wi-Fi 7 and 8," Telecommunications Policy, vol. 48, no. 6, pp. 101–118, Apr. 2024, doi:10.1016/j.telpol.2024.102766.

[4] T. Sasi, A. H. Lashkari, R. Lu, P. Xiong, and S. Iqbal, "A comprehensive survey on IoT attacks: taxonomy, detection mechanisms and challenges," *Journal of Information and Intelligence*, vol. 2, no. 6, pp. 455–513, Dec. 2023, doi:10.1016/j.jiixd.2023.12.001.

[5] F. Cerasuolo, G. Bovenzi, D. Ciuonzo, and A. Pescapè, "Attack-Adaptive Network Intrusion Detection Systems for IoT Networks through Class Incremental Learning," *Computer Networks*, vol. 263, pp. 115–129, Mar. 2025, doi:10.1016/j.comnet.2025.111228.

[6] M. Vanhoef and F. Piessens, "Key Reinstallation Attacks: Forcing Nonce Reuse in WPA2," in *Proc. ACM SIGSAC Conf. on Computer and Communications Security (CCS '17)*, Oct. 2017, pp. 1313–1328, doi:10.1145/3133956.3134027.

[7] C. S. Veerappan, P. Loh, V. Balachandran, and M. S. B. M. Fadilah, "DRAT: A Penetration Testing Framework for Drones," in *Proc. 2021 IEEE 16th Conf. on Industrial Electronics and Applications (ICIEA)*, Aug. 2021, pp. 1–8, doi:10.1109/ICIEA51954.2021.9516f363.

[8] B. Branco, J. Silvestre Serra Silva, and M. Correia, "Cyber Attacks on Commercial Drones: A Review," *IEEE Access*, vol. 13, pp. 9566–9577, 2025, doi:10.1109/ACCESS.2025.3527698.

[9] H. Tiwari and A. Tomar, "Slipper Zero: Exploring Wi-Fi Security Vulnerabilities and Attack Implementations on ESP32 Microcontrollers," in *Proc. IEEE Global Conf. on Wireless and Optical Technologies (GCWOT)*, Sept. 2024, pp. 1–7, doi:10.1109/GCWOT63882.2024.10805625.

[10] L. Calderon and G. T. Salvador, "Detection and Analysis of Flipper Zero Deauthentication Signals Using HackRF One Software-Defined Radio," in *Proc. 2024 Int. Conf. on Issues in Information Systems*, Nov. 2024, doi:10.1109/3ICT64318.2024.10824254.

[11] J. Bellardo and S. Savage, "802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions," in Proc. 12th USENIX Security Symposium, Washington, DC, USA, Aug. 2003, pp. 15–28.

[12] R. Korolkov, S. Kutsak, and V. Voskoboinyk, "Analysis of Deauthentication Attack in IEEE 802.11 Networks and a Proposal for Its Detection," *Bulletin of V. N. Karazin Kharkiv National University – Series: Mathematical Modeling, Information Technology, Automated Control Systems*, vol. 50, pp. 58–70, Jun. 2021, doi:10.26565/2304-6201-2021-50-06.

[13] P. Segeč, R. Kaloč, and R. Dobis, "Portable system for automated audit of Wi-Fi networks in organizations," *in Proc. 19th IEEE Int. Conf. on Emerging Technologies for a Smarter World (ICETA)*, Nov. 2021, pp. 11–15, doi:10.1109/ICETA54173.2021.9726564.

[14] R. *Liu et al*., "Collaborative Radio SLAM for Multiple Robots based on WiFi Fingerprint Similarity," in *Proc. 2021 IEEE Int. Conf. on Robotics and Biomimetics (ROBIO)*, 2021, pp. 795–801, doi:10.1109/ROBIO54168.2021.9739222.

[15] P. Shinde, A. Karve, P. Mandaliya, and S. Patil, "Wireless Security Audit & Penetration Test using Raspberry Pi," in *Proc. 2018 Int. Conf. on Smart City and Emerging Technology (ICSCET)*, Mumbai, India, Jan. 2018, pp. 1–5, doi:10.1109/ICSCET.2018.8537246.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Plataforma escalable de honeypots IoT en Azure: captura, análisis y respuesta automática frente a botnets Mirai y Gafgyt

Cristian Eduardo Calzadilla Cubides, María Jesús Alba Baena, Andrea Villacís Vanegas, Maria-Dolores Cano
Departamento de Tecnologías de la Información y las Comunicaciones
Universidad Politécnica de Cartagena
TRUST Lab, Edificio ELDI, Campus Muralla del Mar, 30202 Cartagena.
cristian.calzadilla@upct.es, maria.alba@edu.upct.es, andrea.villacis@upct.es, mdolores.cano@upct.es

La proliferación de dispositivos de Internet de las Cosas (*Internet of Things*, IoT) ha ampliado la superficie de ataque en redes críticas, facilitando la acción de *botnets* como Mirai y Gafgyt. En este trabajo describimos el diseño, implementación y evaluación de una plataforma de *honeypots* IoT basada en T-Pot y desplegada en Azure, que integra Cowrie (SSH/Telnet), Honeytrap (HTTP/HTTPS) y extensiones para MQTT/CoAP mediante Dionaea. Empleamos un pipeline de ingestión y correlación en ELK combinado con Suricata para la recolección de eventos, cálculo de hashes SHA-256/MD5 y generación de alertas en tiempo real. Además, incorporamos *honeytokens* y automatismos en Azure Firewall y *sandbox* para el bloqueo dinámico de IPs maliciosas y el análisis forense de muestras desconocidas. Tras una semana de pruebas intensivas, ofrecemos métricas de volumen de ataque, distribución geográfica, patrones de fuerza bruta, actividad de *botnets* y eficacia de los *honeytokens*, demostrando la escalabilidad y capacidad de respuesta de la solución.

***Palabras Clave*- IoT, honeypots, Mirai, Gafgyt, ciberseguridad.**

I. Introducción

La rápida adopción de dispositivos Internet de las Cosas (*Internet of Things*, IoT) en entornos domésticos, industriales y de infraestructuras críticas [1] ha ampliado significativamente la superficie de ataque de las redes modernas. Muchas de estas plataformas incorporan *firmware* ligero y credenciales por defecto, lo que las convierte en objetivos atractivos para *botnets* como Mirai y Gafgyt, capaces de reclutar miles de dispositivos para ataques DDoS y propagación de *malware*. La magnitud del problema queda evidenciada tras los ataques analizados por Kolias *et al*. en [2], donde se describe cómo la *botnet* Mirai llegó a orquestar ráfagas superiores al terabit por segundo contra infraestructuras clave, convirtiéndose en un punto de inflexión para la seguridad IoT.

La necesidad de esta investigación surge de las limitaciones críticas observadas en las soluciones existentes de monitorización de amenazas IoT. Los centros de operaciones de ciberseguridad enfrentan dificultades para obtener inteligencia de amenazas específica para dispositivos IoT, mientras que la mayoría de las plataformas de *honeypots* actuales operan de forma aislada, carecen de automatización en la respuesta ante incidentes y presentan gaps significativos en la cobertura de protocolos IoT emergentes como MQTT y CoAP.

Una revisión sistemática reciente [3] confirma que los ataques de denegación de servicio (*Denial of Service*, DoS) siguen siendo uno de los vectores predominantes en IoT, concentrando más del 30 % de los incidentes analizados. Lo que refuerza la urgencia de contar con mecanismos especializados de detección y respuesta. Esta carencia resulta todavía más crítica a la luz de los trabajos de Lakshminarayana *et al*. [4], quienes demuestran que las nuevas funcionalidades de MQTT v5.0 pueden ser explotadas para lanzar ataques DDoS de próxima generación contra despliegues IoT.

Este trabajo presenta una plataforma avanzada de *honeypots* IoT desplegada en Azure, basada en T-Pot e integrada con ELK y Suricata para generación de alertas y respuesta automática en tiempo real. Sus principales aportaciones son: automatización completa del ciclo detección–respuesta sin intervención humana, soporte extendido para protocolos IoT y servicios clásicos, correlación en tiempo real de eventos heterogéneos y arquitectura *cloud-native* escalable y tolerante a fallos. La elección de Azure se fundamenta en su elasticidad para gestionar picos de tráfico malicioso, servicios nativos que proporcionan capacidades de respuesta automática y conectividad global que facilita la exposición controlada a atacantes distribuidos

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

geográficamente. Los resultados tras una semana de pruebas validan la eficacia de la aproximación propuesta.

El resto del artículo se estructura de la siguiente forma. En la Sección 2 presentamos brevemente el estado del arte. En la Sección 3 describimos la arquitectura de la plataforma. La Sección 4 detalla la metodología. En la Sección 5 discutimos los resultados y, finalmente, en la Sección 6 extraemos conclusiones e identificamos trabajos futuros.

II. Trabajos Relacionados

Los avances en defensa frente a amenazas IoT se apoyan en tres pilares: *honeypots* especializados, *frameworks* de análisis adaptativo e integración en infraestructuras *cloud* escalables. Estas líneas de trabajo buscan reducir falsos positivos y automatizar la respuesta, aunque enfrentan limitaciones que motivan el enfoque automatizado de este trabajo.

Honeypots de baja interacción como Conpot [5] y Dionaea [6] han demostrado una alta eficacia en entornos controlados. No obstante, su emulación limitada impide estudiar técnicas evasivas avanzadas. Los de alta interacción, como ThingPot [7] e IoTPOT [8], ofrecen mayor realismo al simular dispositivos IoT completos, pero requieren mayor complejidad operativa.

Si nos centramos en la investigación sobre *honeypot*s para IoT, el IoT Industrial (*Industrial* IoT, IIoT) y los sistemas ciberfísicos (*Cyber Phisical Systems*, CPS), ésta ha progresado desde amplias taxonomías hasta mecanismos defensivos altamente especializados y plataformas de inteligencia de amenazas a gran escala. La bibliografía puede agruparse en cuatro categorías complementarios: estudios y taxonomías, diseño arquitectónico de *honeypots* y *honeynets*, estrategias de defensa basadas en el engaño y la teoría de juegos y, por último, estudios de inteligencia sobre amenazas centrados en el *malware*.

Franco *et al*. ofrecen la síntesis más exhaustiva hasta la fecha en [9], analizando 79 estudios sobre *honeypots* de IoT/IIoT/CPS publicados entre 2002 y 2020 y derivando una taxonomía multiaxial que abarca el propósito, el nivel de interacción, la escalabilidad y los requisitos de recursos. Su estudio pone de relieve el cambio que se está produciendo en este campo, desde las trampas autónomas de baja interacción a los ecosistemas híbridos y adaptables que integran controladores de tipo *Software Defined Networking* (SDN), plataformas en la nube y análisis basados en Inteligencia Artificial (IA). Es importante destacar que también identifican problemas pendientes (principalmente, la emulación realista de dispositivos y el intercambio coordinado de datos), con los que se pueden comparar los principales artículos más recientes que se analizan a continuación.

Los primeros trabajos se centraron en la emulación de servicios individuales, pero Zhang *et al*. introdujeron en [10] un *honeynet* híbrido multipuerto que combina la emulación a nivel de *firmware* con sensores *dockerizados* de baja interacción, capturando ataques desconocidos basados en SOAP que eluden las trampas centradas en Telnet. Sus resultados muestran el valor de la diversidad de servicios. El 79 % de las muestras únicas se obtuvieron solo tras añadir instancias de *firmware* de alta interacción. Chakkaravarthy *et al*. amplían el alcance defensivo de la intrusión en la red al *ransomware* incrustando un señuelo *honeyfolde*r dentro de puntos finales y correlacionando eventos de host con registros de flujo SDN a través de un motor de procesamiento de eventos complejos [11]. El sistema reduce la latencia de detección a menos de 4 s para 20 familias de *ransomware*, lo que ilustra cómo el engaño de los terminales puede complementar a los honeypots de red.

La *et al*. modelan la interacción entre los atacantes y un defensor que puede dirigir el tráfico a los *honeypots* como un juego bayesiano de señalización [12]. Los equilibrios analíticos demuestran que el engaño sólo produce una utilidad mensurable por encima de un umbral de atacantes activos, una idea que más tarde se utilizó para ajustar las políticas de activación adaptativas. De forma complementaria, Du y Wang proponen un juego pseudo-honeypot basado en SDN para la mitigación de DDoS en redes IIoT [13]. Mezclando señuelos de alta y baja interacción y aprovechando el controlador SDN para el rápido redireccionamiento de flujos, la propuesta alcanza un equilibrio Bayesiano-Nash demostrable que minimiza la sobrecarga del controlador. Los experimentos del banco de pruebas revelan una reducción del 27 % en el consumo de energía en comparación con las defensas estáticas, lo que subraya las ventajas de eficiencia del engaño en toda la red.

De forma paralela, Wang *et al*. realizan un análisis del linaje evolutivo de 72 familias de *malware* IoT, combinando inteligencia obtenida de blogs con 38.000 muestras capturadas por un despliegue mundial de Dionea [5]. Su clasificador de conjuntos vincula los nuevos binarios a sus antepasados casi en tiempo real, lo que demuestra que el rastreo a nivel de familia puede automatizarse cuando la telemetría de los *honeypots* es suficientemente rica.

Nakao *et al*. consolidan los datos de los sumideros de la *darknet* y las fuentes de *honeypots* para trazar un mapa de las tendencias globales de exploración (los puertos 23/TCP, 80/TCP y 443/TCP siguen siendo dominantes) y abogan por agentes denominados NICTER distribuidos globalmente para vigilar los abusos emergentes en los bordes de 5G [15]. Su estudio de caso confirma que las redes de observación coordinadas amplifican la utilidad de los *honeynets* individuales para la alerta temprana y la elaboración de políticas.

A pesar de estos avances, persisten tres retos clave. El *fingerprinting* de *honeypots*, la escalabilidad de los modelos IA/*Machine Learning* en dispositivos *edge* y la escasa cobertura de protocolos emergentes como BLE o LoRaWAN. Además, los *pipelines* actuales dependen de datos etiquetados manualmente, presentan baja interpretabilidad y no están integrados de forma *end-to-end*. Frente a estas limitaciones, este trabajo propone una arquitectura automatizada que cubre desde la captura multiprotocolo hasta el bloqueo dinámico de IPs, operando en tiempo real sin intervención humana y avanzando hacia una inteligencia de amenazas IoT completamente operativa y reproducible.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

III. METODOLOGÍA

La metodología se organiza en cuatro fases principales: (1) diseño de la infraestructura IoT, (2) generación y recolección de datos, (3) ingestión y correlación de eventos y (4) automatización de respuesta ante amenazas.

A. Diseño de la plataforma IoT

La plataforma fue desplegada en una máquina virtual Ubuntu LTS sobre Azure, dentro de una red virtual dedicada. Se instaló T-Pot como entorno base, incorporando contenedores *docker* para emular servicios relevantes: Cowrie (SSH/Telnet), Honeytrap (HTTP/HTTPS), Glutton (protocolos TCP genéricos) y Dionaea (especializado en MQTT y CoAP). Cada contenedor expone sus servicios tras un proxy Nginx y persiste *logs* y binarios en volúmenes *docker*, además de almacenar *payloads* en un contenedor centralizado de *Azure Blob Storage* (*iot-honeypot-binaries*).

B. Recolección de datos

El tráfico malicioso fue inducido desde una máquina Kali Linux mediante técnicas como escaneos (*nmap*), fuerza bruta (*hydra*) y descarga de *payloads* (*wget*, *curl*, *tftp*). Se establecieron sesiones interactivas para evaluar la activación de *honeytokens*. Suricata, configurado en modo promiscuo, permitió registrar eventos de alerta (*event_type:alert*), flujos (*event_type:flow*) y metadatos de archivos (*fileinfo*). Las trazas de red completas se archivaron como archivos .pcap semanales.

C. Ingestión y correlación

El pipeline de *Logstash* recibe los eventos de *Filebeat* y los procesa mediante filtros personalizados, extrayendo campos clave como usuario, contraseña, hashes y URLs de descarga. Los eventos se normalizan y se envían a Elasticsearch, donde son accesibles a través de Kibana. Desde allí se exploran índices como cowrie-, suricata- y iot-corr-, que permiten consultar credenciales, *malware*, descargas y geolocalización. También se generan *dashboards* y se ejecutan *Watchers* para activar respuestas automatizadas en caso de actividad anómala.

D. Automatización de respuesta

Un *Watcher* monitoriza eventos de *login* fallidos (*cowrie.login.failed*) y bloquea IPs con más de 20 intentos en 60 segundos mediante la inserción de reglas Deny en Azure Firewall, usando Azure CLI. Asimismo, los hashes SHA-256 desconocidos se envían automáticamente a la plataforma Any.Run, cuyos resultados se indexan en Elasticsearch para su correlación.

E. Extensión a MQTT/CoAP

El soporte para MQTT y CoAP se integró mediante los módulos *mqtt_module* y *coap_module* de Dionaea. Los mensajes capturados se procesan con los mismos filtros de *Logstash* y se identifican mediante los campos *dionaea.mqtt.payload* y *dionaea.coap.payload*. Esta ampliación refuerza la arquitectura modular y automatizada de la plataforma, permitiendo simular entornos IoT reales y detectar amenazas a través de múltiples vectores.

IV. RESULTADOS

Durante una semana de pruebas (22–27 de mayo de 2025) la plataforma registró un alto volumen de actividad maliciosa que se detalla a continuación.

A. Volumen y distribución de ataques

Durante la semana de pruebas la plataforma registró un total de 74.040 conexiones entrantes, donde Cowrie (SSH/Telnet) representó el 99,3 % con 73.785 eventos, seguido por Honeytrap (HTTP/HTTPS) con 153 eventos (0,2%), Tanner con 70 (0,1%) y otros servicios como CiscoASA y Miniprint con apenas 5 eventos (0,01%).

El vector principal de ataque fue Telnet (puerto 23), seguido por SSH (puerto 22), HTTPS (puerto 443) y HTTP (puerto 80). La distribución geográfica mostró que más del 65 % de las conexiones procedieron de IPs ubicadas en China, EE. UU. y Corea del Sur.

La Fig. 1 resume visualmente el número de eventos capturados por cada *honeypot* desplegado, mientras que la Fig. 2 presenta la evolución temporal de los intentos de acceso detectados por Cowrie.

B. Intentos de autenticación y credenciales

De los 20.279 intentos de autenticación registrados por Cowrie, 19.667 resultaron fallidos (97 %) y solo 612 fueron exitosos (3 %). El análisis de credenciales reveló un patrón altamente predecible: los cinco pares usuario:contraseña más comunes fueron root:root (50 %), admin:admin (25 %), root:123456 (15 %), admin:1234 (7 %) y guest:guest (3 %).

La Fig. 3 muestra los nombres de usuario más utilizados durante los ataques, mientras que la Fig. 4 presenta las contraseñas probadas con más frecuencia por los atacantes.

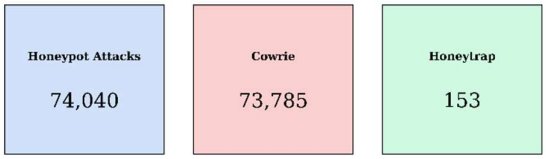


Fig. 1. Resumen de eventos capturados por los *honeypots* desplegados en el entorno T-Pot.

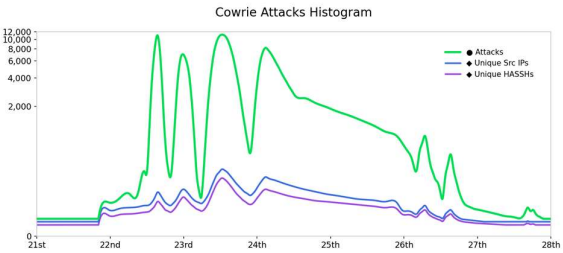


Fig. 2. . Distribución temporal de los intentos de acceso detectados por Cowrie.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

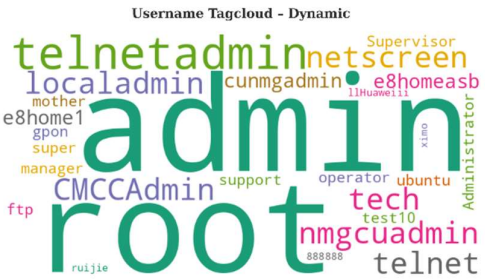


Fig. 3. . Nombres de usuario más frecuentes en ataques contra Cowrie.

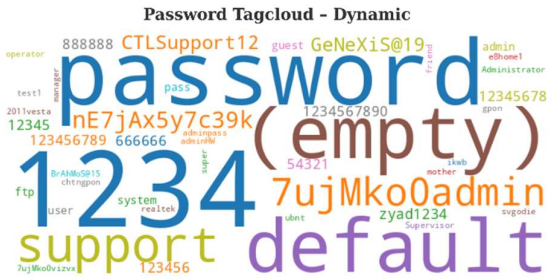


Fig. 4. . Contraseñas más utilizadas en ataques registrados.

C. *Descargas de payloads y hashing*

Se registraron más de 330 eventos de descarga de *malware* a través de Cowrie, todos vinculados al evento *cowrie.session.file_download* y concentrados en cinco URLs principales (véase Fig. 5). Los archivos descargados, almacenados en */data/cowrie/downloads/*, incluían binarios típicos de *botnets* y *scripts* automatizados, muchos de ellos codificados en Base64. Una ´única dirección IP generó más de 19.000 solicitudes, muchas de las cuales correspondieron a intentos repetitivos de obtención de información o ejecución de *scripts* automatizados, sin implicar necesariamente la descarga de *malware*. No obstante, algunos binarios maliciosos sí fueron solicitados entre 100 y 200 veces, lo que evidencia una estrategia activa de propagación de *payloads* por parte de *botnets*.

Del análisis de estas muestras se calcularon 23 hashes SHA-256 únicos: el 80 % coincidió´ con variantes conocidas de Mirai y Gafgyt, mientras que el 20 % restante correspondió a muestras inéditas, enviadas a *sandbox* para su análisis dinámico. Estos resultados confirman una combinación de distribución masiva y experimentación de nuevos binarios maliciosos dentro del entorno IoT simulado.

D. *Detección de botnets*

El IDS Suricata generó un total de 3.521 alertas distribuidas en 121 alertas clasificadas como Mirai, 1.100 alertas de tipo Gafgyt (ET DOS Gafgyt) y 2.300 flujos UDP *flood* reconstruidos, confirmando la presencia activa de estas familias de *botnet* en el tráfico capturado.

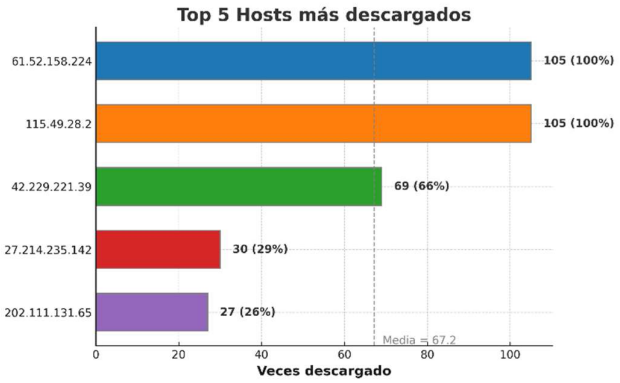


Fig. 5. . Principales URLs utilizadas para descargar *payloads* maliciosos detectas por Cowrie.

E. *Honeytokens y comandos de sesión*

Durante las sesiones Telnet/SSH se ejecutaron 10.116 comandos, de los cuales 860 correspondieron a accesos a archivos señuelo (aproximadamente 8,5% de los comandos totales). La latencia media de ejecución de *honeytokens* fue de 45 ms, demostrando la efectividad del sistema de engaño implementado.

F. *Categorías de alertas IDS*

Durante el periodo de observación, Suricata generó un total de 153.365 alertas, de las cuales las más relevantes asociadas a actividad maliciosa fueron:

- Descarga de ejecutables ELF (ID 2000418-ET INFO *Executable and linking format* ELF *file download*): 870 alertas registradas, la mayoría asociadas a binarios de Mirai/Gafgyt descargados mediante *wget*/*curl*.
- Escaneo automatizado de MySQL (ID 2001689-ET WORM *Potential* MySQL *bot scanning for* SQL *server*): 739 alertas que evidencian intentos de propagación de la *botnet* contra puertos 3306 abiertos.
- Tráfico hacia direcciones internas reservadas (ID 2002752-ET INFO *Reserved Internal* IP *Traffic*): 703 eventos, típicos de *shells* de Mirai que intentan pivotar o evaluar la red local.
- User-Agent *curl* anómalo (IDs 2013028 y 2034567-ET INFO/HUNTING *curl User-Agent*): 1.038 alertas combinadas, asociadas a descargas de *payloads* en *scripts* automatizados.
- Comandos BusyBox sospechosos (IDs 2023016-2023018-ET HUNTING SUSPICIOUS *Busy Box*): 1.446 eventos que revelan ejecución remota de *busybox shell* y *enable*, pasos previos habituales en la instalación de Mirai.

En conjunto, estos indicadores directamente ligados a *malware* suman 5.071 alertas, lo que supone aproximadamente el 3,3% del total generado por Suricata. Aun así, confirman la actividad sostenida de Mirai/Gafgyt en el entorno IoT. Las categorías de control de protocolo (*Generic Protocol Communication*, IDs 2210045, 2210029,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

2210065, 2210041) siguieron concentrando la mayor parte del tráfico anómalo (80%), mientras que los intentos de privilegio administrativo (GPL TELNET *Bad Login*, ID 2101251) sumaron 19.444 alertas (12,7 %).

G. Rendimiento de la plataforma

Los contenedores registraron consumos medios de recursos dentro de parámetros aceptables. Cowrie utilizó aproximadamente 8% de CPU y 360 MB de RAM, Honeytrap requirió 12% de CPU y 430 MB de RAM, mientras que Dionaea (MQTT/CoAP) consumió 7% de CPU y 290 MB de RAM. La latencia media de ingestión en ELK, calculada como la diferencia entre @timestamp y *event.ingested*, fue de aproximadamente 3 segundos.

Estos resultados demuestran la capacidad de la plataforma para capturar, procesar y correlacionar grandes volúmenes de datos IoT y responder de forma automática y eficiente.

V. Conclusiones

Este estudio demuestra la viabilidad y utilidad práctica de una plataforma de *honeypots* IoT desplegada en Azure, capaz de capturar y procesar grandes volúmenes de ataques en tiempo casi real. Durante una semana de observación se registraron aproximadamente 74.000 conexiones maliciosas, se descargaron más de 330 binarios de las familias Mirai/Gafgyt y se bloquearon de forma automática IPs agresoras mediante Azure Firewall. La integración de Suricata y ELK proporcionó una latencia media de ingestión de aproximadamente 3 s, mientras que los *honeytokens* revelaron tácticas de post-explotación en el 8 % de los comandos ejecutados. Hemos visto como se ha logrado una automatización completa del ciclo detección-respuesta (captura, clasificación de malware y bloqueo de fuentes). También se ha dado cobertura multiprotocolo (SSH/Telnet, HTTP/S, MQTT, CoAP) con contenedores ligeros y todo con un modelo de datos unificado que facilita análisis históricos y creación de alertas en Kibana. Como limitaciones cabe indicar que la emulación aún es vulnerable a *fingerprinting* avanzado. Además, la plataforma cubre pocos protocolos emergentes. Como trabajo futuro, se plantea explorar la emulación resistente a *fingerprinting*, la inclusión de detección basada en aprendizaje profundo para variantes *zero-day* y la federación de instancias para intercambio seguro de *threat-intel* en tiempo real.

Agradecimientos

Este trabajo se enmarca en el proyecto “Laboratorio de I+D+i en Ciberseguridad, Privacidad y Comunicaciones Seguras” (TRUST Lab), financiado por el Plan de Recuperación, Transformación y Resiliencia de la Unión Europea (NextGenerationEU) a través del INCIBE.

Referencias

[1] Cisco Systems, Cisco Annual Internet Report (2018 – 2023) White Paper, San Jose, CA, EE. UU., mar. 2020. [En línea]. Disponible en: https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report.html

[2] C. Kolias, G. Kambourakis, A. Stavrou y J. Voas, “DDoS in the IoT: Mirai and Other Botnets,” Computer, vol. 50, no. 7, pp. 80–84, Jul. 2017, doi: 10.1109/MC.2017.201.

[3] R. Lohiya y A. Thakkar, “Application Domains, Evaluation Data Sets, and Research Challenges of IoT: A Systematic Review,” IEEE Internet of Things Journal, vol. 8, no. 11, pp. 8774–8798, Jun. 2021, doi: 10.1109/JIOT.2020.3048439.

[4] S. Lakshminarayana and P. S. Thilagam, “Next-Generation DDoS Attacks on IoT Deployments: Targeting the Advanced Features of MQTT v5.0 Protocol,” IEEE Internet of Things Journal, Early Access, Mar. 2025, doi: 10.1109/JIOT.2025.3549784.

[5] The Conpot Team, “Conpot ICS/SCADA Honeypot,” Conpot Project, 2014. [Online]. Available: https://conpot.org/

[6] The Honeynet Project, “Dionaea: A honeypot designed to trap malware,” Dionaea Project, 2012. [Online]. Available: http://dionaea.carnivore.it/

[7] M. Guarnizo, B. K. Samaniego, J. Echeverría, “ThingPot: An XMPP-based honeypot for IoT devices,” in Proc. of 2nd IEEE Workshop on Security of Emerging Networking Technologies (SENT), 2014. doi: 10.48550/arXiv.1807.04114

[8] A. Pa, M. Murata, and T. Miyoshi, “IoTPOT: Analysing the rise of IoT compromises,” in Proc. of the 9th USENIX Conference on Offensive Technologies (WOOT), 2015. Disponible en: < https://www.usenix.org/system/files/conference/woot15/woot15-paper-pa.pdf >

[9] J. Franco, A. Aris, B. Canberk and A. S. Uluagac, “A Survey of Honeypots and Honeynets for Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2351-2393, 2021. doi: 10.1109/COMST.2021.3106669

[10] W. Zhang, B. Zhang, Y. Zhou, H. He and Z. Ding, “An IoT Honeynet Based on Multiport Honeypots for Capturing IoT Attacks,” IEEE Internet of Things Journal, vol. 7, no. 5, pp. 3991-4002, May 2020. Doi: 10.1109/JIOT.2019.2956173

[11] S. S. Chakkaravarthy et al., “Design of Intrusion Detection Honeypot Using Social Leopard Algorithm to Detect IoT Ransomware Attacks,” IEEE Access, vol. 8, pp. 169944-169962, 2020. Doi: 10.1109/ACCESS.2020.3023764

[12] Q. D. La, T. Q. S. Quek, J. Lee, S. Jin and H. Zhu, “Deceptive Attack and Defense Game in Honeypot-Enabled Networks for the Internet of Things,” *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1025-1039, Dec. 2016. doi: 10.1109/JIOT.2016.2547994

[13] M. Du and K. Wang, “An SDN-Enabled Pseudo-Honeypot Strategy for Distributed Denial of Service Attacks in Industrial Internet of Things,” IEEE Transactions on Industrial Informatics, vol. 16, no. 1, pp. 648-657, Jan. 2020. doi: 10.1109/TII.2019.2917912

[14] H. Wang et al., “An Evolutionary Study of IoT Malware,” IEEE Internet of Things Journal, vol. 8, no. 20, pp. 15422-15438, Oct. 2021. doi: 10.1109/JIOT.2021.3063840

[15] K. Nakao, D. Inoue and K. Yoshioka, “Unveiling IoT Threats: A Case Study on Darknet and Honeypot Analysis,” *IEEE Computer*, vol. 58, no. 4, pp. 37-45, Apr. 2025. doi: 10.1109/MC.2025.3531364

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Desarrollo de un sistema eficiente para asignar bloques de recursos en Slices 5G, utilizando Aprendizaje Por Refuerzo

Manuel Garde, Luis Diez, Ramón Agüero
Departamento de Ingeniería de Comunicaciones,
Universidad de Cantabria
Plaza de la Ciencia s.n. 39005, Santander, Cantabria. España.
manuel.garde@alumnos.unican.es, {ldiez, ramon}@tlmat.unican.es

La quinta generación de redes móviles, *5G*, tiene como uno de sus principales objetivos habilitar la conectividad masiva entre dispositivos que requieren diferentes tipos de calidad de servicio, ya sea latencia, rendimiento o fiabilidad. Es por ello que en 5G se adopta el concepto de *Network Slicing* en la red de acceso, permitiendo aislar recursos de comunicación para diferentes servicios. Así, una correcta asignación de dichos recursos para cada *slice* es fundamental para conseguir la calidad de servicio necesaria, evitando entre otros efectos saturación y degradación de la transmisión. Este trabajo propone el uso de algoritmos de Inteligencia Artificial, basados en Aprendizaje Reforzado, para realizar una asignación dinámica de recursos basada en el estado de las colas de tráfico. Para ello se estima, en tiempo real, el número óptimo de recursos necesarios para mantener la ocupación de los búferes de cada *slice*, analizando la adaptabilidad de este tipo de soluciones.

Palabras Clave—Inteligencia Artificial, Machine learning, Aprendizaje por Refuerzo, Network Slicing, Bloques de Recursos, Colas de tráfico

I. INTRODUCCIÓN

La transición hacia las redes móviles de quinta generación (5G) desde generaciones anteriores, especialmente su predecesora 4G, permite, entre otros aspectos, una conectividad masiva entre dispositivos y máquinas. Este avance hace posible la automatización, el control y la monitorización de procesos que antes eran inviables debido a las limitaciones tecnológicas existentes.

Como consecuencia directa del aumento en la capacidad y flexibilidad de las redes, han aparecido nuevos casos de uso, que demandan requisitos diferentes entre sí en términos de latencia, capacidad y fiabilidad. Por ejemplo, las *Ultra Reliable Low Latency Communications* (URLLC) necesitan conexiones ultra-confiables, con muy poca latencia; las *Enhanced Mobile Broadband* (EMBB) requieren un ancho de banda mayor que el ofrecido por 4G; y para las *Massive Machine-Type Communications* (MMTC) se deben conectar al menos un millón de dispositivos por kilómetro cuadrado.

Una posible solución que se plantea es compartir una única red física para dar soporte a los diferentes flujos de tráfico, que se diferenciarán a través de sus cabeceras, identificadores o aplicando procesos de personalización a cada conexión. Sin embargo, este enfoque presenta una limitación crítica: ante una posible saturación de la red en uno de los flujos de tráfico, el rendimiento global se vería degradado, lo que provoca que los requisitos de los demás servicios se vean afectados.

Por esta razón, el uso de *Network Slicing* (NS) en las redes 5G está cobrando cada vez una mayor relevancia. Esta técnica, formalmente introducida por NGMN Alliance en su *NGMN 5G White Paper* [1], se propuso como solución para gestionar distintos tipos de aplicaciones. Más tarde, el 3GPP formalizó el concepto para la red de acceso en el *Release 15* [2], estableciendo el procedimiento necesario para crear diversas redes lógicas sobre una única infraestructura física de red.

La técnica de NS permite segmentar la red en instancias virtuales independientes, cada una de ellas diseñadas para responder a una demanda de tráfico con requisitos específicos, como alta velocidad, baja latencia o gran volumen de comunicaciones simultáneas. Esta solución asegura una asignación dedicada de recursos, garantizando aislamiento entre *slices* y evitando la interacción entre tráficos de aplicaciones/servicios diferentes.

En el ámbito de las redes de acceso, el concepto de *slice* (o RAN *slice*) se aplica al aislamiento de recursos radio, o *Resource Blocks* (RBs), relacionado con las gestión de recursos de capa MAC (*MAC scheduling*). La especificación de RAN *slice* permite la asignación estática de recursos a diferentes servicios. Sin embargo, un enfoque tan rígido

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

puede limitar el uso eficiente de los recursos, debido a su escasa capacidad de adaptación, tanto a cambios en el volumen de tráfico como a fluctuaciones de la calidad de canal.

La rigidez mencionada anteriormente puede producir sobre-asignaciones o infra-asignaciones innecesarias de recursos, comprometiendo los requisitos de eficiencia y calidad del sistema. Por otro lado, la nula capacidad de aprendizaje previene la optimización en tiempo real de la toma de decisiones, lo que puede comprometer la calidad de los servicios.

Para abordar estas limitaciones, este trabajo explora el uso de soluciones basadas en Inteligencia Artificial (IA), concretamente algoritmos de aprendizaje reforzado (RL). El objetivo es realizar una asignación dinámica de RBs que se encuentre dentro de los márgenes estáticos definidos para cada *slice*, pero que a su vez tenga en cuenta condiciones del canal y el tráfico generado, con el fin de mejorar la gestión de estos recursos.

El resto del documento se estructura como sigue. En la Sección II se realiza una revisión de trabajos relacionados. Posteriormente, en la Sección III se introduce el concepto de *slice*, así como el entorno que se utilizado para analizar el rendimiento de la solución propuesta. Las técnicas desarrolladas se presentan en las Secciones IV y V. Posteriormente, la Sección VI describe los resultados objetivos tras su evaluación. El artículo concluye en la Sección VII, que también identifica posibles líneas futuras que surgen a partir de este trabajo.

II. TRABAJO PREVIO

La adopción del concepto de NS se ha realizado en una amplia variedad de casos de uso, entre los que se puede destacar su aplicación en escenarios de comunicaciones por satélite. En este sentido, Pang Jinkun y Sun Qibo [3] lo aplicaron en un entorno *Low Earth Orbit* (LEO), instanciando múltiples *slices* en un único satélite, para la que se utilizó técnicas de virtualización con Docker y tecnologías de redes tolerantes a retrasos (*Delayed Tolerant Networks, DTN)*. La solución presentada permite la gestión de múltiples flujos de tráfico para aplicaciones IoT, además de servicios de entretenimiento y web.

Otro escenario de aplicación de NS es el de *Vehicle to Everything* (V2X). En este ámbito, Claudia Campolo *et al.* [4] abordan cómo la técnica de NS puede integrarse eficientemente en redes 5G para soportar distintos requisitos de comunicación asociados a escenarios V2X, tales como baja latencia, alta fiabilidad o movilidad extrema. Teniendo en cuenta la necesidad de disponer de una solución flexible y rápida, resulta razonable considerar que el NS puede abordarse mediante técnicas de *Machine Learning* (ML), con el objetivo de aumentar la eficiencia de la comunicación y acelerar la adaptación de la red a los requisitos individuales del tráfico. Siguiendo este razonamiento, Zong-Xun Wu y Yun-Zhe You [5] plantean una solución basada en el aprendizaje de los patrones de tráfico de red para clasificar las correspondientes solicitudes dinámicamente y optimizar así la asignación de recursos.

Una de las técnicas más comunes de ML para la gestión de los recursos son las de RL. Por ejemplo, Mason *et al.* [6] aplican *Distributed Deep Reinforcement Learning* (DDRL) para orquestar recursos en escenarios con NS. Concretamente hacen uso de agentes distribuidos y el algoritmo *Advantage Actor-Critic* (A2C) para asignar recursos de forma dinámica y eficiente, adaptándose a las condiciones cambiantes de la red. Los resultados muestran mejoras frente a métodos estáticos, destacando el potencial del DRL para la gestión inteligente de *slices*.

Bajo el marco de la arquitectura ORAN, se desarrolla DRAGON [7], que combina el uso de NS con técnicas de DRL para aumentar la flexibilidad, escalabilidad e interoperabilidad en el uso de comunicaciones *multiple-input multiple-output* (MIMO). El enfoque se basa en la predicción del esquema de modulación y codificación (Modulation and Coding Scheme, MCS) y el número de flujos de transmisión necesarios para satisfacer los requisitos de SNR en cada instante de tiempo. Sin embargo, no existen muchos trabajos que se hayan centrado en la gestión de los recursos bajo el paradigma de RAN *slicing*.

En los últimos años han aparecido diferentes implementaciones abiertas de redes 5G. Entre ellas, OpenAir-Interface (OAI) [8] emerge como una solución estándar *de-facto*, al ofrecer un marco de código abierto para la implementación y validación de redes móviles 4G y 5G. Desarrollado por OpenAirInterface Software Alliance, proporciona un software conforme a los estándares para la red de acceso y el (CN), permitiendo un despliegue completo de red sobre componentes comerciales físicos. OAI se ha utilizado en los últimos años en multitud de soluciones, tanto para desplegar redes 5G funcionales como para evaluar el rendimiento de soluciones, entre las que destacan aquellas dedicadas a la gestión de recursos. En el ámbito concreto de RAN-slice, Hai Cheng *et al.* presentan ORANSlice [9], que proporciona una implementación funcional del concepto de RAN *slice* sobre OAI y su integración con el entorno O-RAN. En concreto en el trabajo se monitoriza la gestión de recursos entre *slices* mediante una xApp que se comunica con la estación base 5G a través de *near-RT RIC*.

Como se puede observar, aunque existen trabajos que han abordado la gestión de NS en sentido genérico, existen pocos trabajos que se han centrado en la gestión de recursos en escenarios de RAN *slice*.

III. METODOLOGÍA DE ANÁLISIS

En esta sección se introduce de manera más detallada el concepto de *slice*. Posteriormente se presentan las características principales del entorno de trabajo diseñado para el estudio del rendimiento de técnicas de gestión.

A. *RAN* slice

Cada *slice* se define a través de un identificador, tal como se muestra en la Figura 1, llamado *Single Network Slice Assistance Information* (S-NSSAI). A su vez, este identificador está compuesto por dos campos: el *Slice/Service Type* (SST), que define el tipo de tráfico

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

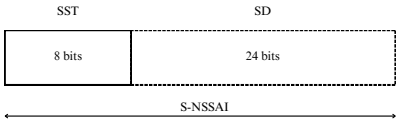


Fig. 1. Estructura de trama de un mensaje S-NSSAI

contenido en la *slice*, y el *Slice Differentiator* (SD), que permite la diferenciación entre *slices* con el mismo SST.

El 3GPP, en su especificación TS 23.501 [2], ha estandarizado valores de SST en el intervalo $[1 \dots 127]$, asignando significados específicos a los primeros cuatro valores, los cuales están asociados a casos de uso clave en redes 5G. Estos valores se presentan en la Tabla I, que también intruduce dichos casos de uso, que son los más comunes: eMBB , URLLC , *massive Internet of Things* (mIoT) y V2X. Por otro lado, los valores en $[128 \dots 255]$ están reservados para uso privado por parte de los operadores de telecomunicación, quienes pueden definir sus propios servicios según sus necesidades.

Tabla I
VALORES ESTANDARIZADOS DEL SST.

Slice/Service Type	Valor del SST
eMBB	SST 1
URLLC	SST 2
MIoT	SST 3
V2X	SST 4

Dentro de la especificación también se indica que un único usuario puede estar vinculado hasta a 8 *slices* diferentes de forma simultánea, en donde cada una de ellas se identifica mediante el campo S-NSSAI. La combinación de todos los S-NSSAI de cada usuario genera el NSSAI, que se incluye en el mensaje *Registration Request*, entre el CN y dicho usuario, habilitando a la red la identificación de cada una de las *slices*.

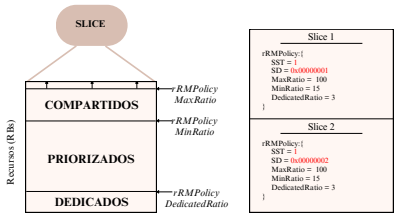


Fig. 2. Estructura de *RRMPolicyRatio*

En el contexto del RAN *slicing*, el 3GPP introduce el concepto de *Radio Resource Management RAN Slicing Policy* (rRMPolicy) [10], que se ilustra en la Figura 2. Permite definir políticas concretas a cada una de las *slices*. Se puede destacar el parámetro `rRMPolicyDedicatedRatio`, que establece el porcentaje de RBs que deben reservarse exclusivamente para una única *slice*, independientemente de si existe tráfico o no en ella.

Asimismo, se definen dos umbrales para gestionar la priorización de recursos: `rRMPolicyMinPolicy` y `rRMPolicyMaxPolicy`. El primero establece el porcentaje mínimo de RBs priorizados que deben estar disponibles para cada *slice*, garantizando un nivel básico de asignación de recursos. El segundo delimita el porcentaje máximo de RBs que se pueden asignar a la *slice*, incluso en condiciones de alta demanda.

Los recursos priorizados asignados a una *slice* pueden compartirse dinámicamente con otras, cuando no existe tráfico activo en dicha *slice*. Sin embargo, cuando hay demanda activa en esa *slice*, estos recursos están garantizados. Por otra parte, los recursos compartidos (aquellos que se encuentran entre los umbrales `rRMPolicyMinPolicy` y `rRMPolicyMaxPolicy`) corresponden a los RBs que no tienen políticas de prioridad, y pueden ser por tanto utilizados libremente por cualquier *slice* en función de la demanda, disponibilidad del sistema y política aplicada.

B. Entorno de análisis

A continuación se describe el diseño seguido para la generación de tráfico y cálculo del *throughput* durante la evaluación de la solución propuesta. A partir de ello se ha implementado un simulador en lenguaje Python que interactúa con las soluciones de DRL que se presentan en las próximas secciones.

El escenario que se utiliza asume una única estación base y tiempo ranurado, donde en cada ranura la estación base toma una decisión de asignación de RBs a cada una de las *slices*, que tienen una configuración individualizada de umbrales (`rRMPolicyDedicatedRatio`, `rRMPolicyMinPolicy` y `rRMPolicyMaxPolicy`). Este enfoque se podría aplicar igualmente a escenarios con múltiples estaciones base. Además de los parámetros definidos por el estándar, para cada una de las *slices* se define un throughput objetivo. Cabe resaltar que, a diferencia del *throughput*, si se utilizara como requisito los RBs asignados, no sería posible asegurar una tasa determinada por *slice*, debido a la dependencia con el estado del canal. Asimismo, en cada ranura temporal el valor de calidad de canal observado por cada *slice* varía. Desde un punto de vista de modelado se asume un estado de canal promedio de los usuarios con servicios de un determinado slice. En cada iteración se genera tráfico de forma aleatoria para cada *slice* siguiendo una distribución uniforme, con función densidad de probabilidad: $p(x) = \frac{1}{b-a}, \quad a \leq x \leq b$, donde a y b son los valores máximos y mínimos.

El tráfico generado por cada *slice* se almacena en un buffer, cuyo tamaño se mide en bytes. A su vez, el número de bytes se convierte a RBs en función de la calidad de canal, tal como se indica en la Ecuación 1:

$$N_{RB} = \frac{N_{Bytes} \cdot 8}{\Delta f \cdot 12 \cdot \eta \cdot T_s} \tag{1}$$

donde Δf corresponde con el ancho de banda de la sub-portadora OFDM, 12 es el número de sub-portadoras en cada RB y η indica la eficiencia espectral en bps/Hz. Finalmente T_s se corresponde con el periodo de símbolo.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Los parámetros Δf y T_s se definen en función de *Sub Carrier Spacing* (SCS), así como el número de subportadoras para cada RB [11]. Por otro lado, η se calcula en cada ranura temporal, siguiendo nuevamente una distribución uniforme, con límites 0.2344 y 7.4063, respectivamente [12], representando la calidad mínima y máxima para conexiones 5G.

Finalmente, el *throughput* que cada *slice* puede alcanzar, se calcula de forma similar como $Th = \frac{N_{RB_{asign}} \cdot \Delta f \cdot 12 \cdot \eta \cdot T_s}{T}$.

En este caso, se hace uso del número de RBs asignados por el algoritmo utilizado ($N_{RB_{asign}}$) y el parámetro T indica el tiempo entre decisiones consecutivas. En general se asumirá $T = 0,01$, siendo una estimación de las fluctuaciones de tráfico producidas en las *slices*, asumiendo que se mantiene constante durante dicho periodo.

Cabe recordar que, en cualquier caso, se asigna un umbral de tasa mínima a cada una de las *slices*, que debe ser capaz de asumir independientemente de la existencia o no de tráfico en cada instante temporal. Con esta modificación se pretende replicar el objetivo de calidad de servicio de las conexiones 5G.

IV. ESQUEMA DE GESTIÓN CON DRL

Esta sección describe la integración de la herramienta de simulación presentada anteriormente con un entorno que permita el uso de agentes DRL. En concreto, se ha hecho uso del entorno Gymnasium [13] para la implementación de los agentes correspondientes. Se trata de una librería de Python diseñada para facilitar la creación, prueba y comparación de algoritmos de RL. Gymnasium proporciona una interfaz unificada para cientos de entornos, permitiendo que los algoritmos desarrollados se prueben sin preocuparse por los detalles internos de cada entorno.

Por otro lado, Gymnasium permite una fácil integración con librerías como *Stable-Baselines3* (SB3), que proporciona múltiples implementaciones de algoritmos de RL basados en redes neuronales. Por esta razón, se ha elegido esta implementación para abordar el problema planteado anteriormente, combinando la creación del módulo DRL mediante Gymnasium con los algoritmos basados en redes neuronales que ofrece SB3.

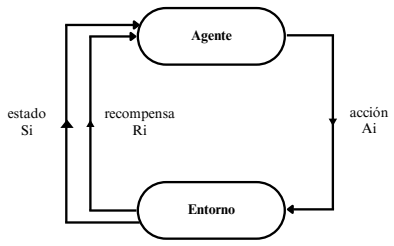


Fig. 3. Estructura básica de un modelo de RL

La Figura 3 representa la estructura básica de un modelo de RL, en el cual un agente interactúa con el entorno. En un instante de tiempo *t*, el agente recibe una observación del estado del entorno, S_t, y procede a ejecutar una acción A_t. Tras aplicar esta acción en el entorno pasa a un nuevo estado S_{t+1} y se obtiene una recompensa R_t, relacionada con la acción aplicada. Esta recompensa cuantifica el resultado de la acción tomada por el agente en el instante de tiempo anterior. Finalmente, este ciclo continúa y el agente trata de aprender una política que maximice la recompensa instantánea a lo largo del tiempo.

A. Entorno

Este trabajo tiene como objetivo asignar correctamente un número de RBs a cada uno de los *slices* sin exceder el número máximo de recursos disponibles en el sistema. Centrando la atención en las observaciones recibidas por el agente, se consideran tres elementos clave: *estimated_RBs*, que representa los RBs correspondientes al tráfico alojado en el búfer de cada *slice*; *max_allowed_RBs*, que informa al agente del número total máximo de RBs que se pueden asignar en total y, finalmente, *min_max_RBs*, que define el número mínimo y máximo de RBs que pueden asignarse a cada *slice* de forma individual.

Dado que el agente debe predecir un número entero de RBs a asignar, se ha utilizado un espacio de acciones discreto múltiple (*Multi Discrete*) proporcionado por la biblioteca *Spaces* de Gymnasium. Este espacio permite representar múltiples acciones discretas, devolviendo un valor entero correspondiente al número de RBs asignados para cada *slice* de la red.

Como se mencionó anteriormente, SB3 ofrece varios modelos de DRL. Sin embargo, no todos ellos son compatibles con todos los espacios de acciones. Para ampliar el alcance del estudio, el mismo entorno también se ha desarrollado en paralelo utilizando el tipo de acciones *Box*, que da lugar a un espacio de acciones continuo dentro de intervalos definidos.

B. Inicialización: función reset

Para iniciar el proceso de aprendizaje del agente, es necesario definir un estado inicial que permita configurar correctamente el entorno de simulación. Durante esta fase, se establecen los parámetros clave del entorno, incluyendo la eficiencia espectral asociada a cada *slice*, el número mínimo y máximo de RBs disponibles y, finalmente, la estimación del resto de recursos compartidos entre todas las *slices*. Este último parámetro se calcula restando del total de RBs disponibles en la red la suma de los RBs previamente asignados, bajo la suposición de que ya se han asignado los recursos dedicados obligatorios para el tráfico esperado.

C. Ejecución: función step

La función *step* se encarga de evaluar las predicciones del agente y actualizar el entorno para modificar el estado. A partir de la acción proporcionada por el modelo, se calculan distintos tipos de penalizaciones y recompensas:

1) Penalización por sub-asignación: la penalización en este caso se produce cuando, habiendo tráfico en una *slice*, el número de RBs es inferior al mínimo disponible (*rRMPolicyMinPolicy*). Se define a través

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

del parámetro γ_1 como sigue:

$$\gamma_1 = \max\{\sum_{i=1}^{N^\circ \text{Slices}} RBs_{min_i} - RBs_i, 0\} \quad (2)$$

donde RBs_{min_i} y RBs_i es el número mínimo y asignado a la *slice* i.

2) Penalización por sobre-asignación: si el agente asigna un número de RBs por encima del máximo permitido (*rRMPolicyMaxPolicy*), el agente recibe la penalización mostrada en Ecuación 3:

$$\gamma_2 = -2 \cdot \sum_{i=1}^{N^\circ \text{Slices}} RBs_i - RBs_{max_i} \quad (3)$$

donde RBs_{max_i} indica el número máximo de RBs correspondiente a la slice i.

3) Penalización por terminación (γ_3): este parámetro tiene valor 0 excepto cuando el buffer de alguna de las *slices* se encuentra saturado. Esto ocurre cuando las colas de los búferes han alcanzado un estado de saturación (fijado en 10^8 bytes) en alguna de las *slices*. Cuando el agente entra en este estado el búfer saturado se vacía completamente y el entorno se reinicia. Con este enfoque, el agente es entrenado para evitar la saturación.

1) Recompensas: Para la definición de recompensas se ha utilizado como base para fijar los objetivos el algoritmo implementado sobre OAI por ORANSlice [9], que sigue el siguiente funcionamiento:

- En primer lugar, se realiza una estimación del número total de RBs necesarios para vaciar completamente las colas del buffer. Tras ello, se ordenan las *slices* de forma ascendente, en función del número de RBs calculado.
- Una vez están ordenadas las *slices*, se calcula el número de RBs que realmente requiere la primera *slice* y se le asignan todos los recursos necesarios, que a su vez son descontados de los totales disponibles en la red. Una vez satisfecho el requisito de la primera, se prosigue con las demás en orden hasta que se acabe la demanda de recursos o hasta que no queden recursos disponibles.

Utilizando la misma lógica se ha definido una recompensa parcial que busca obtener una política similar y se ha complementado con otra adicional. Posteriormente la recompensa total se calcula aglutinando las penalizaciones y cada una de las siguientes recompensas parciales:

1) Recompensas por estabilidad (R_1): con el fin de entrenar al agente para asegurar la estabilidad del buffer de cada *slice*, la predicción debe ser lo más precisa posible. Esto significa predecir el número de RBs que son necesarios para vaciar el buffer, evitando así su crecimiento a lo largo del tiempo. Por esta razón, el agente recibe una recompensa basada en la estabilidad que proporciona su predicción, tal como se muestra en la Ecuación 4.

$$R_1 = \sum_{i=1}^{N^\circ \text{Slices}} [2 - \log_2 \left(1 + \max(0, |RBs_i - RBs_{est_i}| - 10)\right)] \quad (4)$$

donde RBs_{est_i} es el número de RBs necesario para vaciar el buffer del *slice* i.

2) Recompensa por *throughput* (R_{thpt}): para conseguir el requisito de *throughput* de cada *slice*, el agente recibe una recompensa en función de si ha superado el umbral establecido o no.

$$R_i = \begin{cases} 5 & \overline{T}_i > G_i + 10^3 \\ -5 & \overline{T}_i \leq G_i + 10^3 \end{cases} \quad (5)$$

donde G_i es el requisito de *throughput* del *slice* y T_i indica el *throughput* que se ha conseguido como consecuencia de la asignación. Este último parámetro se define como una media móvil a lo largo de un número configurable de ranuras temporales (el caso particular de utilizar una ranura correspondería con el cálculo de *throughput* instantáneo), $\overline{T}_i = \frac{1}{i}\sum_{j=1}^{i_{prev}} T_{ij}$. Finalmente la recompensa total por *throughput*, R_{thpt}, se calcula como la suma de las de cada *slice*.

Para incrementar la fiabilidad del sistema, el cálculo de las recompensas incorpora un *offset* (10^3). Esta opción fuerza al agente a no otorgar únicamente los recursos necesarios, si no a asignar más, con el objetivo de satisfacer el requisito, a pesar de que el canal se encuentre en condiciones desfavorables.

Finalmente, la función *Step* actualiza el estado del entorno, eliminando del buffer de cada *slice* el número de bytes correspondiente a los RBs asignados por el agente, simulando así el vaciado de las colas. Después de este vaciado, se genera nuevo tráfico para continuar con el proceso de entrenamiento. Al final de la función, el agente recibe la recompensa acumulada, que se calcula como:

$$R_{total} = -\gamma_1 - \gamma_2 - \gamma_3 + W \cdot R_1 + \delta \cdot R_{thpt} \quad (6)$$

V. REDES NEURONALES

Esta sección presenta los diferentes agentes DRL utilizados para entrenar los que han incluido en el análisis, de la librería SB3, y que han sido seleccionadas en función de la naturaleza del espacio de acciones

Para los espacios discretos se han seleccionado las siguientes opciones:

- ***Advantage Actor Critic (A2C)***
  Este algoritmo de RL combina estrategias de aprendizaje basadas en actor-crítico y en políticas. La red neuronal del actor se encarga de aprender la política del agente, es decir, genera una distribución de probabilidad sobre las posibles acciones en un estado concreto, mientras que el modelo del crítico estima la probabilidad de tránsito hacia otro estado, con el objetivo de calcular la ventaja, definida como la diferencia entre la recompensa real obtenida por el agente y la recompensa esperada, $A(s,a) = Q(s,a) - V(s)$.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- ***Trust Region Policy Optimization (TRPO)***
  La base del algoritmo TRPO es evitar actualizaciones bruscas en la política del agente, es decir, crear regiones de confianza donde los agentes realicen observaciones para maximizar sus recompensas, sin sobrepasar los límites establecidos en cada estado. TRPO optimiza una política π_θ que busca maximizar una función objetivo que representa la ganancia esperada, pero limitando la distancia entre la nueva política y la anterior, mediante el uso de la divergencia de Kullback-Leibler (KL):

$$\max_\theta \hat{\mathbb{E}}[\frac{\pi_\theta(a|s)}{\pi_{\theta'}(a|s)} A^{\pi_{\theta'}}(s,a)] \; ; \; \hat{\mathbb{E}}\left[D_{\mathrm{KL}}\left(\pi_{\theta'} \,\|\, \pi_\theta\right)\right] \leq \delta \tag{7}$$

- ***Proximal Policy Optimization (PPO)***
  La última red neuronal utilizada es PPO, que combina las ideas de A2C usando múltiples agentes, así como los conceptos de TRPO, generando regiones de confianza para evitar actualizaciones agresivas de la política de un paso al siguiente.

Por su parte, los algoritmos seleccionados para espacios continuos son los siguientes:

- ***Deep Deterministic Policy Gradient (DDPG)***
  DDPG se basa en el enfoque actor-crítico, pero trabajando sobre espacios continuos, similar a A2C. Combina las ideas de *Deep Q-Network (DQN)* con el método *Deterministic Policy Gradient (DPG)*.
- ***Twin Delayed DDPG (TD3)***
  TD3 es una mejora directa del algoritmo DDPG, orientada a aumentar la estabilidad del aprendizaje y reducir errores comunes en los agentes. Las principales diferencias respecto a DDPG son tres: (i) TD3 emplea dos redes críticas (Twin Critics) y utiliza el valor mínimo de ambas para actualizar la política, mitigando así la sobreestimación del valor de la acción; (ii) la política no se actualiza en cada paso, sino cada **d** iteraciones, lo que mejora la estabilidad; y (iii) se añade ruido gaussiano a las acciones del actor objetivo durante el cálculo de la actualización, evitando que el agente aprenda políticas sub-óptimas causadas por errores de aproximación.
- ***Soft Actor Critic (SAC)***
  SAC hereda la técnica de Double Q-Learning, utilizada previamente en TD3, con el objetivo de evitar la sobre-estimación de los valores en cada acción. A diferencia de otros algoritmos actor-crítico, el objetivo de SAC es maximizar simultáneamente la recompensa esperada y la entropía de la política. Esto permite que el agente no solo aprenda comportamientos óptimos, sino también políticas más robustas y exploratorias, adecuadas para entornos complejos con un alto grado de incertidumbre.

VI. RESULTADOS

Esta sección presenta los resultados obtenidos de todas las pruebas llevadas a cabo. En primer lugar se realiza un análisis de la capacidad de aprendizaje con todos los agentes, utilizando una función de recompensa simplificada, para posteriormente comparar el rendimiento de alguno de ellos con la recompensa completa que se ha presentado anteriormente. En todos los casos los resultados se han obtenido con 2 slices.

A. Análisis de convergencia

Como primera prueba del entorno de simulación desarrollado, se ha aplicado una simplificación de la recompensa total entregada a los agentes con el objetivo de estudiar la convergencia de cada uno, estudiando si son adecuados para este escenario concreto o, por el contrario, son incapaces de asignar dinámicamente recursos de forma correcta. En este caso el sistema se ha configurado con 106 RBs, y el tráfico generado por cada *slice* está en torno a 30Mbps.

Los valores de W y δ de la Ecuación 6 son 1 y 0, respectivamente. Con esto se pretende que los agentes prioricen el vaciado de los buffer de cada *slice*, intentando no desperdiciar recursos.

En primer lugar, se presentan las recompensas obtenidas por los agentes entrenados sobre espacios discretos (A2C, TRPO y PPO), reflejando cómo de acertadas han sido sus predicciones en cada iteración.

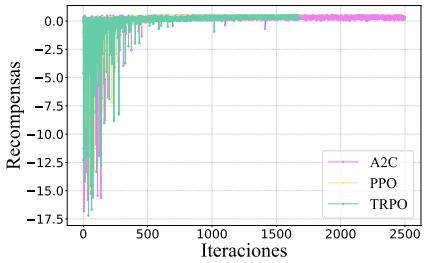


Fig. 4. Recompensas cada 100 iteraciones en espacio discreto

Como se puede observar en la Figura 4, todos los modelos ponen de manifiesto una convergencia adecuada, alcanzando una tendencia estable que todos los agentes logran después de 50000 pasos. Esto implica que los agentes son capaces de aprender a partir de las observaciones y predecir, en primer lugar, una solución correcta.

Como se puede ver en la Figura 5, que muestra las recompensas obtenidas por los agentes entrenados sobre espacios continuos (DDPG, TD3 y SAC), no se aprecia una convergencia clara en los entornos continuos para ningún agente. Los resultados ponen de manifiesto una baja consistencia en las recompensas obtenidas entre iteraciones cercanas. Además, las recompensas obtenidas al inicio y al final del entrenamiento son muy similares, lo que refuerza la conclusión sobre la falta de aprendizaje por parte de los agentes.

Tras completar el proceso de entrenamiento, los agentes son evaluados bajo nuevas condiciones de red, con el fin de analizar su rendimiento y verificar si han aprendido de manera efectiva los patrones necesarios para asignar de manera óptima los recursos correspondientes. Para medir

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

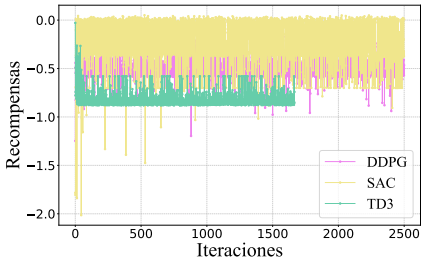


Fig. 5. Recompensas cada 100 iteraciones en espacio continuo

la precisión en la asignación de cada agente, se calcula el error como la diferencia entre los recursos asignados y los estrictamente necesarios: $Error = RBs_{asig} - RBs_{nec}$.

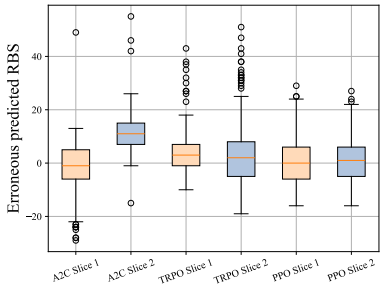


Fig. 6. Error de predicción para cada agente discreto

La Figura 6 representa la distribución del error obtenida por los agentes discretos para cada uno de las *slices* involucradas. Como se puede observar, todos demuestran una alta precisión en sus predicciones, con valores medianos del error concentrados principalmente alrededor de 0.

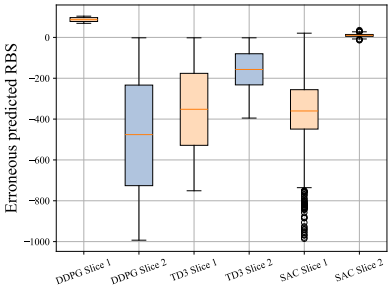


Fig. 7. Error de predicción para cada agente continuo

A diferencia de los agentes discretos, los errores obtenidos en la fase de evaluación de los agentes entrenados en espacios continuos aumentan significativamente. Como se muestra en la Figura 7, predominan los errores superiores a 100 RBs, así como distribuciones de error ampliamente dispersas, resultados que se corresponden con el rendimiento observado durante el entrenamiento.

Estos resultados no son sorprendentes, ya que el comportamiento observado durante el entrenamiento permite anticipar, en cierta medida, la calidad de las predicciones realizadas por el agente.

B. Análisis de adaptación

A continuación se presenta un análisis en el que se han tenido en cuenta las recompensas relativas al requisito de *throughput* (δ = 1, Ec. 6). Además, se ha modificado el peso relativo entre ambas recompensas, para observar la capacidad de adaptación del agente DRL utilizado a diferentes políticas. Para llevar a cabo esta nueva prueba se ha utilizado el modelo A2C que, como se ha observado anteriormente, es capaz de converger y asignar recursos correctamente.

En este caso se cuenta nuevamente con dos *slices*, pero se aumenta la capacidad del sistema a 306RBs. Además, el requisito de *throughput* para la primera es variable, a fin de observar los cambios en la asignación de recursos que realiza el agente, mientra que para la otra se mantiene constante.

Adicionalmente, para cada nivel de requisito de *throughput*, se modifica el valor del peso W asociado a la recompensa R_1. Esta variación permite analizar cómo, al incrementar la importancia relativa de la recompensa por precisión en la asignación de recursos, el cumplimiento de los objetivos de *throughput* tiende a verse comprometido. Por el contrario, cuando el peso asignado a dicha recompensa disminuye, se espera que el agente priorice el cumplimiento de los objetivos de tasa establecidos.

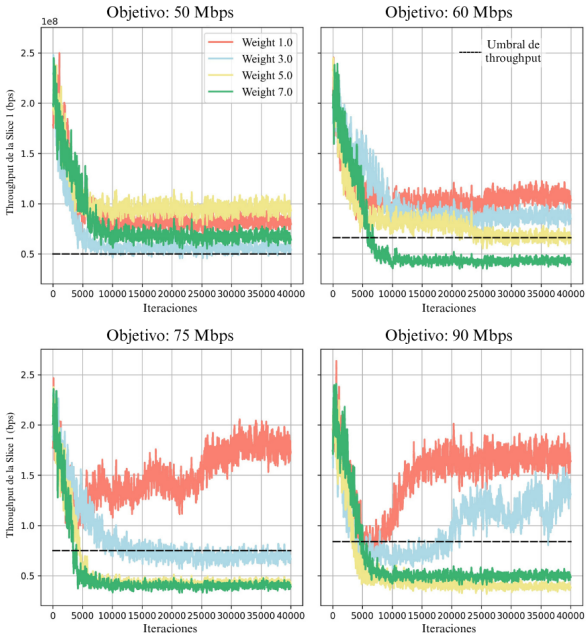


Fig. 8. *Thrpughput* instantáneo a lo largo del entrenamiento

El comportamiento obtenido se ilustra en la Figura 8. Como primera prueba, se establece un objetivo de *throughput* de 50Mbps para la *slice* 1, cuyo tráfico generado presenta una tasa media muy cercana a dicho valor. En este escenario, el modelo no se ve forzado por la restricción de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

rendimiento (throughput), lo que le permite cumplir con ambos requisitos.

En los experimentos posteriores, se incrementa progresivamente el objetivo de *throughput* para la *slice* 1. Se analiza así cómo el agente adapta la asignación de recursos a medida que el cumplimiento del objetivo de tasa se vuelve más exigente. A partir del umbral de 60 Mbps, se observa que el agente entrenado con $W = 7.0$ deja de alcanzar el objetivo de *throughput*, mostrando así el comportamiento esperado: priorización de la precisión frente a la tasa.

Finalmente, al seguir aumentando el requisito de *throughput*, los agentes se enfrentan a un mayor compromiso entre ambos objetivos, lo que conduce a una diferenciación clara entre aquellos que priorizan el cumplimiento del umbral de tasa y aquellos que maximizan la precisión en el vaciado de las colas.

VII. CONCLUSIONES

En este artículo se ha explorado el uso de técnicas de inteligencia artificial para la asignación de recursos en escenario de RAN *slicing*. A través del uso de aprendizaje reforzado, se ha demostrado la capacidad del agente para adaptarse a diferentes condiciones de tráfico, cumpliendo con los requisitos de *throughput* establecidos y evitando la asignación innecesaria de bloques de recursos.

Como primer experimento, se han entrenado satisfactoriamente los agentes de RL sobre espacios discretos, mostrando la capacidad de decidir en base a las condiciones cambiantes del entorno, y amoldar sus predicciones para no sobre-asignar recursos innecesariamente. Por el contrario, los agentes entrenados sobre entornos continuos no han sido capaces de ofrecer el comportamiento deseado.

Este análisis se ha complementado con un segundo experimento en el que se añaden requisitos de *throughput*, de forma que se obtiene una política que extiende las definidas en soluciones similares. Los resultados han demostrado que los agentes DRL son capaces de adaptarse a diferentes configuraciones, utilizando políticas con múltiples parámetros.

El estudio presentado en este trabajo se extenderá de dos formas diferentes. En primer lugar se pretende profundizar en el análisis de soluciones de DRL para la gestión de recursos en entornos RAN *slicing*, mediante el uso de objetivos más complejos y añadiendo técnicas de predicción de tráfico. Como complemento, también se están analizando las requisitos para llevar a cabo una implementación real, utilizando como base la implementación ORANSlice basada en OAI.

AGRADECIMIENTOS

Este trabajo ha sido financiado parcialmente por el Ministerio de Asuntos Económicos y Transformación Digital de España mediante los proyectos SITED (PID2021-125725OB-I00) y 6GBLUR/JOINT y 6GBLUR/SMART (TSI-063000-2021-56/57). También ha recibido ayuda del Gobierno de Cantabria, Fondo Europeo de Desarrollo Regional (FEDER), a través del programa "Ayudas a proyectos de investigación con alto potencial industrial de agentes tecnológicos de excelencia para la competitividad industrial (TCNIC)" (2023/TCN/002).

REFERENCIAS

[1] NGMN, "NGMN 5G White Paper)," Next Generation Mobile Networks (NGMN) Alliance, Tech. Rep., 2015. [Online]. Available: {https://www.ngmn.org/wp-content/uploads/NGMN_5G_White_Paper_V1_0.pdf}

[2] 3GPP, "5G; System architecture for the 5G System (5GS) (Release 18)," 3GPP, Tech. Rep. TS 23.501 V18.9.0 (2025-04), 2025.

[3] P. Jinkun and S. Qibo, "Network slicing in leo satellite network," in *2021 2nd International Conference on Artificial Intelligence and Information Systems*, ser. ICAIIS 2021. New York, NY, USA: Association for Computing Machinery, 2021.

[4] C. Campolo, A. Molinaro, A. Iera, and F. Menichella, "5g network slicing for vehicle-to-everything services," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 38–45, 2017.

[5] Z.-X. Wu, Y.-Z. You, C.-C. Liu, and L.-D. Chou, "Machine learning based 5g network slicing management and classification," in *2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2024, pp. 371–375.

[6] F. Mason, G. Nencioni, and A. Zanella, "Using distributed reinforcement learning for resource orchestration in a network slicing scenario," *IEEE/ACM Transactions on Networking*, vol. 31, no. 1, pp. 88–102, 2023.

[7] Q. An, R. Doost-Mohammady, R. Yang, and K. Sridhar, "Dragon: A drl-based mimo layer and mcs adapter in open ran 5g networks," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, ser. ACM MobiCom '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 2323–2328.

[8] F. Kaltenberger, G. d. Souza, R. Knopp, and H. Wang, "The openairinterface 5g new radio implementation: Current status and roadmap," in *WSA 2019; 23rd International ITG Workshop on Smart Antennas*, 2019, pp. 1–5. [Online]. Available: {https://ieeexplore.ieee.org/document/8727207}

[9] H. Cheng, S. D'Oro, R. Gangula, S. Velumani, D. Villa, L. Bonati, M. Polese, T. Melodia, G. Arrobo, and C. Maciocco, "Oranslice: An open source 5g network slicing platform for o-ran," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, ser. ACM MobiCom '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 2297–2302.

[10] 3GPP, "5G; Management and orchestration; 5G Networ Resource Model (NRM);," 3GPP, Tech. Rep. TS 28.541 V16.8.0 (2021-04)), 2021.

[11] 3GPP, "5g; nr; physical channels and modulation (release 18)," 3GPP, Tech. Rep. TS 38.211 V18.6.0 (2025-04), 2025.

[12] ——, "5g; nr; physical layer procedures for data (release 18)," 3GPP, Tech. Rep. TS 38.214 V18.6.0 (2025-04), 2025.

[13] M. Towers, A. Kwiatkowski, J. K. Terry, J. U. Balis, G. D. Cola, T. Deleu, M. Goulão, A. Kallinteris, M. Krimmel, A. KG, R. Perez-Vicente, A. Pierré, S. Schulhoff, J. J. Tai, H. Tan, and O. G. Younis, "Gymnasium: A standard interface for reinforcement learning environments," *CoRR*, vol. abs/2407.17032, 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Riesgos de seguridad en Open RAN: Vulnerabilidades, deficiencias y mitigación

Johnny Choque, Alberto García, Luis Diez, Ramón Agüero
Departamento Ingeniería de Comunicaciones,
Universidad de Cantabria
Edificio Ingeniería de Telecomunicación, Plaza de la Ciencia, 39005 Santander.
{jchoque, agarcia, ldiez, ramon}@tlmat.unican.es

La desagregación de la red de acceso radio ofrece beneficios, tales como mayor interoperabilidad, mejor coordinación entre nodos y reducción de costes, pero también plantea nuevos desafíos de seguridad en interfaces abiertas y mecanismos de control distribuidos. Este trabajo se centra en analizar vulnerabilidades de uno de los protocolos utilizados en dichas interfaces, SCTP, mediante dos ataques: uno de denegación de servicio y otro de suplantación de identidad. Ambos se han implementado y validado experimentalmente sobre la plataforma FlexRIC, en el marco de la arquitectura Open RAN, en contraste con enfoques puramente teóricos predominantes en la literatura. A partir de estos experimentos, se realiza una evaluación crítica de diversas soluciones de seguridad propuestas para SCTP. Aunque algunas ofrecen mecanismos de mitigación parcial, presentan limitaciones estructurales, operacionales y de compatibilidad que afectan su eficacia. Esto resalta la necesidad de adoptar protocolos de transporte más modernos y eficientes, capaces de integrar de manera nativa funciones de seguridad adecuadas para entornos Open RAN.

***Palabras Clave*—Open RAN, seguridad, SCTP, FlexRIC**

I. INTRODUCCIÓN

El avance hacia las redes móviles de quinta generación (5G) ha impulsado la adopción de arquitecturas abiertas como Open RAN (*Radio Access Network*) que promueven la interoperabilidad y la innovación mediante la desagregación funcional y el uso de interfaces estandarizadas. Dentro de este ecosistema, el *near Real-Time RAN Intelligent Controller* (near-RT RIC) desempeña un papel central en la gestión inteligente de los nodos de acceso, comunicándose con ellos a través de la interfaz E2. Esta interfaz se sustenta sobre el protocolo de transporte *Stream Control Transmission Protocol* (SCTP), diseñado para garantizar comunicaciones fiables en contextos de alta exigencia. El protocolo SCTP, si bien aporta ventajas significativas, presenta debilidades reconocidas en materia de seguridad, como se recoge en el RFC 5062 [4] y otros estudios posteriores. Estas vulnerabilidades son especialmente preocupantes cuando SCTP se implementa sin mecanismos de protección adicionales, como ocurre en diversas soluciones abiertas actualmente en uso.

La motivación principal del presente trabajo radica así en la identificación y análisis experimental de vulnerabilidades de seguridad del protocolo SCTP en implementaciones reales de arquitecturas Open RAN, con especial énfasis en la plataforma FlexRIC. Frente al enfoque predominantemente teórico de muchos estudios previos, este trabajo adopta una perspectiva completamente experimental, orientada a la validación práctica de los riesgos de seguridad en entornos funcionales. Esta aproximación permite detectar no sólo las limitaciones intrínsecas del protocolo SCTP, sino también deficiencias específicas en su implementación dentro de *frameworks* abiertos como FlexRIC, contribuyendo así al diseño de propuestas de mejora más realistas y aplicables. Para garantizar la replicabilidad de los experimentos presentados en este estudio, todo el código relevante, algunas mediciones y la descripción de las herramientas utilizadas se han puesto a disposición en un repositorio de acceso abierto[1].

El resto del artículo se organiza como sigue. En la Sección 2 se presenta el estado del arte sobre la seguridad de SCTP en contextos Open RAN. La Sección 3 describe los fundamentos tecnológicos necesarios para la comprensión del entorno de pruebas, incluyendo la arquitectura de FlexRIC y la interfaz E2. En la Sección 4 se analiza el funcionamiento detallado de SCTP en FlexRIC, sentando las bases para la identificación de sus vulnerabilidades. La Sección 5 expone la metodología seguida para la ejecución de los ataques, mientras que la Sección 6 describe la plataforma experimental utilizada. La Sección 7 detalla la implementación de los ataques, acompañada de análisis pormenorizados de su impacto. En la Sección 8 se evalúan diferentes soluciones de seguridad

[1]https://doi.org/10.5281/zenodo.15847555

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

propuestas para proteger SCTP, y finalmente, en la Sección 9, se presentan las conclusiones del estudio, destacando tanto los hallazgos como las recomendaciones para futuras mejoras.

II. ESTADO DEL ARTE

El protocolo SCTP ha sido objeto de análisis en términos de seguridad desde hace más de una década. Uno de los primeros documentos en este ámbito es el RFC 5062, publicado en 2007, donde se identifican múltiples vectores de ataque y se proponen contramedidas basadas en configuraciones seguras y un diseño robusto del protocolo. Estos ataques continúan siendo relevantes, como lo demuestra el estudio más reciente en [7], en el cual se demuestra experimentalmente que varias implementaciones actuales que usan SCTP siguen siendo vulnerables a alguno de los ataques descritos originalmente en el RFC 5062.

En los últimos años, se ha intensificado el interés en la seguridad del protocolo SCTP dentro del contexto de redes abiertas como Open RAN. No obstante, gran parte de la literatura reciente se centra en un análisis teórico. Por ejemplo, en [3] se presentan las mejores prácticas de seguridad en Open RAN, destacando sus beneficios de forma conceptual, sin incluir validaciones prácticas. De manera similar, en [5] se lleva a cabo un estudio formal para sintetizar posibles ataques sobre SCTP, utilizando herramientas de análisis simbólico, pero sin realizar pruebas experimentales.

Algunos trabajos han intentado cerrar esta brecha entre teoría y práctica. En [2], se realiza un modelado de amenazas de las interfaces Open RAN, proponiendo escenarios de ataque y llevando a cabo experimentos sobre la plataforma O-RAN Software Community (OSC). Sin embargo, este último trabajo carece de información técnica detallada sobre la implementación concreta de los ataques, lo que limita su utilidad como referencia práctica replicable.

Existen pocos estudios que presenten ataques experimentales sobre SCTP en entornos Open RAN con un enfoque técnico profundo. Entre las contribuciones más destacadas se encuentran [6] y [8]. En [6], los autores exploran un ataque de denegación de servicio mediante la explotación de una vulnerabilidad en el proceso de *E2 subscription request* del protocolo E2AP, implementado sobre FlexRIC y OSC. Por otro lado, en [8], se describe una implementación práctica de un ataque de *ARP spoofing* sobre la interfaz E2, utilizando srsRAN y FlexRIC como plataforma de pruebas. Ambos estudios constituyen ejemplos representativos de trabajos experimentales, pero no abordan específicamente los vectores de ataque sobre SCTP investigados en el presente trabajo.

III. FUNDAMENTOS TECNOLÓGICOS

A. Arquitectura Open RAN

Open RAN representa un paradigma de desagregación de las funciones de la red de acceso radio tradicional, estableciendo interfaces abiertas y estandarizadas que permiten la interoperabilidad entre componentes de diferentes proveedores. Esta arquitectura se fundamenta en la separación funcional entre el plano de control, el plano de datos y el plano de gestión, facilitando la implementación de soluciones multi-fabricante e impulsando la competencia en el ecosistema de telecomunicaciones.

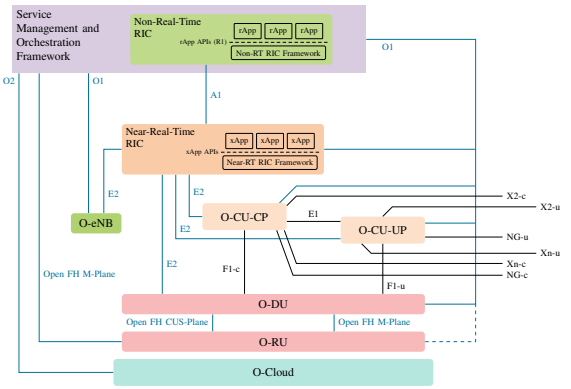


Fig. 1. Arquitectura Open RAN.

Dentro de la arquitectura Open RAN mostrada en la Fig. 1, el near-RT RIC, que opera con latencias entre 10 ms a 1 segundo, constituye el núcleo de procesado para aplicaciones de control y optimización de red. El near-RT RIC alberga las xApps (*eXtensible Applications*), que son aplicaciones específicas diseñadas para ejecutar algoritmos de control, optimización y análisis de rendimiento de la red. Estas aplicaciones pueden implementar funcionalidades como gestión de movilidad, control de interferencias, optimización de recursos radio y balanceado de carga, entre otras.

Los nodos E2, que típicamente corresponden a estaciones base (gNB) en redes 5G, representan los elementos controlados dentro de la arquitectura Open RAN, tales como O-CU, O-DU o O-eNB de la Fig. 1. Estos nodos implementan las funciones de protocolo de acceso radio y procesamiento de banda base, mientras mantienen interfaces estandarizadas que permiten su gestión y control por parte del near-RT RIC. Estos nodos recopilan y reportan información de estado y métricas de rendimiento al near-RT RIC, al tiempo que ejecutan las decisiones de control que este les transmite mediante mensajes de control específicos.

B. Interfaz E2 y protocolo E2AP

La interfaz E2 constituye el elemento de conectividad crítico entre el near-RT RIC y los nodos E2, definiendo los procedimientos de intercambio de información y control. Esta interfaz opera sobre el protocolo *E2 Application Protocol* (E2AP), que establece los mecanismos de comunicación asíncrona y garantiza la entrega confiable de mensajes entre las entidades de la arquitectura.

Cada procedimiento elemental definido por el protocolo E2AP está diseñado para soportar casos de uso específicos, desde la configuración inicial de la conexión hasta la ejecución de políticas de control en tiempo real. El procedimiento *E2 Setup* representa el mecanismo fundamental de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

establecimiento de conexión y negociación de capacidades entre el nodo E2 y el near-RT RIC. Durante este proceso, el nodo E2 informa al RIC sobre sus funcionalidades soportadas, los modelos de servicio disponibles, y los parámetros de configuración específicos. El procedimiento *E2 Setup* reviste especial importancia en este estudio, ya que en el transcurso de dicho proceso el nodo E2 transmite al near-RT RIC su información de identificación.

C. FlexRIC como implementación de Open RAN

FlexRIC [1] emerge como una implementación de código abierto de la arquitectura near-RT RIC, diseñada específicamente para facilitar la investigación, desarrollo y validación de conceptos en el ámbito de Open RAN. Esta plataforma proporciona un entorno flexible y extensible que permite la experimentación con algoritmos de control inteligente y la evaluación de su impacto en el rendimiento de la red.

A diferencia de otras plataformas RIC similares, FlexRIC minimiza la sobrecarga al adoptar un diseño orientado a eventos, evita funciones innecesarias y permite personalizar controladores mediante aplicaciones internas (iApps) y modelos de servicio.

La elección de FlexRIC como plataforma experimental se fundamenta en su capacidad para operar tanto con equipos de radio reales como con simuladores y emuladores, proporcionando flexibilidad en términos de configuración experimental. Esta versatilidad resulta especialmente valiosa para entornos de investigación, donde la disponibilidad de equipos comerciales puede ser limitada, permitiendo la validación de conceptos antes de su implementación en infraestructuras de producción.

IV. ANÁLISIS DEL PROTOCOLO SCTP EN FLEXRIC

La comunicación entre los componentes fundamentales de la arquitectura FlexRIC se sustenta en el protocolo SCTP, el cual proporciona las funcionalidades de transporte necesarias para el intercambio confiable de mensajes de control entre el near-RT RIC y los nodos E2. Esta sección presenta un análisis detallado del flujo de comunicación establecido entre estos componentes, examinando los mecanismos específicos del protocolo SCTP que facilitan tanto el establecimiento como el mantenimiento de las asociaciones de red, así como la señalización de la interfaz E2 involucrada durante estos procesos.

La Fig. 2 muestra el intercambio de mensajes entre dos componentes clave del sistema FlexRIC: el near-RT RIC, que actúa como servidor SCTP, y el E2 Node (gNB), que desempeña el rol de cliente. La comunicación entre ambos se establece utilizando exclusivamente el protocolo SCTP.

1) Establecimiento de la conexión SCTP: El proceso comienza con la iniciación de una conexión SCTP entre el RIC y el gNB, el cual sigue el mecanismo de *four-way handshake*. En primer lugar, el gNB inicia la conexión enviando un mensaje INIT al near-RT RIC. Este mensaje especifica parámetros como los flujos soportados y un número de verificación inicial. El RIC responde con un INIT ACK que incluye un *State Cookie* generado localmente para protegerse contra ataques de tipo *SYN flooding*. Posteriormente, el gNB responde con un COOKIE ECHO, que devuelve el *cookie* como prueba de que recibió correctamente el mensaje anterior y está dispuesto a establecer la asociación. Finalmente, el near-RT RIC envía un COOKIE ACK, confirmando la validez del *cookie* y completando así el establecimiento de la asociación SCTP desde el punto de vista del cliente, quien puede entonces comenzar el envío de mensajes DATA.

2) Configuración de la interfaz E2: De manera eficiente, el mensaje COOKIE ECHO enviado por el gNB incluye también un *chunk* DATA que encapsula el mensaje *E2 Setup Request* del protocolo E2AP. Aunque el RFC 9260 prohíbe el envío de datos antes de recibir el COOKIE ACK, se permite que un *chunk* DATA acompañe al COOKIE ECHO, siempre y cuando dicho dato no sea procesado por el servidor hasta después de validar correctamente el *cookie*. Esta práctica permite reducir la latencia en la fase de establecimiento, optimizando así la inicialización de la interfaz E2. En este contexto, el mensaje *E2 Setup Request* permite al gNB comunicar su identidad (PLMN ID, Node ID) y las funciones E2 que soporta, lo que habilita al RIC para construir su modelo de control sobre el nodo. Posteriormente, una vez que la asociación ha sido confirmada con el COOKIE ACK, el near-RT RIC responde con un *chunk* DATA que contiene el mensaje *E2 Setup Response*, donde se confirma la aceptación (total o parcial) de las funciones RAN propuestas y se formaliza la relación funcional entre el controlador y el nodo. Durante este intercambio, los mensajes SACK (*Selective Acknowledgment*) juegan un papel crucial, ya que permiten a cada extremo confirmar de manera precisa la recepción de los *chunks* DATA.

3) Monitoreo de la conexión: Durante el tiempo de vida de la asociación, ambos extremos —tanto el gNB como el RIC— participan activamente en el mantenimiento de la conexión mediante el uso de mensajes HEARTBEAT.

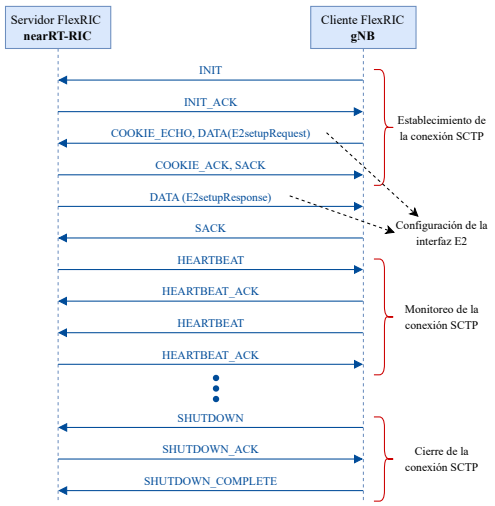


Fig. 2. Procesos SCTP y mensajes E2AP.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Cada extremo envía periódicamente este tipo de paquetes para verificar la accesibilidad de las rutas IP disponibles, a los cuales la contraparte responde con HEARTBEAT ACK. Esta verificación mutua es esencial, ya que SCTP puede operar en modo *multihoming*, lo que permite redirigir el tráfico a rutas alternativas en caso de fallos, garantizando así la continuidad del servicio.

4) Cierre de la conexión: Finalmente, el proceso concluye con la terminación ordenada de la asociación SCTP, iniciada por el cliente (gNB) mediante el envío del mensaje SHUTDOWN. Esta acción indica que el nodo cliente ha finalizado la transmisión de datos y solicita el cierre controlado de la sesión. El near-RT RIC responde con SHUTDOWN ACK, asegurando que todos los datos han sido recibidos, y el cliente completa el procedimiento con el mensaje SHUTDOWN COMPLETE. Este cierre ordenado evita la pérdida de información y asegura que todos los mensajes pendientes hayan sido entregados antes de liberar los recursos de red.

En conjunto, el intercambio observado refleja el funcionamiento coordinado entre los planos de transporte (SCTP) y de aplicación (E2AP), mostrando cómo FlexRIC utiliza mecanismos estandarizados para lograr una comunicación fiable, eficiente y segura entre el controlador y los nodos de acceso en redes móviles avanzadas.

V. METODOLOGÍA DE LOS ATAQUES

La metodología empleada en este trabajo se fundamenta en un modelo de amenaza específico que establece las condiciones y capacidades del atacante. El análisis de seguridad se realiza bajo el supuesto de que el adversario ha logrado previamente obtener acceso a la red donde operan los componentes críticos de FlexRIC, específicamente el near-RT RIC y el nodo E2. Esta premisa, aunque restrictiva, refleja un escenario realista en el contexto de ciberseguridad donde los atacantes pueden comprometer la seguridad perimetral mediante diversas técnicas de penetración.

Se asume que el atacante ha conseguido establecer una posición privilegiada dentro de la infraestructura de red, posiblemente mediante la intrusión en un nodo existente a través de técnicas como *reverse shell*, explotación de vulnerabilidades de software, o ingeniería social. Esta capacidad de acceso interno permite al atacante realizar actividades de monitorización pasiva del tráfico de red, así como la generación y transmisión de paquetes maliciosos dirigidos específicamente a los componentes de FlexRIC.

Los ataques desarrollados en este estudio explotan fundamentalmente las vulnerabilidades inherentes al protocolo SCTP, así como las deficiencias identificadas en la implementación específica de FlexRIC. El protocolo SCTP, aunque robusto en términos de confiabilidad y gestión de conexiones, presenta ciertas características que pueden ser aprovechadas maliciosamente cuando no se implementan medidas de seguridad adecuadas.

Para la implementación práctica de los ataques, se han desarrollado *scripts* especializados utilizando el lenguaje de programación *Python* en combinación con la biblioteca *Scapy*. Esta elección tecnológica se justifica por la flexibilidad que ofrece *Scapy* para la construcción y manipulación de paquetes de red a bajo nivel, permitiendo el control preciso sobre los campos de cabecera de los protocolos involucrados, incluyendo SCTP y E2AP. Los *scripts* implementados tienen la capacidad de interceptar, analizar y generar paquetes SCTP maliciosos, replicando los parámetros necesarios para suplantar identidades y manipular el estado de las asociaciones entre los componentes de FlexRIC.

La metodología adopta un enfoque sistemático que abarca desde la fase de reconocimiento y captura de tráfico legítimo, hasta la generación de ataques específicos dirigidos a interrumpir o comprometer la comunicación entre el near-RT RIC y los nodos E2. Esta aproximación permite no solo demostrar la viabilidad de los ataques, sino también cuantificar su impacto sobre la disponibilidad e integridad del sistema FlexRIC.

A. *Plataforma experimental*

Los experimentos relacionados con los ataques de seguridad fueron llevados a cabo utilizando una plataforma de pruebas diseñada específicamente para este propósito, tal como se ilustra en la Fig. 3. Dicha plataforma está compuesta por tres máquinas virtuales (VM) desplegadas sobre el hipervisor *VirtualBox*, todas interconectadas mediante una red de tipo *Host-only*.

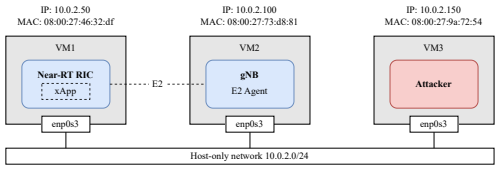


Fig. 3. Plataforma experimental.

Cada máquina virtual ejecuta un sistema operativo basado en Ubuntu Linux. La VM1 alberga el componente near-RT RIC, que actúa como el nodo servidor dentro del marco FlexRIC. Aunque también se encuentra desplegado un módulo xApp en esta VM, dicho componente no fue utilizado activamente durante la ejecución de los experimentos presentados en este trabajo.

En la VM2 se instancia el cliente de FlexRIC mediante un agente E2 que simula las funcionalidades de una gNB. La comunicación entre VM1 y VM2 se establece a través de sus respectivas interfaces de red virtuales *enp0s3*, las cuales permiten la implementación de la interfaz E2, responsable del intercambio de mensajes entre el near-RT RIC y el nodo gNB emulado.

Finalmente, la VM3 representa el nodo atacante dentro del entorno de pruebas. En este nodo se alojan los *scripts* desarrollados en *Python* que llevan a cabo distintos tipos de ataques de seguridad dirigidos al sistema. Este componente tiene como objetivo simular comportamientos maliciosos en el plano de control del sistema, con el fin de evaluar la robustez y posibles vulnerabilidades de la arquitectura FlexRIC frente a ataques externos.

La disposición estructural de la plataforma, así como la asignación de funciones específicas a cada máquina

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

virtual, permite una evaluación sistemática y reproducible de amenazas potenciales en entornos de redes 5G basadas en controladores RIC.

VI. IMPLEMENTACIÓN DE ATAQUES

Los ataques implementados explotan características específicas del protocolo SCTP y deficiencias en la implementación de FlexRIC para demostrar cómo un adversario con acceso a la red puede disrumpir o manipular las comunicaciones entre el near-RT RIC y los nodos E2. En este contexto, en esta sección se describe la implementación experimental de dos categorías principales de ataques dirigidos contra el protocolo SCTP en el entorno FlexRIC.

A. Ataque de denegación de servicio

Un ataque DoS (*Denial of Service*, o Denegación de Servicio) es un tipo de ciberataque cuyo objetivo es interrumpir la disponibilidad de un sistema, servicio o red, impidiendo que los usuarios legítimos puedan acceder a él. En este estudio se utiliza un mecanismo específico (el mensaje ABORT de SCTP) para llevar a cabo el ataque DoS.

1) Mensaje ABORT: El mensaje ABORT en el protocolo SCTP tiene como función principal terminar de manera abrupta una asociación SCTP existente, liberando todos los recursos asociados sin seguir el procedimiento de cierre ordenado (como el que ocurre con los mensajes SHUTDOWN). Este mensaje se envía para notificar al otro extremo que la asociación debe finalizar de inmediato debido a una condición de error o una situación que impide continuar la comunicación.

Uno de los parámetros principales dentro de la estructura básica del *chunk* ABORT es el *Bit T* y su relación con el *Verification Tag* de la cabecera SCTP. Cuando el *Bit T* se establece en 1, el mensaje ABORT se interpreta como respuesta a un *chunk* recibido previamente, y el *Verification Tag* debe coincidir con el del mensaje que originó la respuesta, asegurando su correcta vinculación con el contexto de error. En cambio, si el *Bit T* está en 0, el ABORT se genera de forma autónoma, y el *Verification Tag* debe corresponder al valor esperado por el receptor, identificando al emisor dentro de la asociación. Esta interacción asegura que el receptor pueda validar el mensaje sin ambigüedades, manteniendo la integridad y trazabilidad del protocolo incluso ante eventos críticos.

2) Ataque mediante el mensaje ABORT: La Fig. 4 ilustra un ataque ejecutado por un nodo malicioso mediante el uso del mensaje ABORT. Este nodo monitoriza la red de manera activa, capturando paquetes con el objetivo de interceptar el mensaje HEARTBEAT transmitido por el gNB. Una vez identificado el paquete deseado, el nodo malicioso extrae los parámetros críticos de la cabecera SCTP, incluyendo la dirección IP origen y destino, el puerto origen y destino, así como el *Verification Tag*. Con esta información, el nodo malicioso construye un paquete ABORT, configurando el *Bit T* en 0 para señalar una terminación unilateral de la asociación. Este paquete es enviado al near-RT RIC, simulando ser el gNB al emplear la IP, el puerto y el *Verification Tag* extraídos de este último. Al recibir el paquete ABORT, el near-RT RIC interpreta que el gNB ha solicitado el cese de la asociación SCTP en el lado del servidor. Paralelamente, el gNB, ajeno a la suplantación, continúa enviando paquetes HEARTBEAT al near-RT RIC. Al detectar la recepción de un HEARTBEAT, tras haber procesado un ABORT presuntamente originado por el gNB, el near-RT RIC identifica una anomalía. En respuesta, envía un paquete ABORT al gNB, estableciendo el *Bit T* en 1 e incluyendo en la cabecera SCTP el *Verification Tag* del mensaje que provocó la inconsistencia. Finalmente, el gNB, al recibir este nuevo ABORT, procede a cerrar la asociación en el lado del cliente.

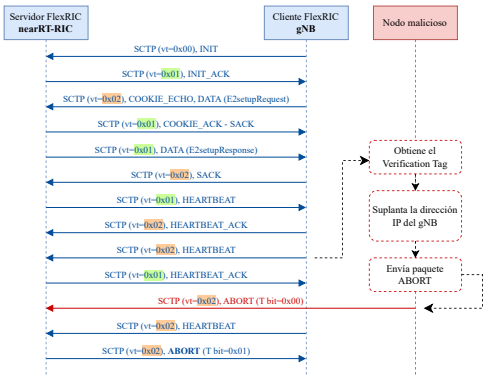


Fig. 4. Ataque DoS mediante mensaje ABORT.

3) Impacto del ataque DoS en FlexRIC: Para el caso de FlexRIC, el intercambio de paquetes ABORT entre el cliente y el servidor resulta únicamente en el cierre de la asociación a nivel de la capa SCTP. No obstante, la configuración de la interfaz E2, establecida previamente mediante los paquetes *E2 Setup Request* y *E2 Setup Response*, permanece almacenada en los registros del near-RT RIC. En caso de que el gNB intente restablecer una nueva conexión SCTP, el near-RT RIC, al procesar un nuevo *E2 Setup Request*, detectaría que el identificador de nodo E2 coincide con uno ya registrado. Como consecuencia, el near-RT RIC rechaza la conexión enviando un paquete SHUTDOWN, lo que impide que el gNB restablezca la comunicación con el controlador.

Adicionalmente, en la versión de FlexRIC empleada para este análisis (v2.0.0-87), el intento de reconexión por parte del gNB provoca una terminación abrupta del near-RT RIC debido a un error crítico (*Aborted*). Este comportamiento sugiere la presencia de una deficiencia en la implementación del near-RT RIC, potencialmente relacionada con una gestión inadecuada de los registros antiguos, lo que podría comprometer su estabilidad operativa.

El ataque mediante el mensaje ABORT puede considerarse una forma específica de ataque DoS. En este caso, el nodo malicioso utiliza paquetes ABORT de SCTP para destruir la asociación entre el near-RT RIC y el gNB, interrumpiendo su comunicación. Esto no solo afecta la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

conexión SCTP, sino que también impide reconexiones posteriores debido a errores en la gestión de registros del near-RT RIC, causando una denegación de servicio prolongada al evitar que el gNB se comunique con el controlador. Además, la terminación abrupta del near-RT RIC (*Aborted*) agrava el impacto, dejando potencialmente inoperables otros servicios que dependan de este componente.

B. Ataque de suplantación de identidad

Un ataque de suplantación de identidad (*Identity Spoofing*) es un tipo de ciberataque en el que un adversario se hace pasar por una entidad legítima del sistema, adoptando fraudulentamente sus credenciales y parámetros de identificación para establecer comunicaciones no autorizadas.

En este apartado se implementa un ataque de suplantación de identidad que utiliza el mensaje SHUTDOWN como vector inicial para desconectar al nodo legítimo, creando posteriormente una oportunidad para que el atacante establezca una nueva asociación suplantando la identidad del nodo E2 desconectado.

1) Mensaje SHUTDOWN: El cierre ordenado de una asociación SCTP constituye un procedimiento esencial para asegurar la entrega completa y confiable de los datos antes de finalizar la conexión. Conforme al RFC 9260, este proceso se ejecuta mediante una secuencia de tres mensajes: SHUTDOWN, SHUTDOWN ACK y SHUTDOWN COMPLETE. El mensaje SHUTDOWN, enviado por el extremo que inicia el cierre, incluye un campo *Cumulative TSN Ack* que confirma la recepción de todos los datos con TSN (*Transmission Sequence Number*) iguales o inferiores al especificado, permitiendo la liberación de recursos por parte del receptor. Posteriormente, el mensaje SHUTDOWN ACK confirma que el receptor ha procesado correctamente el cierre y entregado todos los datos pendientes. Finalmente, el mensaje SHUTDOWN COMPLETE certifica la finalización completa del proceso, tras lo cual ambos extremos liberan los recursos de la asociación. Este procedimiento incluye mecanismos de retransmisión para garantizar su fiabilidad frente a pérdidas, y ofrece ventajas sustanciales sobre el cierre abrupto con ABORT, al preservar la integridad de los datos y permitir una liberación controlada de recursos, características críticas en entornos que exigen alta confiabilidad.

2) Ataque mediante el mensaje SHUTDOWN: La Fig. 5 presenta el procedimiento llevado a cabo por un nodo malicioso para ejecutar un ataque utilizando el mensaje SHUTDOWN de SCTP. Inicialmente, el nodo malicioso realiza un monitoreo activo de la red con el propósito de capturar un paquete SCTP que contenga un *chunk* SACK. De dicho paquete, extrae el valor del *Cumulative TSN Ack*, un parámetro esencial para la construcción del mensaje malicioso. Posteriormente, el nodo malicioso continúa supervisando la red para interceptar un paquete HEARTBEAT enviado por el gNB. De manera análoga al procedimiento empleado en el ataque de DoS descrito en la sección anterior, el nodo malicioso extrae de este paquete los parámetros críticos de la cabecera SCTP, incluyendo las direcciones IP de origen y destino, los puertos de origen y destino, y el *Verification Tag*.

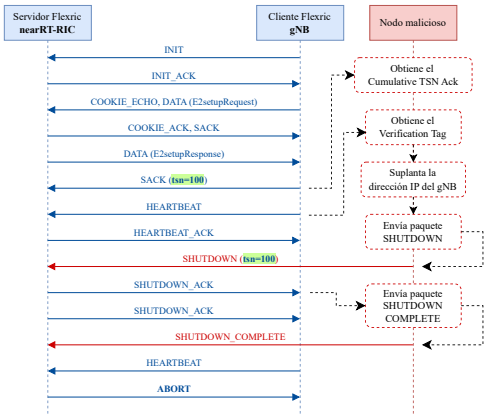


Fig. 5. Ataque a SCTP mediante mensaje SHUTDOWN.

Con esta información, el nodo malicioso construye un paquete SCTP que incorpora un *chunk* SHUTDOWN, utilizando el valor del *Cumulative TSN Ack* previamente obtenido. En este paquete, el nodo malicioso no emplea su propia dirección IP como origen, sino que adopta la dirección IP del gNB, suplantando su identidad. De esta forma, el near-RT RIC interpreta que el paquete SHUTDOWN proviene legítimamente del gNB. Al recibir el mensaje SHUTDOWN, el near-RT RIC responde enviando un paquete SHUTDOWN ACK al gNB. Sin embargo, el gNB descarta este paquete, ya que, según su estado interno, no espera un SHUTDOWN ACK debido a que no ha iniciado un proceso de cierre de la conexión SCTP.

Dado que el gNB no responde al SHUTDOWN ACK, el near-RT RIC, asumiendo que el paquete se ha perdido, reenvía el SHUTDOWN ACK. Mientras tanto, el nodo malicioso, que ha continuado monitorizando la red, detecta el paquete SHUTDOWN ACK enviado por el near-RT RIC. En respuesta, genera un paquete SHUTDOWN COMPLETE, nuevamente utilizando la dirección IP del gNB como origen, perpetuando la suplantación. Puesto que el gNB no tiene conocimiento de que la conexión SCTP ha sido cerrada por el nodo malicioso, prosigue enviando paquetes HEARTBEAT como parte de su mecanismo estándar para mantener activa la conexión. Al recibir uno de estos paquetes HEARTBEAT, el near-RT RIC identifica una anomalía en el proceso de comunicación con el gNB, ya que considera que la conexión ya había sido cerrada. En consecuencia, el near-RT RIC envía un paquete ABORT al gNB, finalizando abruptamente la conexión SCTP y completando el impacto del ataque.

3) Impacto del ataque SHUTDOWN en FlexRIC: A diferencia de lo observado en el ataque DoS, el cierre de la conexión SCTP entre el near-RT RIC y el gNB mediante el mensaje SHUTDOWN no afecta de manera significativa el funcionamiento de ambos componentes. El mensaje SHUTDOWN induce al near-RT RIC a eliminar el identificador del gNB de sus registros internos, permitiendo así que el gNB pueda intentar restablecer la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

conexión posteriormente. No obstante, al igual que en el ataque DoS, el mensaje ABORT recibido finalmente por el gNB solo tiene efecto a nivel de la capa SCTP. A nivel de aplicación, el gNB no detecta que la conexión ha sido cerrada, lo que sugiere una posible deficiencia en la implementación del nodo E2 en FlexRIC. Esta limitación podría estar relacionada con una falta de notificación adecuada a la capa superior, es decir, a la aplicación (gNB), sobre el estado de la conexión. Dicha notificación, que podría manifestarse como un error o un evento específico, depende de la implementación de la API de SCTP utilizada por la aplicación y su ausencia indica un posible aspecto de mejora de la implementación.

4) Ejecución del ataque de suplantación: El ataque basado en el mensaje SHUTDOWN ha generado tres consecuencias significativas. En primer lugar, se ha producido la eliminación de la asociación SCTP establecida entre el near-RT RIC y el gNB. En segundo lugar, el near-RT RIC ha eliminado al gNB de sus registros internos, liberando los identificadores asociados a la interfaz E2. Por último, el gNB permanece ajeno a estos eventos, lo que resulta en su desconexión inadvertida del near-RT RIC.

Estas circunstancias son aprovechadas por el nodo malicioso para ejecutar un ataque de suplantación. Con este propósito, el nodo malicioso realiza un monitoreo continuo de la red desde su inicio, con el objetivo de interceptar un paquete SCTP que contenga el mensaje COOKIE ECHO. Este paquete incorpora un *chunk* DATA que incluye un mensaje *E2 Setup Request* del protocolo E2AP. El propósito principal de dicho mensaje es iniciar la configuración de una conexión entre un nodo E2, como un gNB, y un near-RT RIC, permitiendo al nodo E2 notificar su presencia y capacidades al controlador, así como establecer una conexión destinada a la gestión y control de la red dentro del marco de la arquitectura Open RAN. Este mensaje incluye parámetros clave para identificar al nodo E2, entre los cuales destacan el *Global gNB ID*, que lo define de forma única mediante la combinación de un identificador de red móvil pública (PLMN ID) y un identificador específico del nodo (gNB ID). Estos parámetros garantizan que el near-RT RIC pueda reconocer y administrar de manera única al nodo E2.

El ataque de suplantación se inicia tras la ejecución del ataque mediante el mensaje SHUTDOWN descrito en la sección anterior. Aprovechando la desconexión total entre el near-RT RIC y el gNB, el nodo malicioso establece una nueva conexión SCTP con el near-RT RIC mediante el envío de un paquete INIT inmediatamente después del ataque DoS. Dado que se trata de una nueva asociación, no resulta imprescindible realizar una suplantación (*spoofing*) de las direcciones IP, a diferencia de lo requerido en el ataque basado en el mensaje ABORT. Tras recibir el INIT ACK, el nodo malicioso envía un duplicado del mensaje *E2 Setup Request* previamente capturado, el cual contiene los parámetros de conexión asociados al gNB relacionados con la interfaz E2. De esta manera, el nodo malicioso logra suplantar la identidad del gNB para esta nueva asociación. Los paquetes subsiguientes, como SACK y HEARTBEAT, forman parte de los procesos habituales de gestión de la conexión SCTP. Sin embargo, tanto a nivel de la asociación SCTP como de la configuración de la interfaz E2, el nodo malicioso opera efectivamente como si fuera el gNB legítimo.

VII. EVALUACIÓN DE SOLUCIONES

A pesar de sus beneficios, el protocolo SCTP no incorpora, por defecto, mecanismos de seguridad robustos tales como la confidencialidad, la autenticación y la integridad de los datos, elementos esenciales para garantizar una protección efectiva frente a los ataques descritos en las secciones precedentes. Con el objetivo de subsanar estas limitaciones, a continuación se analizan diferentes opciones.

A. TLS sobre SCTP

TLS proporciona autenticación, cifrado e integridad a nivel de aplicación sobre asociaciones SCTP. Sin embargo, su uso presenta limitaciones importantes: el *handshake TLS* ocurre tras establecer la asociación SCTP, dejando una ventana temporal desprotegida vulnerable a ataques DoS y suplantación. Aunque el RFC 3436 define TLS 1.0 sobre SCTP, esta versión está obsoleta; TLS 1.3 es preferible, aunque su implementación puede requerir tratar SCTP como un único flujo, sacrificando algunas de sus ventajas. Una vez establecida la sesión, TLS ofrece alta eficacia mediante validación criptográfica continua, protegiendo incluso interfaces como E2. No obstante, exige infraestructura PKI y conlleva un costo computacional significativo, especialmente en entornos con muchas conexiones. Además, su implementación sigue siendo experimental, con soporte limitado en bibliotecas criptográficas como OpenSSL.

B. SCTP-AUTH

SCTP-AUTH extiende SCTP proporcionando autenticación e integridad mediante HMAC para *chunks* específicos. Se activa más rápidamente que TLS, al requerir solo parámetros en la negociación inicial. Sin embargo, su autenticación es selectiva, dependiendo de una configuración adecuada, y no ofrece cifrado, exponiendo los datos a ser interceptados. Puede autenticar mensajes ABORT y SHUTDOWN, previniendo suplantaciones y ataques DoS, y protege de forma granular los aspectos críticos de la asociación. A nivel operativo, enfrenta retos en la distribución y gestión de claves simétricas sin un canal seguro, afectando la escalabilidad. En entornos con muchas asociaciones, el cálculo de HMAC puede ser un cuello de botella, aumentando la latencia y facilitando ataques DoS criptográficos.

C. DTLS sobre SCTP

DTLS adapta TLS a protocolos de datagrama y puede aplicarse sobre SCTP. No obstante, solo protege los datos de aplicación, dejando expuestos los mensajes de control SCTP, como ABORT y SHUTDOWN, a ataques. Además, la coexistencia de mecanismos de confiabilidad en DTLS y SCTP introduce complejidad técnica. En escenarios reales,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

su efectividad es limitada, ya que los atacantes pueden eludir la protección DTLS atacando directamente SCTP. Operacionalmente, la sobrecarga criptográfica de DTLS se suma a la de SCTP, y la gestión de ambas capas dificulta la configuración y resolución de problemas. Además, el estándar RFC 6083 solo contempla TLS 1.2, impidiendo la adopción de DTLS 1.3 y sus mejoras.

D. SCTP sobre IPsec

IPsec actúa a nivel de red, proporcionando autenticación, integridad y cifrado para todo el tráfico, incluido SCTP. Su implementación requiere una infraestructura sólida, y presenta dificultades en redes con NAT o topologías complejas. Además, sus principios arquitectónicos entran en conflicto con características avanzadas de SCTP, como la *multihoming*, lo que complica su configuración y puede degradar el rendimiento [9]. A pesar de ello, IPsec ofrece una protección robusta contra ataques DoS y suplantación, asegurando criptográficamente todos los mensajes SCTP. A nivel operacional, su principal ventaja es la transparencia para las capas superiores, permitiendo una integración sin modificaciones en las aplicaciones, alineada con políticas de seguridad ya existentes y sin disrupciones en procesos operacionales.

E. QUIC como alternativa a SCTP sobre IPSec

A pesar de los diversos mecanismos de seguridad evaluados para proteger las asociaciones SCTP, todos presentan limitaciones significativas en términos de cobertura, eficacia frente a ataques o viabilidad operativa. Entre estas opciones, IPsec se posiciona como la alternativa más robusta y completa, al ofrecer una protección integral del tráfico SCTP sin requerir modificaciones en las capas superiores del sistema. Sin embargo, su implementación se ve obstaculizada por incompatibilidades arquitectónicas, complejidad de configuración y degradaciones de rendimiento. En este contexto, el protocolo QUIC se identifica como una alternativa prometedora, al integrar de forma nativa seguridad y transporte confiable, superando las limitaciones de soluciones híbridas como IPsec sobre SCTP. Su simplicidad operativa, eficiencia y madurez lo posicionan como un candidato viable para reemplazar SCTP en la interfaz E2 [10].

VIII. CONCLUSIONES

Este estudio ha evidenciado, mediante validación práctica y análisis técnico, la existencia de vulnerabilidades críticas en el protocolo SCTP al ser empleado en arquitecturas Open RAN, particularmente en la implementación de FlexRIC. Se demostraron dos ataques efectivos —denegación de servicio mediante mensajes ABORT y suplantación de identidad a través de mensajes SHUTDOWN— que comprometen la disponibilidad y confiabilidad de la interfaz E2, aprovechando la ausencia de autenticación e integridad en los mensajes SCTP.

Respecto a las medidas de seguridad evaluadas, ninguna alternativa proporciona una protección integral sin incurrir en compromisos importantes. TLS protege solo después del establecimiento de sesión; SCTP-AUTH enfrenta dificultades de gestión de claves y escalabilidad; DTLS no protege mensajes de control; e IPsec, aunque robusto, presenta problemas técnicos, arquitectónicos y de rendimiento al implementarse sobre SCTP. En este contexto, el protocolo QUIC surge como una solución prometedora que no solo supera las deficiencias específicas de IPsec, sino que también aborda las limitaciones estructurales y operacionales presentes en todas las alternativas previamente consideradas.

Los resultados obtenidos subrayan la necesidad de fortalecer la seguridad en las interfaces de control de redes móviles abiertas y ofrecen una base empírica para futuras investigaciones orientadas a soluciones más seguras, escalables y eficientes en entornos 5G y 6G.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Gobierno de Cantabria a través del proyecto “Tecnologías habilitadoras de gemelos digitales y su aplicación a los sectores químico y de comunicaciones” (GDQuiC) del programa TCNIC (2023/TCN/002), así como por la Comisión Europea, European Cybersecurity Competence Centre (ECCC), a través del proyecto “5G-TACTIC: 5G Trusted and Secure Network Services” (DIGITAL-ECCC-2022-CYBER-03-101127973).

REFERENCIAS

[1] R. Schmidt et al., ”FlexRIC: an SDK for next-generation SD-RANs.” In Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies (CoNEXT ’21). Association for Computing Machinery, New York, NY, USA, 411–425. 2021. doi: 10.1145/3485983.3494870.

[2] P. Baguer et al., ”Attacking O-RAN Interfaces: Threat Modeling, Analysis and Practical Experimentation,” in IEEE Open Journal of the Communications Society, vol. 5, pp. 4559-4577, 2024, doi: 10.1109/OJCOMS.2024.3431681.

[3] M. Liyanage et al., “Open RAN security: Challenges and opportunities,” in Journal of Network and Computer Applications, Volume 214, 2023, 103621, ISSN 1084-8045, doi: 10.1016/j.jnca.2023.103621.

[4] R. Stewart et al., ”Security Attacks Found Against the Stream Control Transmission Protocol (SCTP) and Current Countermeasures”, RFC 5062, September 2007, doi 10.17487/RFC5062.

[5] J. Ginesin et al., “A Formal Analysis of SCTP: Attack Synthesis and Patch Verification,” in 33rd USENIX Security Symposium (USENIX Security 24, 2024, isbn 978-1-939133-44-1.

[6] F. Klement et al., “Endless Subscriptions: Open RAN is Open to RIC E2 Subscription Denial of Service Attacks” in IEEE 10th European Symposium on Security and Privacy (EuroS&P), 2025.

[7] L. Chianese et al., ”Leveraging Threat Modelling for Effective Penetration Testing in 5G Systems,” 2024 IEEE International Conference on Cyber Security and Resilience (CSR), London, United Kingdom, 2024, pp. 180-185, doi: 10.1109/CSR61664.2024.10679437.

[8] J. Kim et al., ”Simulation of an ARP Spoofing Attack on the E2 Interface in Open RAN,” 2024 IEEE 35th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, 2024, pp. 1-6, doi: 10.1109/PIMRC59610.2024.10817456.

[9] J. Groen et al., ”Securing O-RAN Open Interfaces,” in IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 11265-11277, Dec. 2024, doi: 10.1109/TMC.2024.3393430.

[10] V. -L. Nguyen et al., ”Security and Privacy for 6G: A Survey on Prospective Technologies and Challenges,” in IEEE Communications Surveys & Tutorials, vol. 23, no. 4, pp. 2384-2428, 2021, doi: 10.1109/COMST.2021.3108618.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Tráfico de sensado en la interfaz E2: integración de nodos radio 6G en O-RAN

Fátima Khan†, Luis Diez†, Jesús Gutiérrez¶, Ramón Agüero†
†*Universidad de Cantabria*, Santander, Spain. khanf@unican.es, {ldiez, ramon}@tlmat.unican.es
¶*IHP GmbH – Innovations for High Performance Microelectronics*, Germany. teran@ihp-microelectronics.com

En el proceso de evolución de las generaciones móviles, la integración real del sensado con las comunicaciones aparece como un pilar fundamental de la futura 6G, complementado por el esfuerzo de Open-RAN (O-RAN) en establecer un paradigma abierto. En este trabajo se presenta el desarrollo de una extensión de la interfaz E2 para soportar de manera natural el tráfico de sensado dentro de la arquitectura O-RAN. La propuesta se basa en la solución FlexRIC, un software abierto que implementa un near-Real Time RAN Intelligent Controller (near-RT RIC), de manera que un agente E2 podría desplegarse en cualquier elemento de red: nodos O-DU y O-CU, o en arquitecturas monolíticas de estaciones base (gNB o eNB). Además, es capaz de enviar métricas al near-RT RIC, las cuales pueden ser recolectadas posteriormente por una xApp. Finalmente, se propone habilitar una nueva interfaz en el agente E2 que permita la comunicación con la parte radio de una infraestructura desagregada sensing-Radio Unit (sRU), en la cual se recopilan datos de sensado del entorno.

***Palabras Clave*—Sensing, O-RAN, E2, xApp, near-RT RIC**

I. INTRODUCCIÓN

El ecosistema Open-RAN (O-RAN) da lugar a un paradigma abierto, inteligente, programable y virtualizado de las redes celulares. Con el fin de dotar a la red de inteligencia, se habilita el intercambio de Key Performance Indicators (KPI), haciendo uso de diferentes algoritmos de control para optimizar el rendimiento. Con este objetivo, y para garantizar la aplicabilidad de las soluciones que se propongan, es fundamental respetar las interfaces estandarizadas, permitiendo así la interoperabilidad entre diferentes componentes y fabricantes.

Uno de los componentes clave de la arquitectura O-RAN es el RAN Intelligent Controller (RIC), una funcionalidad basada en el paradigma *software-defined networking* que proporciona un punto de control centralizado para la Radio Access Network (RAN). Para habilitar la comunicación entre el RIC y la propia red de acceso, se define el protocolo E2 Application Protocol (E2AP), encargado de encapsular los mensajes del Service Model (SM) y establecer un mecanismo de comunicación estandarizado entre las eXternal App. (xApp) y la RAN Function (RF). Por su parte, la especificación E2 Service Model (E2SM) define los elementos comunes de información asociados a cada uno de los procedimientos contemplados en el protocolo E2AP [1], [2].

Además, la propia O-RAN define cuatro SMs: (i) RAN Function Network Interface (NI), que expone las interfaces de red; (ii) Key Performance Metrics (KPM), que monitoriza estadísticas de los nodos E2; (iii) RAN Control (RC), que define la interfaz para controlar los *Radio Bearers*, mapear los flujos de Quality of Service (QoS) a los Data Rate Bearers (DRBs), gestionar la asignación de recursos radio, el control de movilidad en modo conectado, el control de acceso radio, la conectividad dual y la agregación de portadoras, entre otras funcionalidades; y (iv) Cell Configuration and Control (CCC), que permite controlar y monitorizar la configuración de las celdas. Adicionalmente, se han definido otros SMs basados en las especificaciones del 3GPP, como los modelos para las capas MAC, RLC y PDCP [3]. Asimismo, en [4] se propone un SM personalizado para el control de tráfico Traffic Control (TC) compatible con O-RAN. En concreto, se introduce un *pipeline* de control de tráfico ubicado en el plano de usuario, entre SDAP y PDCP, compuesto por los bloques *classifier*, *policer*, *queue*, *scheduler*, *shaper* y *pacer*, definiéndose así un E2SM que encapsula dichos elementos de información.

Por otro lado, el IMT-2030, que establece los requisitos identificados para la próxima generación de comunicaciones inalámbricas, destaca la integración de las comunicaciones con capacidades de sensado de la red Integrated Sensing and Communication (ISAC), con el objetivo de dotarla de inteligencia basada en la información recolectada. Diversos trabajos han abordado la visión, las técnicas y las aplicaciones derivadas de la integración del

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

sensado y las comunicaciones en 6G. Estudios previos han analizado las posibles arquitecturas para incorporar ISAC, así como los KPI más relevantes, identificando los retos que plantea su implementación y proponiendo soluciones para afrontarlos, incluyendo tecnologías habilitadoras y aspectos de estandarización [5], [6], [7]. En [8] se discuten sus beneficios, desafíos técnicos, aplicaciones potenciales (como vehículos conectados, monitorización ambiental y hogares inteligentes), y el estado actual de su desarrollo y estandarización. Además, se hace especial hincapié en los compromisos y equilibrios necesarios entre diferentes objetivos o métricas al compartir recursos inalámbricos entre las funciones de sensado y comunicación. Por su parte, en [9] se resalta el papel de ISAC en la mejora de la eficiencia, la precisión en la localización y la percepción del entorno. También explora diversas técnicas, arquitecturas y aplicaciones potenciales, como la localización precisa y el establecimiento de representaciones (mapas) precisas del entorno.

Para lograr una integración efectiva, es necesario incorporar la información de sensado como un elemento clave dentro de la arquitectura de O-RAN, no solo para la toma de decisiones, sino también para la monitorización del entorno físico. Sin embargo, son pocos los trabajos que abordan esta convergencia. Así, en [10] se presenta ORAN-Sense, una arquitectura para la localización de transmisores no cooperativos mediante sensores Internet of Things (IoT) de espectro que están integrados en redes 5G a través de O-RAN. La propuesta emplea algoritmos Time Difference of Arrival (TDOA) con corrección de desfases de reloj y sincronización basada en señales de oportunidad. La solución se implementa como una xApps en el RIC. Resultados experimentales en entornos urbanos consiguen niveles de precisión del orden de decenas de metros.

A diferencia de estudios previos, el presente trabajo propone la extensión de los E2SMs existentes para incorporar de manera real tráfico de sensado, con el propósito de recopilar información del entorno físico. Esto permitirá la toma de decisiones, que podrían afectar tanto al propio SM como a otros, con el objetivo de optimizar el rendimiento de la red. Para ello, se han llevado acabo las siguientes acciones: (i) análisis de la arquitectura O-RAN para la integración de comunicaciones de sensado, (ii) diseño preliminar de un SM independiente del dispositivo de sensado, y (iii) validación de la implementación mediante una prueba de concepto en un escenario real.

El resto del documento se organiza en tres secciones. En la sección II, se describe la arquitectura O-RAN en lo relativo a la interfaz E2, ecosistema en el que se enmarca este trabajo, con especial énfasis en la integración de elementos de sensado dentro de la red. A continuación, en la sección III, se detallan los desarrollos realizados y se presenta un caso de uso en el que se monitoriza el entorno físico mediante una antena radar. Finalmente, en la sección IV, se presentan las conclusiones del trabajo junto con los siguientes pasos que se van a abordar.

II. ARQUITECTURA

En esta sección se introduce la arquitectura de O-RAN junto con la integración de elementos de sensado activos y pasivos. En primer lugar, como se observa en la Fig. 1, se considera una arquitectura multi-RAT, que contempla tanto 5G-RAN como conectividad *non-3GPP* (Sub-6, Wi-Fi, mmWave, radar). De esta manera se pretende obtener una representación precisa del entorno.

La información de sensado puede lograrse adoptando enfoques basados en Channel State Inofrmation (CSI) o radar pasivo. En ambos casos, la información de sensado se expone, a la RAN, a través de la interfaz E2 del RIC, que establece una conexión entre el Near-Real-Time RAN Intelligent Controller (near-RT RIC) y los nodos E2, como son Open-Distributed Unit (O-DU), Open-Central Unit (O-CU), así como con infraestructuras monolíticas, tales como gNB y eNB. La comunicación se realiza mediante mecanismos de publicación y suscripción para facilitar el intercambio de información entre las xApps y las funciones de la RAN. Adicionalmente, se propone extender la arquitectura para permitir la integración de nodos de sensado externos, representados por la unidad sensing-Radio Unit (sRU), que representa tecnologías *non-3GPP* enfocadas en tareas de percepción del entorno.

Para introducir inteligencia en la red, la arquitectura O-RAN incorpora dos controladores clave: el near-RT RIC, que opera en escalas de tiempo del orden de milisegundos a segundos, optimizando el comportamiento de la RAN mediante la interfaz E2, y el Non-Real-Time RAN Intelligent Controller (Non-RT RIC), que proporciona control de políticas a más largo plazo, optimización de recursos y análisis avanzados. La comunicación entre el Non-RT RIC y el near-RT RIC se realiza a través de la interfaz A1. Por su parte, el Non-RT RIC forma parte del marco funcional del Service Management and Orchestration (SMO), el cual coordina los recursos de la RAN mediante la interfaz O1, garantizando una gestión eficiente de la configuración y del rendimiento del sistema.

Las xApps que operan en el near-RT RIC, son responsables de recoger información de sensado, analizarla y almacenarla en bases de datos para su procesado en tiempo real. Estos datos se utilizan para optimizar parámetros clave de los diferente bloques funcionales que componen la RAN, mediante la interacción con los SMs definidos en el RIC, que permiten aplicar políticas y controlar de manera dinámica el comportamiento de la red. Además, la información de sensado puede ser transferida al SMO, que, al contar con una visión integral del sistema, es capaz de tomar decisiones de reconfiguración óptima de la red.

III. DESARROLLO

Este trabajo se ha centrado en extender las capacidades del RIC para permitir el intercambio de datos de *sensing*, independientemente del tipo de elemento conectado del que provenga dicha información. La interfaz E2 se instancia sobre el protocolo Stream Transmission Control Protocol (SCTP) y da soporte a dos servicios fundamentales: E2AP y E2SM. A través de un mecanismo de publicación

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

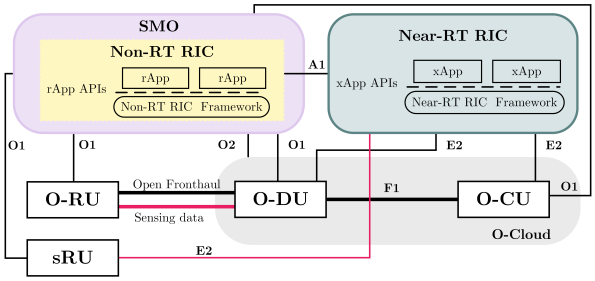


Fig. 1. Arquitectura O-RAN con integración de sensado.

y suscripción, los nodos E2 pueden publicar sus datos, mientras que las xApps desplegadas en el near-RT RIC pueden suscribirse a una o más funcionalidades de la RAN, como se ilustra en el diagrama que muestra la Fig. 2.

La interfaz E2 se estructura en términos de servicios, procedimientos e Information Elements (IE). Los servicios ofrecidos por el near-RT RIC comprenden múltiples procedimientos, cada uno de los cuales está compuesto por distintos IEs. Estos elementos de información determinan el tipo y los rangos de los datos intercambiados entre el near-RT RIC y los nodos E2.

Los mensajes definidos por el protocolo E2AP, también conocidos como *procedures*, encapsulan diferentes E2SMs. Estos mensajes permiten implementar funcionalidades asociadas tanto con el control como con la monitorización de métricas de la RAN, algunos de ellos reflejados en la Fig. 2. Dentro de estos procedimientos destacan los Elemenetary Procedures (EPs), distinguiéndose dos clases, aquellos con y sin respuesta por parte de la RAN. Asimismo, existen dos subcategorías principales de procedimientos: (i) *Global Procedures*, los cuales no están vinculados a una aplicación especifica, y (ii) *near-RT RIC Functional Procedures*, representando los mensajes específicos intercambiados entre el near-RT RIC y los nodos E2.

En la Fig. 2 se representan claramente ambos grupos. Dentro de los *Global Procedures*, destaca el procedimiento `E2 Setup`, con el objetivo de establecer la señalización entre el nodo E2 y near-RT RIC. Este procedimiento se inicia por el nodo E2 y transmite información relativa a su propia configuración, así como acerca de las funcionalidades RAN que soporta, las cuales se especifican mediante el *RAN Functions Added List IE*. Cada función RAN está definida a través de su correspondiente SM, e incluye, en caso de ser necesario, la notificación sobre la disponibilidad de datos de sensado. Si el near-RT RIC acepta la configuración propuesta, el campo `reject`, correspondiente a cada una de las funcionalidades, toma el valor 0 en el mensaje de respuesta. En cuanto a los *near-RT RIC Functional Procedures*, hay que tener en cuenta los que se enumeran: `Subscription Request` permite establecer suscripciones E2, siendo el iniciador de una secuencia de eventos y acciones; `Subscription Delete`, utilizado para eliminar suscripciones previamente establecidas en los nodos E2, y finalmente los procedimientos de `Indication`, cuyo propósito es transferir un mensaje asociado a los servicios *Report* y/o *Insert* desde el RIC hacia el near-RT RIC, como consecuencia de una suscripción exitosa y la posterior detección de un evento. Por su parte, los E2SMs especifican los distintos servicios disponibles en el RIC, entre los que se incluyen: *report*, *insert*, *control*, *policy* y *query*. Asimismo, definen los distintos IEs, que pueden ser comunes o específicos de cada SM.

Con el fin de extender las funcionalidades del sistema y habilitar la transmisión de datos de sensado, se ha diseñado y desarrollado un nuevo SM, respetando la arquitectura de comunicaciones establecida por la interfaz E2. Para ello, se ha utilizado la implementación de FlexRIC[1] de OAI, desarrollando el código correspondiente en lenguaje C.

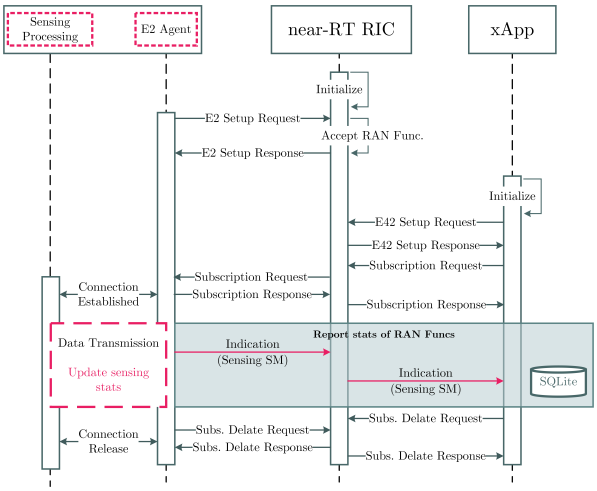


Fig. 2. Diagrama de secuencia de las operaciones en el near-RT RIC.

En particular, se ha desarrollado un nuevo SM orientado al sensado, el cual se integra como una librería en cada uno de los agentes involucrados: el near-RT RIC, los xApps y el agente E2. Una vez definidos los datos soportados por dicho SM, estructurados conforme a un modelo de *heatmap*, se ha adaptado la lógica interna de cada entidad para incorporar y dar soporte funcional a este nuevo modelo.

Seguidamente, en el propio nodo E2, se ha habilitado la posibilidad de recibir datos de otro servicio, concretamente de la unidad de procesado de datos de sensado, como se observa en las Figs. 2 y 3. Este procesado se realiza localmente y genera información de sensado ligera (en particular, mapas de calor), que es enviada al agente E2. De esta manera, se garantiza que, en caso de que la unidad de procesado de datos de sensado no esté alojada en el mismo dispositivo que el agente E2, no se sobrecargue la red. Finalmente, se ha habilitado la xApp con el fin de que se encargue de recolectar la información, a través de *callbacks* que se activan cuando se reciben datos mediante el procedimiento `Indication`, y de almacenarlos en una base de datos mediante el empleo de consultas `MySQL`.

Finalmente, como se puede observar en la Fig. 3, se

[1]https://gitlab.eurecom.fr/mosaic5g/flexric.git

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

ha desplegado la implementación en un entorno real, a modo de prueba de concepto, en un escenario de localización basado en radar. En concreto, el sistema de sensado se ha construido utilizando cuatro dispositivos USRP N321, además de un PC que se encarga, a su vez, de la transmisión, recepción y procesado de los datos. Una vez procesados, estos datos se envían al agente E2, encargado de comunicarse con la propia RAN. Los datos se transmiten a través de la interfaz E2 mediante el empleo del SM dedicado a los datos de sensado. Seguidamente, el near-RT RIC se encarga de orquestar la recolección de esta información en tiempo real a través de la xApp, la cual la almacena en una base de datos para su posterior análisis y visualización. Finalmente, mediante el empleo de Grafana, y realizando consultas a la base de datos, se puede visualizar el mapa de calor generado por la propia xApp, como se refleja en la Fig. 3.

El mapa de calor representa la intensidad de la señal recibida: cuanto más fuerte es la señal, más tiende el color hacia el rosa, para un cierto ángulo y distancia dados. El ángulo se representa en el borde de la circunferencia, mientras que el radio indica la distancia. La resolución angular representa el ángulo, y la del rango la distancia correspondiente, que determinan el nivel de detalle con el que se perciben los objetos. Estas resoluciones permitirían entender en qué medida los objetos "vistos" aparecen más grandes que sus dimensiones físicas reales, permitiendo estimar su localización.

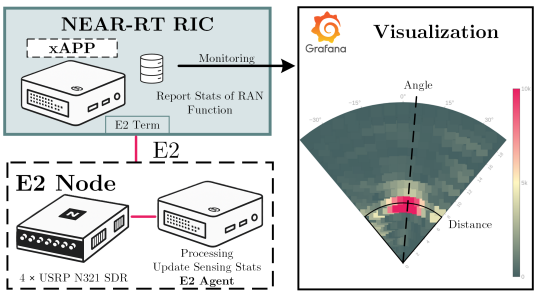


Fig. 3. Prueba de concepto integrando comunicaciones de sensado.

IV. CONCLUSIONES Y LÍNEAS FUTURAS

En este trabajo se han presentado las bases de una línea de investigación incipiente, que aúna la integración del sensado y comunicaciones en las futuras generaciones móviles (6G), dentro del marco de la arquitectura O-RAN, independientemente de si se emplean tecnologías 3GPP o non-3GPP. Con este fin, se ha mostrado cómo interactúa el elemento de sensado con la red desde el punto de vista de arquitectura, se han explorado las implicaciones de llevar la implementación correspondiente a un contexto real, y se ha desplegado un escenario a modo de prueba de concepto para su validación.

Como trabajo futuro, se propone extender la interfaz E2 para habilitar que el SM dedicado al sensado pueda recopilar nuevas estadísticas, definiendo así las métricas de interés y diseñando e implementando políticas de control que permitan actuar directamente sobre el elemento de sensado. Además, se plantea la elaboración de una xApp que, mediante una lógica específica, facilite la interconexión de las diferentes capas que conforman el segmento de red RAN a través de los diferentes SMs.

Por otro lado, el diseño del SM dedicado al sensado permitirá caracterizar los flujos de tráfico que podrían ser inyectados en la red, facilitando de esta manera acometer un estudio analítico más detallado del comportamiento de la red. En ese sentido se tiene que tener en cuenta que en dicho segmento de red (fronthaul) convergen tráficos provenientes de otras tecnologías o interfaces (open-fronthaul, backhaul). El objetivo final sería acometer el análisis de configuraciones óptimas según las condiciones de la red y realizar un modelado preciso del tráfico para un correcto estudio del sistema.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el programa Horizonte Europa Smart Networks and Services Joint Undertaking (SNS JU) bajo el acuerdo #101139282 (6G-SENSES). También ha recibido fondos del Ministerio de Asuntos Económicos y Transformación Digital de España mediante el proyecto SITED (PID2021-125725OB-I00) y del Gobierno Regional de Cantabria a través del programa TCNIC (2023/TCN/002).

REFERENCIAS

[1] O-RAN, "O-ran e2 application protocol (e2ap)," O-RAN Alliance, Alfter, Germany, Tech. Rep., 2025.

[2] ——, "O-ran e2 service model (e2sm)," O-RAN Alliance, Alfter, Germany, Tech. Rep., 2025.

[3] R. Schmidt, M. Irazabal, and N. Nikaein, "Flexric: an sdk for next-generation sd-rans," in *Proceedings of the 17th International Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 411–425. [Online]. Available: https://doi.org/10.1145/3485983.3494870

[4] M. Irazabal and N. Nikaein, "Tc-ran: A programmable traffic control service model for 5g/6g sd-ran," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 2, pp. 406–419, 2024.

[5] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6g wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[6] A. Kaushik, R. Singh, S. Dayarathna, R. Senanayake, M. Di Renzo, M. Dajer, H. Ji, Y. Kim, V. Sciancalepore, A. Zappone, and W. Shin, "Toward integrated sensing and communications for 6g: Key enabling technologies, standardization, and challenges," *IEEE Communications Standards Magazine*, vol. 8, no. 2, pp. 52–59, 2024.

[7] Y. Zhang, Y. Lin, Q. Bie, R. Zhang, and Y. He, "Toward integrated sensing and communication network: The evolutionary architecture for 5g-advanced and beyond," in *2024 IEEE 24th International Conference on Communication Technology (ICCT)*, 2024, pp. 856–862.

[8] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[9] N. González-Prelcic, M. Furkan Keskin, O. Kaltiokallio, M. Valkama, D. Dardari, X. Shen, Y. Shen, M. Bayraktar, and H. Wymeersch, "The integrated sensing and communication revolution for 6g: Vision, techniques, and applications," *Proceedings of the IEEE*, vol. 112, no. 7, pp. 676–723, 2024.

[10] Y. Lizarribar, R. Calvo-Palomino, A. Scalingi, G. Santaromita, G. Bovet, and D. Giustiniano, "Oran-sense: Localizing non-cooperative transmitters with spectrum sensing and 5g o-ran," in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications*, 2024, pp. 1870–1879.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Comparación de Algoritmos de Aprendizaje Automático para Predicción de Contaminación en la Ciudad de Cartagena

Emilio Ramos-Sorroche, Esteban Egea-Lopez, Jose Santa

Escuela Técnica Superior de Ingeniería de Telecomunicación
Universidad Politécnica de Cartagena
Plaza Cronista Isidro Valverde, 30202 Cartagena, Murcia
emilio.ramos@upct.es, esteban.egea@upct.es, jose.santa@upct.es

La contaminación atmosférica es una preocupación mundial con implicaciones sanitarias y ambientales. Este trabajo aborda la predicción de concentraciones de monóxido de carbono (CO) mediante análisis de series temporales, usando datos de sensores urbanos en Cartagena. Se prueban métodos estadísticos y de aprendizaje automático con las librerías Python *sktime* y *Optuna* para ajustar hiperparámetros. Se comparan modelos clásicos (ARIMA, ETS), regresiones k-NN y redes profundas (CNNRegressor) bajo ventanas de entrenamiento de una semana a varios meses. La precisión se evalúa con SMAPE, RMSE, MAE y R^2, y también se mide el tiempo de ejecución. Los resultados indican que modelos sencillos como ETS o ARIMA ofrecen un rendimiento sólido, mientras que k-NN destaca por su ejecución rápida con precisión razonable, equilibrando velocidad y exactitud. Estos hallazgos evidencian que algoritmos bien ajustados permiten obtener pronósticos de CO precisos e instantáneos, útiles para la gestión de la calidad del aire en ciudades.

***Palabras Clave*—Contaminación atmosférica, Predicción de CO, Series temporales, Aprendizaje automático, Modelos ARIMA/ETS, k-nearest neighbors**

I. INTRODUCCIÓN

La contaminación atmosférica sigue siendo uno de los retos medioambientales más acuciantes en todo el mundo, afecta a millones de personas y plantea riesgos sustanciales para la salud pública, los ecosistemas y el desarrollo económico. La rápida urbanización, la expansión industrial y el aumento del tráfico de vehículos contribuyen al deterioro de la calidad del aire en ciudades de todo el mundo. La vigilancia y la previsión de las concentraciones de contaminantes permiten a los responsables políticos y a las partes interesadas llevar a cabo intervenciones oportunas, proteger el bienestar de los ciudadanos y tomar decisiones basadas en datos sobre el control de las emisiones.

Los recientes avances en las tecnologías de sensores han dado lugar a densas redes de vigilancia que captan datos en tiempo real sobre contaminantes como el CO, NO_2, O_3, SO_2, VOCs, y CO_2. Esta enorme afluencia de información también plantea retos en términos de heterogeneidad, valores perdidos y la variabilidad inherente a las series temporales medioambientales. Aunque este estudio se centra en la ciudad de Cartagena (España), donde el turismo y las actividades industriales ponen de relieve la importancia del aire limpio, estas cuestiones son de interés universal en las áreas metropolitanas de todo el mundo.

Se han propuesto varios enfoques para la predicción de la contaminación atmosférica, desde modelos estadísticos tradicionales (p. ej., ARIMA, SARIMA) hasta métodos avanzados de aprendizaje automático o aprendizaje profundo (p. ej. LSTM, CNN). No obstante, sigue habiendo dudas sobre qué algoritmos pueden captar mejor las fluctuaciones a corto plazo de los contaminantes en condiciones urbanas heterogéneas, al tiempo que se equilibra el coste computacional. En particular, los datos de monóxido de carbono (CO) suelen mostrar una compleja interacción de fuentes de emisión locales, factores meteorológicos y patrones diurnos.

En este trabajo, proponemos una comparación sistemática de múltiples algoritmos de Machine Learning para predecir concentraciones de CO. Basándonos en la librería *sktime*, podemos implementar aproximaciones clásicas de series temporales (ARIMA, ETS, TBATS, etc.), métodos basados en regresión (*k-nearest neighbors* via *reduction*), y arquitecturas de redes neuronales (CNNRegressor) bajo un marco unificado. Además, *Optuna* se utiliza para realizar el ajuste automatizado de hiperparámetros, con el objetivo de maximizar la precisión y minimizar el tiempo de experimentación manual. Las principales **contribuciones** de este trabajo se resumen a continuación:

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- **Benchmark exhaustivo:** Evaluamos algoritmos de aprendizaje automático clásicos y modernos para la predicción de CO sobre múltiples conjuntos de datos de sensores, detallando las métricas de error y los tiempos de ejecución.
- **Optimización de hiperparámetros:** La optimización automática mediante *Optuna* se aplica de forma sistemática a cada algoritmo, ofreciendo información sobre las configuraciones óptimas y reduciendo la conjetura.
- **Impacto del tamaño de la ventana de entrenamiento:** Investigamos cómo la longitud de los datos históricos afecta a la precisión de la predicción, poniendo de manifiesto la relación entre coste computacional y rendimiento.
- **Conclusiones prácticas:** Nuestros resultados destacan qué algoritmos logran un mejor equilibrio entre velocidad y precisión para la predicción a corto plazo de la calidad del aire, orientando los despliegues reales en redes de sensores urbanas.

El resto de este artículo se organiza del siguiente modo. La Sección II revisa el *Estado del Arte* en la predicción de series temporales de contaminantes, identificando las lagunas de los enfoques existentes. La Sección III describe los conjuntos de datos y los pasos de preprocesado (incluida la imputación de valores perdidos). En la Sección IV, se detallan los algoritmos seleccionados, el procedimiento de ajuste de hiperparámetros y el diseño experimental global. Los resultados y su discusión se presentan en la Sección V, con énfasis en las métricas de precisión, las comparaciones de tiempos de ejecución y el efecto del tamaño de la ventana de entrenamiento. Finalmente, la Sección VI recoge las conclusiones y las líneas futuras de investigación en una discusión conjunta.

II. ESTADO DEL ARTE

La predicción de la contaminación atmosférica, y en particular del CO, ha sido un tema de investigación relevante debido al impacto directo de este contaminante en la salud pública. Los métodos estadísticos clásicos, como ARIMA y SARIMA, suelen emplearse como referencia por su interpretabilidad [1]. Sin embargo, pueden presentar dificultades ante las fuertes no linealidades y los patrones cambiantes propios de los entornos urbanos. Estudios más recientes adoptan enfoques de Aprendizaje Automático para manejar grandes conjuntos de datos de sensores [2], aunque muchos de ellos se centran en predicciones de una sola estación o no incluyen la optimización de hiperparámetros.

Las redes neuronales, especialmente las arquitecturas profundas, han despertado gran interés por su capacidad para modelar dependencias temporales complejas. Xie et al. [3] proponen un modelo CNN-LSTM que ofrece mejoras significativas en escenarios con alta variabilidad. De igual modo, Zhan et al. [4] investigan una LSTM con atención para la predicción de CO en zonas costeras, subrayando la integración de información espaciotemporal. A pesar de estos avances, los largos tiempos de entrenamiento y la necesidad de grandes bases de datos bien etiquetadas pueden limitar su aplicabilidad en entornos reales.

Más allá de los enfoques neuronales, los métodos basados en regresión—como los modelos de ventana deslizante o los algoritmos basados en árboles—han mostrado su eficacia al capturar fluctuaciones a corto plazo en la contaminación urbana [5, 6]. Estos suelen depender de una selección cuidadosa de variables de entrada (por ejemplo, factores meteorológicos o datos de tráfico) y pueden desplegarse con una menor carga computacional. No obstante, muchas soluciones no evalúan de forma sistemática cómo diferentes longitudes de datos históricos afectan al rendimiento, ni aprovechan de modo exhaustivo la optimización sistemática de hiperparámetros.

Herramientas como *Optuna* reducen el tiempo dedicado al ajuste de hiperparámetros, minimizando las fases de prueba y error y mejorando la reproducibilidad [7]. Combinadas con bibliotecas como *sktime* [8], facilitan la comparación amplia entre métodos clásicos (ARIMA, ETS) y arquitecturas más avanzadas en un único entorno. Sin embargo, pocos trabajos profundizan en los compromisos computacionales que resultan cruciales en contextos de tiempo real o cuasi tiempo real.

Además, ciertos estudios subrayan las complejidades espaciales y meteorológicas de la distribución del CO, abogando por enfoques multivariantes que tengan en cuenta otras interacciones de contaminantes y variables climáticas [9, 10]. Aunque estos métodos multivariantes pueden mejorar la precisión, también introducen mayor complejidad, lo que hace pertinente considerar si enfoques univariantes más sencillos (con un ajuste adecuado) son suficientes para predicciones fiables a corto plazo. El presente estudio aborda estas lagunas mediante la evaluación sistemática de un amplio conjunto de algoritmos sobre múltiples conjuntos de datos de sensores, cada uno con diferentes extensiones de datos históricos fiables, y otorgando prioridad tanto a la precisión como al tiempo de ejecución.

III. DATOS Y PREPROCESADO

A. Descripción de los datos

Nuestro conjunto de datos procede de 15 sensores medioambientales desplegados en el casco histórico de Cartagena, cada uno de los cuales registra múltiples contaminantes —incluidos CO, NO_2, O_3, SO_2, COV y CO_2. La Figura 1 muestra un mapa de esta zona urbana, destacando la ubicación de los sensores. El Ayuntamiento de Cartagena instaló originalmente una red más amplia de sensores para monitorizar la calidad del aire en todo el municipio. Para este estudio seleccionamos 15 dispositivos cuya distribución geográfica ofrece una visión representativa de las zonas urbanas clave (p. ej., calles comerciales, áreas residenciales y entornos industriales), garantizando al mismo tiempo un mínimo de datos ausentes en sus registros de CO. Este subconjunto pretende así capturar dinámicas de contaminación diversas dentro de un alcance manejable para experimentos comparativos de predicción.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Fig. 1: Mapa del centro histórico de Cartagena que muestra la ubicación de los sensores.

El monóxido de carbono (CO) se eligió no solo por su abundancia relativa y la continuidad de sus registros, sino también porque actúa como indicador crítico del tráfico local y de las emisiones industriales, factores especialmente relevantes en el entorno urbano de Cartagena. Según la calidad de los datos, el intervalo final seleccionado para cada sensor oscila entre 15 días y siete meses.

B. Pasos de preprocesado

El preprocesado de datos se diseñó para garantizar una frecuencia horaria coherente y huecos mínimos. Los pasos seguidos fueron:

1) *Selección de periodos por sensor*: identificar segmentos con mínimos valores perdidos y lecturas continuas.
2) *Remuestreo horario*: calcular promedios horarios de CO para unificar la tasa de muestreo.
3) *Imputación de valores perdidos*: los huecos cortos se rellenaron propagando hacia adelante la última observación válida, preservando las tendencias locales. Para ello se ha usado la función *pandas.DataFrame.ffill* del framework Pandas.
4) *Tratamiento de valores atípicos*: las mediciones claramente inválidas se eliminaron o sustituyeron, según el conocimiento del dominio. Dada la preselección de los sensores y las ventanas temporales con mejor continuidad y calidad de mediciones, se minimizó la cantidad de estos valores atípicos y, en caso de quedar alguno, esta eliminación y posterior sustitución con *pandas.DataFrame.ffill* se realizó de forma manual.
5) *Formateo para sktime*: las series limpias se almacenaron en *DataFrames* de *pandas*, indexadas por tiempo, para cumplir los requisitos de entrada de los experimentos posteriores.

Estos pasos garantizan la coherencia de los datos entre sensores, a pesar de las diferencias en la duración total. La Figura 2 muestra ejemplos típicos de las series horarias tras la limpieza.

IV. METODOLOGÍA

A. Algoritmos de Predicción Seleccionados

Investigamos un conjunto diverso de enfoques de predicción disponibles en *sktime*:

- **Estadísticos clásicos**: ARIMA/*autoARIMA*, ETS/*AutoETS*, Suavizado Exponencial, BATS, TBATS y *UnobservedComponents*.
- **Basados en regresión**: *make_reduction* con *k*-nearest neighbors (k-NN).
- **Neuronales e híbridos**: *CNNRegressor* con estrategia recursiva y *Prophet* (aunque *Prophet* es parcialmente model-based, suele tratarse por separado).

Si bien estos métodos están ampliamente descritos en la literatura, la novedad radica en aplicar una optimización sistemática de hiperparámetros a cada uno y, posteriormente, comparar tanto la exactitud como la eficiencia computacional. Además, examinamos cómo la variación de la ventana de entrenamiento (de una semana a varios meses) afecta al rendimiento de la predicción, lo que permite evaluar la adaptabilidad de cada enfoque bajo diferentes niveles de disponibilidad de datos.

B. Métricas de Evaluación

Para valorar la precisión de las predicciones y comparar los modelos utilizamos las siguientes métricas de error:

- **SMAPE** (Error Porcentual Absoluto Medio Simétrico): mide el error relativo medio, penalizando sobre- y sub-predicciones grandes y mitigando los problemas cuando valores reales o predichos se acercan a cero.
- **RMSE** (Raíz del Error Cuadrático Medio): refleja la raíz cuadrada de las diferencias cuadráticas medias entre predicción y observación, siendo especialmente sensible a los valores atípicos.
- **CVRMSE** (Coeficiente de Variación del RMSE): normaliza el RMSE por la media de los datos observados, facilitando la comparación entre conjuntos con distinta escala.
- **MAE** (Error Absoluto Medio): calcula el valor medio de los residuos sin tener en cuenta su signo, indicando cuánto se desvían, en promedio, las predicciones.
- R^2 (Coeficiente de Determinación): cuantifica la proporción de varianza explicada por el modelo, variando de $-\infty$ (ajuste muy deficiente) a 1 (ajuste perfecto).

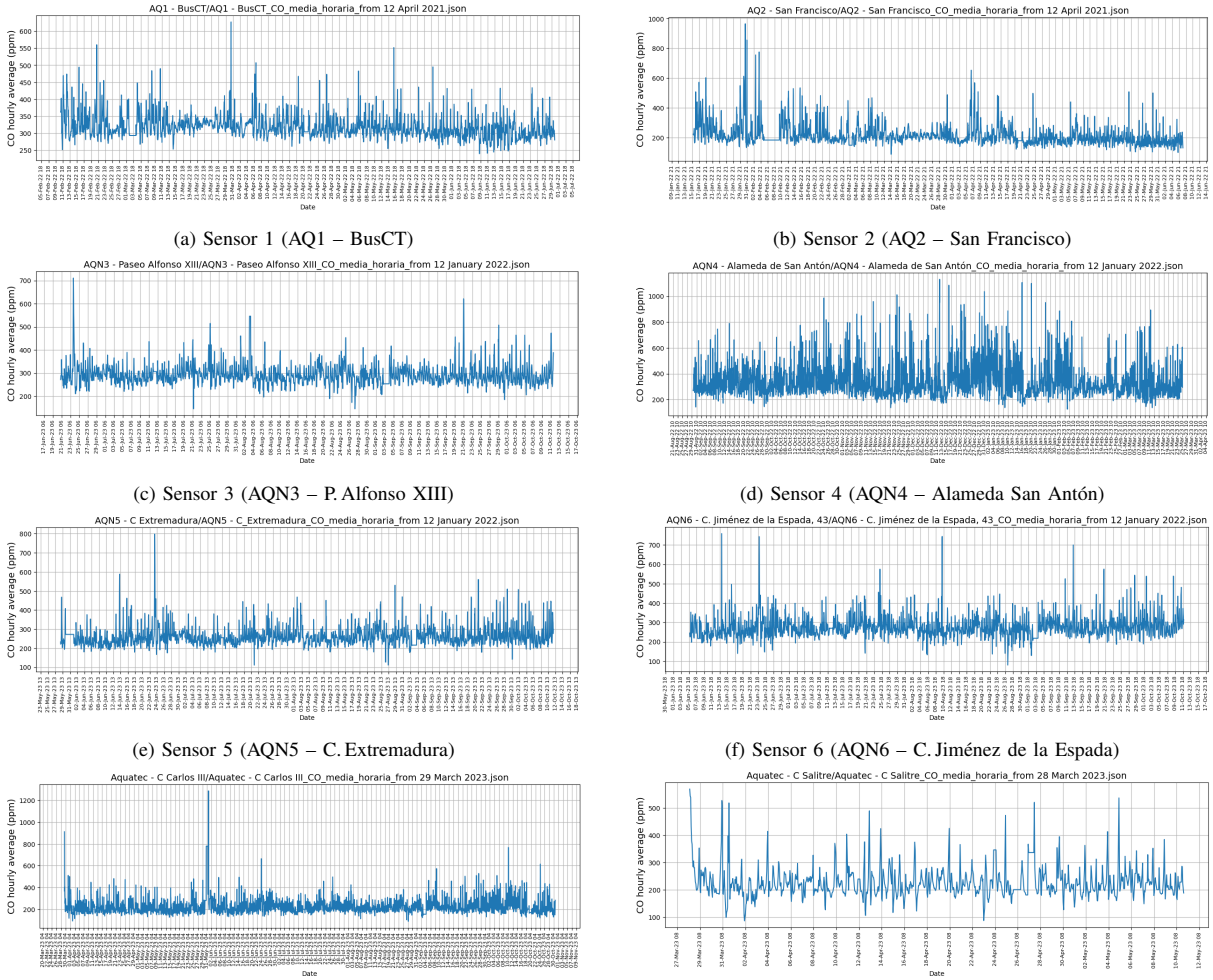
SMAPE guía la optimización con Optuna, mientras que RMSE, MAE y R^2 ofrecen perspectivas complementarias sobre la calidad de la predicción.

C. Optimización de Hiperparámetros con Optuna

Empleamos *Optuna* para automatizar el ajuste de hiperparámetros de manera coherente en todos los algoritmos. El procedimiento general es el siguiente:

1) **Función objetivo**: minimizar SMAPE sobre un conjunto de validación.
2) **Espacio de búsqueda**: definir rangos adecuados para cada modelo (p. ej., $\{p, d, q\}$ en ARIMA, número de vecinos en k-NN, detalles de arquitectura en *CNNRegressor*, etc.).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) Sensor 1 (AQ1 – BusCT)

(b) Sensor 2 (AQ2 – San Francisco)

(c) Sensor 3 (AQN3 – P. Alfonso XIII)

(d) Sensor 4 (AQN4 – Alameda San Antón)

(e) Sensor 5 (AQN5 – C. Extremadura)

(f) Sensor 6 (AQN6 – C. Jiménez de la Espada)

(g) Sensor 7 (Aquatec – C. Carlos III)

(h) Sensor 8 (Aquatec – C. Salitre)

Fig. 2: Serie horaria de CO de los sensores 1–8 tras el preprocesado.

3) **Ejecuciones** (***trials***): en cada ejecución *Optuna* muestrea una configuración y evalúa su rendimiento.
4) **Mejor configuración**: se conserva la combinación de hiperparámetros con menor SMAPE.

Optuna se configuró con varias mejoras: (i) detección de configuraciones duplicadas para evitar evaluaciones redundantes; (ii) número de ejecuciones adaptado a la complejidad del modelo (menos pero más costosas para *CNNRegressor*; más y ligeras en métodos estadísticos); y (iii) SMAPE como métrica principal, definida por:

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \tag{1}$$

Estas adaptaciones resultaron esenciales dada la variabilidad de la serie de CO y los distintos espacios de hiperparámetros.

D. Entrenamiento, Configuración de Predicción y Comparación de Modelos

Con los hiperparámetros óptimos fijados, cada serie se divide en entrenamiento y prueba, reservando las últimas 24 h para test. Se generan predicciones de 24 h que se comparan punto a punto con los valores reales. Los experimentos se automatizan mediante *scripts* que abarcan desde la carga de datos hasta el cálculo de errores y la generación de gráficos superpuestos (predicción vs. observación).

Para minimizar sesgos, repetimos el proceso en los 15 sensores y promediamos las métricas. Además, registramos el tiempo de ejecución de cada algoritmo, aspecto crítico para monitorización ambiental en tiempo (casi) real.

Finalmente, analizamos la influencia del tamaño de la ventana de entrenamiento en el Sensor 7 (calle Carlos III) con siete meses de datos. Se replicó la predicción de 24 h usando ventanas de seis meses, tres meses, un mes y una semana. Los resultados, presentados en la Sección V, muestran cómo cambia la precisión y la carga computacional al variar el histórico disponible.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

V. RESULTADOS Y DISCUSIÓN

A. Valores predichos frente a reales con CO (Sensor de ejemplo)

Con el fin de visualizar cómo se comporta cada algoritmo de predicción en un conjunto de datos concreto, elegimos el sensor situado en la calle Carlos III (Sensor 7), ya que ofrece una representación clara del comportamiento general de los métodos. La Fig. 3 muestra los últimos siete días de mediciones de CO, destacando las 24 h finales usadas como conjunto de prueba. Durante ese intervalo se comparan los valores previstos por cada método (naranja) con las observaciones reales (verde). Al analizar estas gráficas se aprecia hasta qué punto la trayectoria de cada modelo coincide o se desvía de las concentraciones reales de CO, aportando una visión complementaria a las métricas numéricas que se comentan más adelante.

Aunque la mayoría de los algoritmos capturan razonablemente bien las tendencias generales, existen diferencias notables en la rapidez con que responden a cambios bruscos. Algunas predicciones subestiman la altura de picos repentinos (ETS, BATS y k-NN), mientras que otras sobrecorrigen o muestran cierto desfase (Suavizado Exponencial, ETS y *UnobservedComponents*). En varios casos las curvas previstas resultan más suaves que la serie real (Prophet, StatsForecastAutoARIMA), lo que indica que esos modelos imponen una regularización más fuerte o suponen patrones estacionales más simples. Por el contrario, los métodos que reaccionan muy deprisa a los nuevos datos pueden sobreajustarse a anomalías breves y generar picos transitorios en sus predicciones (ARIMA).

Esta comparación visual ofrece una perspectiva práctica sobre los puntos fuertes y las limitaciones de cada técnica. Los métodos más suaves pueden ser preferibles para pronósticos generales a largo plazo, mientras que los más reactivos podrían ser útiles para alertas a corto plazo siempre que no generen excesivos falsos positivos. Las secciones siguientes presentan un análisis cuantitativo, agrupando los errores de múltiples sensores y discutiendo los compromisos entre precisión y requisitos computacionales.

B. Rendimiento global en todos los sensores

La Tabla I recoge el rendimiento medio de cada algoritmo sobre los 15 conjuntos de datos de sensores, usando las métricas descritas en la Sección B. Además de la precisión, registramos el tiempo necesario (en segundos) para entrenar y generar predicciones de 24 h en los 15 sensores. Estos tiempos son orientativos, pues dependen del hardware y del entorno de ejecución.

Incluimos también una *puntuación combinada* (**SCORE**) que penaliza tanto ejecuciones lentas como valores altos de SMAPE:

$$\text{SCORE} = \frac{1}{\text{Tiempo (s)}} \times \frac{1}{\text{SMAPE}} \qquad (2)$$

de modo que los algoritmos con SMAPE bajo y tiempos cortos obtienen valores mayores. La SCORE es ilustrativa (no un ranking definitivo) y ayuda a visualizar el compromiso entre precisión y eficiencia.

Tabla I:
Comparación de métricas de error, tiempos de ejecución y puntuación combinada.

Algoritmo	SMAPE	RMSE	CVR	MAE	R^2	T (s)	SCORE
k-NN (red.)	7.329	43.212	0.195	32.323	–0.681	**0.30**	**0.454**
Suav. Exp.	6.439	35.269	0.168	27.657	0.449	0.80	0.194
ARIMA	7.439	40.543	0.183	32.534	0.815	3.60	0.037
ETS	**6.084**	**34.714**	**0.159**	**27.315**	–0.230	5.80	0.028
SF-AutoARIMA	8.370	44.970	0.215	35.521	1.032	101	0.001
Unobs. Comp.	6.989	39.822	0.183	31.168	–0.579	722	0.0002
BATS	7.234	40.432	0.189	31.840	–0.733	1322	0.0001
Prophet	8.452	45.860	0.219	36.739	–1.169	1490	8e-5
AutoETS	7.770	43.987	0.198	34.555	0.768	3362	4e-5
TBATS	8.148	43.735	0.208	35.213	–1.046	6114	2e-5
CNN + Rec.	12.438	54.801	0.311	46.409	–2.730	7560	1e-5
autoARIMA	7.851	43.217	0.206	33.565	–0.779	60075	2e-6

Desde el punto de vista de la *precisión*, ETS obtiene el SMAPE más bajo (6.084) y un RMSE relativamente reducido (34.714), lo que muestra que sus predicciones reproducen con gran fidelidad los valores observados. Suavizado Exponencial también ofrece un SMAPE bajo (6.439) y requiere un tiempo mínimo de cómputo (0.80 s). No obstante, la mejor SCORE global corresponde al k-NN (`make_reduction`), que combina un tiempo ínfimo (0.30 s) con un SMAPE moderado (7.329).

Modelos como *autoARIMA*, AutoETS, TBATS, Prophet o CNNRegressor presentan SMAPE altos, tiempos de ejecución elevados, o ambas cosas, lo que se traduce en puntuaciones bajas. Por ejemplo, *autoARIMA* registra un SMAPE de 7.851 pero necesita unas 60 075 s de proceso, mientras que CNNRegressor muestra el peor SMAPE (12.438), el mayor RMSE (54.801) y un tiempo considerable (7 560 s).

En conjunto, ETS y Suavizado Exponencial destacan por su precisión, mientras que k-NN ofrece el mejor equilibrio precisión–velocidad. La elección final dependerá de si la prioridad es *minimizar el error* o *reducir el tiempo de cómputo*.

C. Comparación de distintas ventanas de entrenamiento

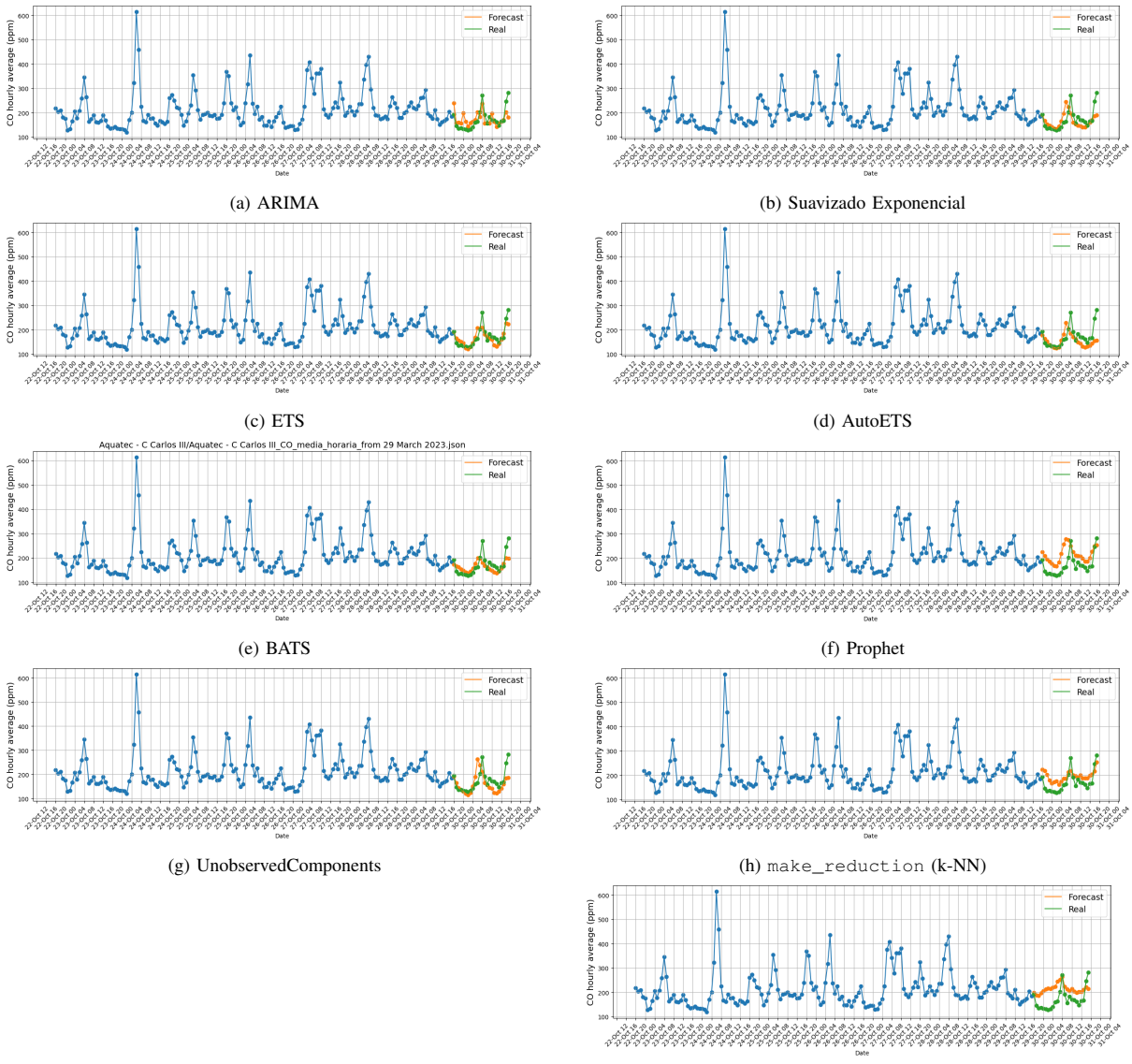
Para analizar cómo la cantidad de histórico influye en la precisión de las predicciones, realizamos un experimento adicional con el sensor 7 (calle Carlos III), que dispone de siete meses de mediciones horarias de CO. Cada algoritmo se entrenó con cuatro longitudes de ventana (seis meses, tres meses, un mes y una semana), manteniendo para prueba las mismas 24 h finales. Así observamos cómo varía el rendimiento cuando se dispone de menos información histórica.

La Fig. 4 ilustra el comportamiento de `make_reduction` (k-NN) bajo estas cuatro ventanas. Elegimos este algoritmo porque con ventanas cortas muestra pronósticos visualmente más reactivos. La subfigura (a) corresponde a seis meses de entrenamiento; las subfiguras (b), (c) y (d) usan respectivamente tres meses, un mes y una semana.

La Tabla II resume SMAPE, RMSE, MAE, R^2 y el *Tiempo (s)* de entrenamiento para cada algoritmo y ventana.

Nota: con una semana de histórico, **AutoETS** no generó predicciones válidas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) ARIMA

(b) Suavizado Exponencial

(c) ETS

(d) AutoETS

(e) BATS

(f) Prophet

(g) UnobservedComponents

(h) `make_reduction` (k-NN)

(i) StatsForecastAutoARIMA

Fig. 3: Últimos siete días de datos de CO del Sensor 7 (calle Carlos III) y predicciones de las 24 h finales para nueve algoritmos: ARIMA, Suavizado Exponencial, ETS, AutoETS, BATS, Prophet, UnobservedComponents, `make_reduction` (k-NN) y StatsForecastAutoARIMA.

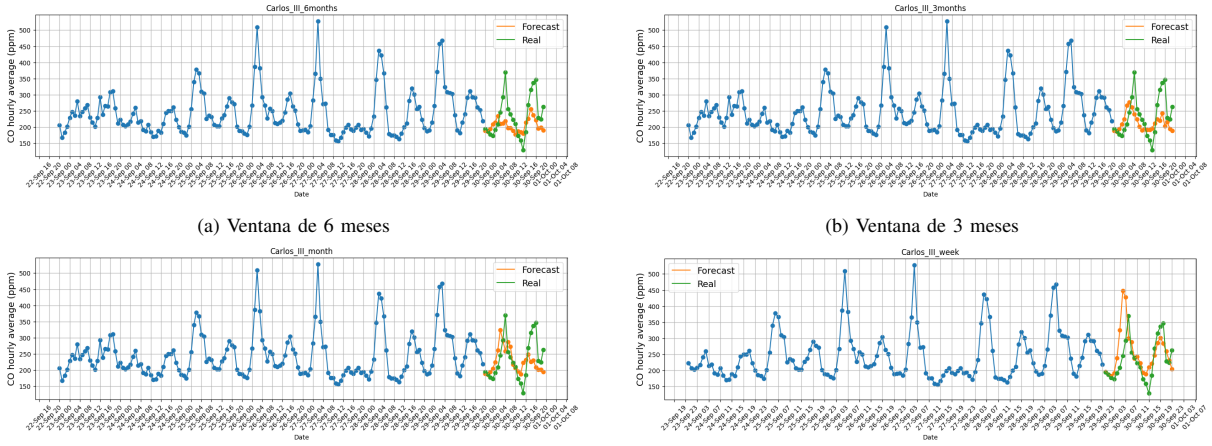
En algunos casos, reducir la ventana de seis a tres meses—o incluso a un mes—no empeora (e incluso mejora) la predicción a corto plazo. Es posible que la ventana corta se centre en patrones recientes más relevantes, mientras que una historia larga incluya componentes estacionales obsoletos. El tiempo de entrenamiento aumenta de forma apreciable en métodos como *autoARIMA* cuando la ventana crece, mientras que modelos más simples (p. ej., Suavizado Exponencial) se mantienen rápidos con cualquier tamaño; esto resulta crítico si el reentrenamiento debe ser frecuente en sistemas en tiempo real.

Algunos métodos, como Suavizado Exponencial o *autoARIMA*, se degradan notablemente con menos datos, mientras que otros—BATS o k-NN—toleran bien ventanas cortas e incluso mejoran ligeramente. En definitiva, la elección del modelo óptimo depende de la cantidad y estabilidad del histórico disponible: con datos escasos o recientes, conviene un método robusto a ventanas cortas; con historiales extensos y estables, modelos más complejos que aprovechen el contexto completo pueden ofrecer mejores predicciones. Para horizontes de varios días o semanas, las ventanas largas suelen producir pronósticos más "conservadores"; verificar este efecto en horizontes mayores queda como trabajo futuro.

D. Discusión global y consideraciones en tiempo real

Los resultados anteriores indican que métodos relativamente simples (p. ej., ETS, ARIMA o incluso Suavizado Exponencial) pueden ser firmes candidatos para la predicción de CO a corto plazo cuando se ajustan de forma

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(a) Ventana de 6 meses

(b) Ventana de 3 meses

(c) Ventana de 1 mes

(d) Ventana de 1 semana

Fig. 4: Valores reales vs. predichos para `make_reduction` (k-NN) con cuatro longitudes de ventana de entrenamiento. El resto de algoritmos se analizó de forma análoga.

Tabla II:
Métricas de error y tiempos según la ventana de entrenamiento (Sensor 7). El mejor valor por fila aparece en **negrita**.

Algoritmo	Vent.	SMAPE	RMSE	CVR	MAE	R^2	T (s)
ARIMA	6 m	7.07	39.15	0.168	32.17	0.599	0.20
	3 m	**7.03**	39.19	0.168	**32.09**	0.598	0.11
	1 m	7.17	39.11	**0.168**	32.35	0.600	0.06
	1 sem	7.48	**39.76**	0.171	32.94	**0.586**	**0.04**
autoARIMA	6 m	8.66	53.21	0.228	41.84	0.259	416
	3 m	10.61	58.04	0.249	51.09	0.118	532
	1 m	**8.64**	**53.12**	**0.228**	**41.99**	**0.261**	662
	1 sem	10.40	60.26	0.259	51.44	0.050	**25**
Exponential Smoothing	6 months	**9.484**	54.175	0.232	44.986	0.232	0.059
	3 months	9.954	58.903	0.253	48.047	0.092	0.030
	1 month	10.942	66.829	0.287	54.231	-0.169	**0.013**
	1 week	10.024	**73.281**	**0.314**	**52.333**	**-0.406**	0.025
ETS	6 months	7.556	43.512	0.187	34.409	0.504	0.679
	3 months	7.407	43.235	0.186	33.833	0.511	0.228
	1 month	**7.433**	**42.031**	**0.180**	**34.405**	**0.538**	0.084
	1 week	7.694	45.959	0.197	37.273	0.447	**0.037**
AutoETS	6 months	**7.031**	47.945	0.206	34.022	0.398	135.580
	3 months	7.696	**46.767**	**0.201**	35.110	**0.428**	37.870
	1 month	7.596	49.140	0.211	**36.463**	0.368	**7.420**
	1 week	-	-	-	-	-	-
BATS	6 months	7.783	44.982	0.193	35.894	0.470	30.920
	3 months	7.882	45.631	0.196	36.465	0.455	15.570
	1 month	**6.883**	**42.280**	**0.181**	**31.983**	**0.532**	**2.360**
	1 week	7.645	46.171	0.198	36.659	0.442	0.478
TBATS	6 months	8.151	46.416	0.199	37.439	0.436	61.970
	3 months	9.377	51.412	0.221	43.323	0.308	15.130
	1 month	8.441	46.525	0.200	38.970	0.433	12.640
	1 week	**7.438**	**45.843**	**0.197**	**35.867**	**0.450**	**1.770**
Prophet	6 months	8.480	51.409	0.221	40.372	0.308	1.590
	3 months	**8.348**	**50.418**	**0.216**	39.711	**0.335**	0.291
	1 month	9.034	52.704	0.226	**43.266**	0.273	**0.130**
	1 week	10.329	67.997	0.292	52.414	-0.210	0.145
StatsForecastAutoARIMA	6 months	8.620	50.418	0.216	40.879	0.335	14.510
	3 months	9.936	53.620	0.230	47.135	0.247	7.210
	1 month	8.985	50.587	0.217	42.564	0.330	3.900
	1 week	**9.510**	**69.055**	**0.296**	**49.415**	**-0.248**	**0.529**
UnobservedComponents	6 months	9.120	51.424	0.221	43.177	0.308	10.740
	3 months	9.114	52.097	0.224	43.564	0.290	6.440
	1 month	**10.014**	**56.323**	**0.242**	**48.308**	**0.170**	**1.840**
	1 week	9.884	66.996	0.287	50.475	-0.175	0.406
k-NN (make_reduction)	6 months	10.038	59.088	0.254	46.985	0.086	0.028
	3 months	9.703	58.587	0.251	45.801	0.102	0.020
	1 month	9.315	58.911	0.253	44.710	0.092	0.026
	1 week	**8.545**	**65.195**	**0.280**	**45.109**	**-0.113**	**0.007**

adecuada. En cambio, los enfoques neuronales o híbridos más avanzados suelen requerir volúmenes mayores de datos de entrenamiento—y más tiempo de cálculo—para superar el rendimiento de las técnicas clásicas, algo a menudo poco viable en escenarios dinámicos de tiempo real. Además, métodos como *k*-NN ofrecen una muy buena relación velocidad–precisión, lo que resulta valioso en sistemas de alerta inmediata donde es necesario reentrenar con frecuencia o inferir con rapidez.

En la práctica, la elección del algoritmo dependerá de:

- **Disponibilidad de datos**: si sólo se dispone de unas pocas semanas de datos fiables, los modelos más sencillos suelen ser más robustos.
- **Restricciones computacionales**: los sistemas en tiempo real se benefician de enfoques rápidos de entrenar y actualizar.
- **Requisitos de precisión**: ciertas aplicaciones sanitarias o normativas pueden exigir errores mínimos, lo que justifica ventanas de entrenamiento más largas y ajustes más complejos.

VI. CONCLUSIONES

En este trabajo comparamos de forma sistemática diversos enfoques de predicción de series temporales para estimar concentraciones de CO, aplicando una optimización exhaustiva de hiperparámetros. A diferencia de muchos estudios previos que recurren a ajustes manuales o parciales, nuestra canalización automatiza la selección de parámetros tanto en algoritmos clásicos como avanzados, reduciendo el ensayo–error y garantizando una estrategia de búsqueda uniforme entre modelos. Esta ventaja metodológica mejora la reproducibilidad y ofrece pautas prácticas para la predicción a corto plazo de CO en despliegues reales de sensores urbanos.

Los resultados muestran que los métodos clásicos, como ETS y ARIMA, siguen siendo muy competitivos: suelen igualar o superar a modelos más complejos con unos requisitos computacionales moderados. Por su parte, los métodos de reducción *k*-NN destacan por su velocidad de ejecución y precisión aceptable, cualidades especialmente valiosas en escenarios en tiempo (casi) real. Frente a técnicas más exigentes en datos—como las redes neuronales profundas—estos enfoques ofrecen un equilibrio favorable entre precisión y rapidez, requiriendo ventanas de entrenamiento más cortas e infraestructuras sencillas.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Además, al examinar distintos tamaños de ventana demostramos que la precisión puede mejorar con historiales más recientes, sugiriendo que datos actuales y de calidad pueden ser más útiles que registros extensos pero obsoletos.

Pese a estos resultados alentadores, nuestro estudio presenta algunas limitaciones. Primero, los experimentos se centraron en el CO como contaminante único (predicción univariante), lo que impide aprovechar posibles sinergias con variables meteorológicas u otros contaminantes (NO_2, O_3, etc.). Segundo, algunos sensores sólo aportaron series cortas (hasta 15 días), lo que afecta a algoritmos sensibles a tendencias de largo plazo. Tercero, el estudio se realizó en el contexto urbano de Cartagena, por lo que los modelos podrían requerir recalibración para generalizar en zonas con patrones o estacionalidades distintas. Por último, aunque incluimos varias técnicas clásicas, basadas en regresión y neuronales, quedan por explorar arquitecturas profundas más avanzadas (Transformers o Graph Neural Networks), que podrían mejorar la precisión a costa de mayor complejidad computacional.

De cara al futuro, sería útil incorporar más contaminantes y variables meteorológicas en un marco multivariante, clarificando las relaciones entre factores ambientales. Explorar arquitecturas profundas como LSTMs, Transformers o GNNs puede aportar mejoras cuando se disponga de conjuntos de datos amplios y de alta calidad. Asimismo, integrar sensores móviles o aprovechar flotas de vehículos urbanos ampliaría la cobertura espacial, permitiendo pronósticos más granulares y revelando puntos críticos de contaminación. Estas líneas de trabajo ayudarán a los responsables municipales a diseñar estrategias de intervención más eficaces, protegiendo la salud pública y fomentando la sostenibilidad ambiental.

AGRADECIMIENTOS

Este trabajo ha sido financiado por las ayudas *CNS2022-136150* y *PID2023-148104OB-C42*, otorgadas por el MICIU/AEI/10.13039/501100011033 y cofinanciadas por el FEDER/UE; por la ayuda FPU21/06991 concedida por el Ministerio de Universidades; por la ayuda *22228/PDC/23*, concedida por la Fundación Séneca – Agencia de Ciencia y Tecnología de la Región de Murcia; y por el proyecto *GREENTBEAT BG-RRP-2.005-0002*, financiado por el Plan Nacional de Recuperación y Resiliencia de Bulgaria y la Comisión Europea. Agradecimiento al Ayuntamiento de Cartagena por su colaboración en el acceso a los datos del estudio.

REFERENCIAS

[1] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018. [Online]. Available: https://otexts.com/fpp2/

[2] B. Yu, Y. Ma, and J. Wang, "Air quality forecasting using machine learning methods: A systematic review," *Applied Sciences*, vol. 10, no. 15, p. 5202, 2020.

[3] X. Xie, J. Zhou, J. Qiao, and L. Zhu, "A novel deep learning model for air quality forecasting based on cnn-lstm," *Science of The Total Environment*, vol. 718, p. 137310, 2020.

[4] S. Zhan, Y. Zhu, and B. Wang, "Spatiotemporal forecasting of co concentrations using an attention-based lstm model in a coastal urban environment," *Environmental Pollution*, vol. 300, p. 118954, 2022.

[5] G. S. Babu and B. E. Reddy, "A moving-window time series analysis technique for short-term forecasting of air pollution in hyderabad," *Environmental Modelling & Software*, vol. 52, pp. 119–127, 2014.

[6] Y. Zheng, F. Liu, and H.-P. Hsieh, "U-air: When urban air quality inference meets big data," in *KDD '13: Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2013, pp. 1436–1444.

[7] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2623–2631.

[8] M. Löning, A. Bagnall, S. Ganesh, V. Kazakov, J. Lines, and A. Bostrom, "sktime: A unified interface for machine learning with time series," Workshop on Systems for ML at NeurIPS 2019, 2019. [Online]. Available: https://arxiv.org/abs/1909.07872

[9] G. Ibarra-Berastegi, I. Madariaga, A. Serra, and et al., "Assessment of origin, distribution and local meteorological influence of carbon monoxide concentrations in a mid-sized city," *Chemosphere*, vol. 193, pp. 1062–1075, 2018.

[10] T. Chari, B. Bekbulat, K. Lee, and J. S. Apte, "Assessing air quality model performance for multiple pollutants: A comprehensive approach," *Environmental Science & Technology*, vol. 55, no. 15, pp. 10 251–10 262, 2021.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Políticas avanzadas de scheduling MAC para tráfico XR en redes 5G y 6G

Neco Villegas*, Ana Larrañaga-Zumeta†, Luis Diez*, Katerina Koutlia†, Sandra Lagén†, Ramón Agüero*
*Departamento Ingeniería de Comunicaciones, Universidad de Cantabria
neco.villegas@unican.es, {ldiez, ramon}@tlmat.unican.es
†Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA),
{alarranaga, kkoutlia, slagen}@cttc.es

Las aplicaciones eXtended Reality (XR) imponen cada vez más presión sobre las redes móviles para satisfacer sus estrictos requisitos de Quality of Service (QoS). En particular, los servicios XR exigen un throughput alto, además de una latencia baja, con patrones de tráfico que difieren considerablemente de los que caracterizan aplicaciones más tradicionales. Para afrontar las exigentes demandas impuestas por el tráfico XR, se propone un nuevo scheduler Medium Access Control (MAC), que aprovecha la segmentación del contenido XR en conjuntos de Packet Data Units (PDUs), conocidos como *PDU sets*. Estos PDU sets constituyen la unidad básica de tráfico para los servicios XR y deben ser entregados respetando sus requisitos específicos de QoS, los cuales son establecidos dinámicamente en tiempo real por la propia aplicación XR. La solución propuesta se apoya en la Teoría de Lyapunov, en particular en el algoritmo *drift-plus-penalty*, para asignar recursos eficientemente, estabilizar colas de tráfico y garantizar la entrega en plazo de los PDU sets. Se utiliza una implementación en el marco de ns-3 5G LENA y los resultados obtenidos ponen de manifiesto una mejora notable en comparación con otras alternativas, en cuanto a uso de recursos y cumplimiento de QoS.

***Palabras Clave*—5G, PDU set, XR, MAC, Lyapunov, ns-3**

I. Introducción

En respuesta a los retos que plantean las nuevas aplicaciones, con requisitos de QoS cada vez más estrictos, las redes 5G se han diseñado para soportar un abanico amplio de servicios, cada uno con sus propias necesidades. Además de abordar las demandas de alta capacidad de las aplicaciones enhance Mobile Broadband (eMBB), las redes 5G también deben acomodar el crecimiento de los servicios massive Machine-Type Communications (mMTC) y Ultra Reliable Low Latency Communications (URLLC), los cuales requieren un soporte adecuado para asegurar un número masivo de dispositivos conectados y una conectividad extremadamente fiable y de baja latencia, respectivamente. En este escenario, el 3GPP está inmerso en el desarrollo de 5G-Advanced, la próxima evolución de 5G, incorporando mecanismos y técnicas que permitan hacer frente a las complejidades introducidas por servicios de próxima generación.

Entre ellos, XR destaca por su combinación única de requisitos. Los dispositivos XR suelen ser incapaces de gestionar localmente todas las tareas de procesamiento necesarias y, en su lugar, dependen del apoyo de nodos externos. Un enfoque habitual es la división de la renderización, de manera que la carga computacional de la generación y el procesado multimedia se distribuye entre un servidor XR de alto rendimiento y el propio dispositivo. Las tareas de procesamiento intensivo, como el renderizado de escenas, se descargan en el servidor XR, mientras que las funciones más ligeras, como la corrección de la *pose* (posición y orientación), se gestionan localmente en el dispositivo. Además, el tráfico XR muestra patrones distintivos impulsados por la interacción del usuario en tiempo real, que normalmente tiene como consecuencia la existencia de varios flujos de vídeo que deben estar estrechamente sincronizados con frecuentes actualizaciones de movimiento y control. Debido a esta combinación de necesidades (velocidad de datos y estrictas limitaciones de latencia), los servicios XR podrían situarse entre eMBB y URLLC en términos de requisitos de red, lo que los convierte en un caso de uso clave y único para las innovaciones a introducir en la tecnología 5G-Advanced, así como en la próxima evolución de redes móviles, 6G.

En los servicios XR, cada unidad multimedia (como un fragmento de vídeo o un fotograma) se segmenta en PDUs antes de transmitirse a los usuarios. En el lado del cliente, se tiene que recibir un número mínimo de PDU de cada unidad multimedia antes de que pueda llevarse a cabo la descodificación o el procesamiento posterior. Así, este grupo mínimo de PDUs, que ha recibido el nombre de PDU set en las últimas especificaciones del 3GPP [1], constituye la unidad fundamental del tráfico

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

para servicios XR, por lo que debe considerarse en los mecanismos de asignación de recursos, teniendo en cuenta sus restricciones específicas de QoS.

El scheduler MAC es responsable de determinar cómo se asignan los recursos radio entre los flujos de tráfico. Aunque 5G proporciona el marco para soportar tráfico XR, no se definen algoritmos específicos en relación al scheduling MAC, dejando abiertos su diseño e implementación. Así, el principal objetivo de este trabajo es el diseño, implementación y evaluación de un scheduler MAC para redes 5G/6G, que tenga en cuenta las características específicas del tráfico XR, en concreto el concepto de PDU sets. Para ello se aprovecharán técnicas de control.

En su trabajo seminal [2], Neely introdujo un marco integral que utiliza técnicas de Lyapunov para la estabilización de colas. Abarca tanto los métodos de *max-weight* (o *min-drift*) como el problema de control *drift-plus-penalty*. El planteamiento no se centra únicamente en la estabilización de colas, sino que también permite integrar objetivos promediados en el tiempo, dentro del proceso de toma de decisiones. En particular, la optimización de Lyapunov de drift-plus-penalty es una técnica bien establecida, que se ha aplicado ampliamente a las redes de colas y a otros sistemas estocásticos. Estabiliza colas reales o virtuales al tiempo que optimiza el rendimiento de la red minimizando la función de penalización de la red promediada en el tiempo. En el contexto de scheduling MAC, en este trabajo se propone aprovechar esta metodología para diseñar un nuevo mecanismo de scheduling sensible a los PDU sets, que optimice la asignación de recursos, estabilice las colas de tráfico y garantice el cumplimiento de los requisitos de QoS de los servicios XR.

En general, las principales contribuciones de este trabajo son:

- La propuesta de un scheduler MAC drift-plus-penalty que mejora la optimización de la asignación de recursos entre flujos con estrictos requisitos, teniendo en cuenta las características únicas del tráfico XR.
- La solución que se propone aborda múltiples objetivos simultáneamente, pudiendo configurar la importancia de cada uno de ellos. Se demuestra que es capaz de garantizar los requisitos específicos de cada aplicación XR, asegurando al mismo tiempo un uso eficiente de los recursos asignados y la estabilización de las colas Radio Link Control (RLC).
- El scheduler MAC propuesto se integra dentro de ns-3 5G LENA [3]. Esto permite un detallado análisis de su rendimiento, a través de una extensa campaña experimental, comparándolo con otros schedulers del estado del arte.

El resto del documento se organiza de la siguiente manera. En la Sección II se presentan los antecedentes del marco de gestión de QoS de 5G NR, con especial énfasis en el concepto de PDU set. A continuación, la Sección III revisa los trabajos relacionados y destaca cómo nuestro enfoque realiza una contribución clara respecto al estado del arte. La Sección IV presenta la propuesta de scheduler basado en drift-plus-penalty. La Sección V evalúa el rendimiento de la solución implementada, a través de simulaciones realizadas en 5G-LENA, y comparándola con estrategias de scheduling alternativas sobre escenarios con tráfico XR mixto. Finalmente, en la Sección VI se concluye el trabajo y se esbozan futuras líneas de investigación.

II. Antecedentes

El modelo de calidad de servicio de 5G se estructura en torno al concepto de flujos de QoS, los cuales se clasifican en Guaranteed Bit Rate (GBR), Delay Critical GBR (DC-GBR) (ambos requieren un Guaranteed Flow Bit Rate (GFBR) específico) y Non-GBR (que no lo requieren). Un flujo QoS representa el nivel más alto de granularidad para diferenciar la QoS dentro de una sesión PDU. Cada flujo se identifica mediante un QoS Flow Indicator (QFI), que se utiliza en todo el 5G System (5G-S) para aplicar un tratamiento adecuado del tráfico, por ejemplo scheduling y control de admisión, para flujos del plano de usuario que compartan el mismo QFI en una sesión PDU determinada. El QFI se incluye en una cabecera de encapsulación a través de la interfaz N3 (y N9), garantizando que las cabeceras de los paquetes de extremo a extremo permanezcan inalteradas.

Las características de cada flujo de QoS vienen definidas en su perfil de QoS. Cada flujo de QoS se asocia a un 5G QoS Identifier (5QI), que define el tratamiento que debe recibir dicho flujo a través del 5G-S. Un perfil de QoS incluye, entre otros:

- **GFBR**. Especifica la tasa binaria mínima que debe proporcionarse al flujo.
- **Packet Delay Budget (PDB)**. Indica el retardo máximo que pueden experimentar los paquetes entre el UE y la User Plane Function (UPF), que termina la interfaz N6.
- **Packet Error Rate (PER)**. Define un límite superior para la tasa de PDUs que son procesadas por el emisor en la capa de enlace (por ejemplo, RLC) pero que no consiguen ser entregadas con éxito a la capa superior (por ejemplo, PDCP) en el receptor correspondiente dentro de la Radio Access Network (RAN).

El 5QI puede estar o no estandarizado (personalizado según las necesidades del operador). Desempeña un papel clave a la hora de permitir una gestión diferenciada de QoS en una amplia gama de servicios, incluidos el streaming de vídeo, Voice over Internet Protocol (VoIP) y las aplicaciones XR.

Tradicionalmente, el control de QoS en redes móviles se ha basado en dos conceptos principales: flujos y paquetes individuales. Sin embargo, los enfoques de scheduling pasan por alto la interdependencia entre paquetes, que es un aspecto clave para servicios multimedia interactivos como XR, en los que una única unidad de capa multimedia se transmite a menudo a través de múltiples paquetes. Para permitir la transmisión integrada y diferenciada de unidades multimedia (por ejemplo, un fotograma o un fragmento de vídeo), se ha introducido recientemente el

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

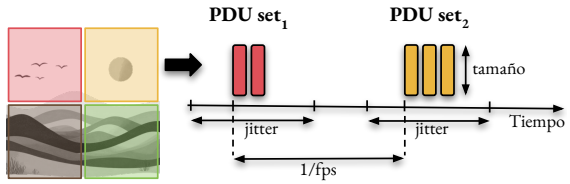


Fig. 1. Modelo de tráfico XR en DL del 3GPP.

concepto de PDU set. Un PDU set representa, por tanto, el grupo de paquetes que transportan colectivamente una unidad multimedia. Todas las PDU de un PDU set tienen la misma importancia en la capa de aplicación. Además, la capa de aplicación necesita el conjunto completo de PDU para procesar la unidad de información correspondiente. Sin embargo, en determinados escenarios, la aplicación puede seguir asumiendo que faltan algunas PDU [4].

Una de las distribuciones de tráfico más relevantes es la correspondiente al flujo de fotogramas de un vídeo, cuyo patrón sigue una distribución gaussiana truncada. Esta distribución caracteriza el tamaño de las unidades multimedia, es decir, los PDU sets, así como la fluctuación o variación del tiempo entre llegadas (en UpLink (UL), la fluctuación suele despreciarse, ya que el vídeo se genera más cerca o en el propio UE) [5]. Cada PDU set se transmite periódicamente con un tiempo medio entre llegadas igual al inverso de la frecuencia de fotograma (1/fps), pero sujeto a posibles fluctuaciones (*jitter*). Este aspecto debe ser gestionado cuidadosamente por el scheduler para cumplir los estrictos requisitos de retardo y fiabilidad de los servicios XR. Así, la dinámica temporal y la distribución del tamaño de los PDU sets son factores clave que influyen en la QoS extremo a extremo. La Fig. 1 ilustra el modelo de tráfico XR.

El 5G Core (5G-C) adquiere cierta información para soportar la gestión de la QoS a nivel de PDU set, pudiendo clasificarse en metadatos de PDU set y parámetros de QoS a nivel de PDU set.

Los metadatos correspondientes al PDU set son los siguientes:

- **PDU Set Sequence Number (PSSN) y PDU Sequence Number (PSN)**. El PSSN representa el número de secuencia del PDU set al que pertenece la PDU actual, sirviendo como identificador del PDU set. El PSN denota el número de secuencia de la PDU actual dentro de ese PDU set.
- **End PDU indicator (EPDU)**. Especifica si la PDU actual es la última del PDU set.
- **PDU Set Size (PSSize)**. Corresponde al tamaño total, en bytes, de todas las PDU del PDU set al que pertenece la PDU. Este campo es opcional.

Por su parte, los parámetros de QoS a nivel de PDU set son:

- **PDU Set Delay Budget (PSDB)**. Especifica el retardo máximo que puede experimentar un PDU set durante su transferencia entre el UE y el punto de terminación N6 en el UPF. Este retardo se mide desde el momento en que llega la primera PDU para su transmisión (en el punto de terminación N6 para DownLink (DL) o en el UE para UL) hasta que se han recibido con éxito todas las PDU del PDU set (en el UE para DL o en el punto de terminación N6 para UL).
- **PDU Set Error Rate (PSER)**. Define un límite superior para la tasa de pérdida de PDU sets no relacionados con la congestión.
- **PDU Set Importance (PSI)**. Establece la importancia relativa de un PDU set dentro de un flujo de QoS, en comparación con otros conjuntos de PDU. Esta información puede utilizarse, por ejemplo, para descartar paquetes a nivel de PDU set en episodios de congestión. En general, se da mayor prioridad a las fotogramas/slices intracodificados que a los fotogramas P/B o a los slices P/B [4].
- **PDU Set Integrated Handling Indicator (PSIHI)**. Indica si se requiere recibir correctamente todas las PDU del conjunto en la capa de aplicación del receptor.

Para los flujos que utilizan la gestión de QoS basada en PDU sets, el 5G-C proporciona uno o más parámetros (relacionados con PDU set) a la RAN. En concreto, los establece el Policy Control Function (PCF) y el Session Management Function (SMF) los incluye en el perfil QoS proporcionado a la RAN, siendo la fuente original de estos parámetros la propia aplicación XR.

Para el tráfico DL se requiere que el UPF identifique los PDU sets y sus características, en particular la relevancia y su tamaño. El UPF utiliza una descripción de protocolo que indica cómo realizar la identificación de un flujo de tráfico concreto, y trasladar posteriormente a la RAN la información del PDU set asociado. La Fig. 2 ilustra cómo se transfiere la información sobre los PDU sets a través de 5G-S, tal y como se ha descrito anteriormente. Además, como se puede ver, el UPF mapea los PDU sets identificados a flujos de QoS que requieren una gestión basada en PDU sets, que luego son mapeados a Data Radio Bearers (DRBs) por la RAN. 3GPP define la notación L:M:K para describir las configuraciones de asignación entre PDU sets (L), flujos de QoS (M) y DRB (K). Debido a la visibilidad limitada de la información sobre PDU sets entre capas de protocolo, sólo se admiten las correspondencias específicas 1:1:1, N:1:1 y N:N:1. El mapeo 1:1:1 permite un control detallado de la QoS por flujo de aplicación, mientras que los mapeos N:1:1 y N:N:1 ofrecen una coordinación más sencilla, al tratar las PDU agrupadas con la misma prioridad.

III. Trabajos Relacionados

Esta sección revisa los pocos trabajos recientes que se centran en el scheduling MAC en redes 5G con un enfoque particular en los servicios XR, identificando las diferencias respecto a este trabajo, para destacar las aportaciones más relevantes.

Aunque la investigación sobre la gestión de QoS para el tráfico XR en 5G NR ha sido intensa en los últimos meses, los estudios existentes relacionados con scheduling

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

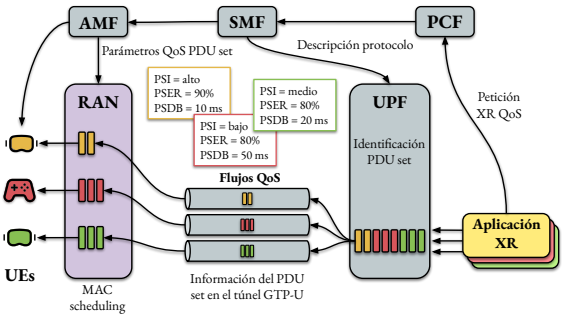


Fig. 2. Transferencia de información del PDU set a través de los componentes 5G-S para habilitar la gestión adecuada de la QoS.

MAC son muy pocos. Los autores de [6] proponen un scheduler Proportional Fairness (PF) basado en la demanda para diferenciar entre tráfico de VR y UHD en escenarios eMBB. El scheduler sensible al QoS en [7] realiza una clasificación del tráfico, incluyendo servicios XR y mixtos, basándose en el 5QI asociado. Los autores de [8] introducen una estrategia de scheduling que combina razones de respuesta ponderadas por canal y retardo, para gestionar tráfico heterogéneo en tiempo real y eMBB. Esta propuesta incorpora una función de priorización que considera la proximidad al límite de retardo, el tamaño de los paquetes y la calidad del canal.

Además, la gestión del tráfico basado en PDU sets es un concepto relativamente nuevo, introducido recientemente en las últimas especificaciones de la tecnología 5G-Advanced, por lo que la literatura existente sobre este tema es aún muy escasa. Los pocos trabajos relacionados destacan la necesidad de investigar más a fondo cómo podrían aprovecharse los mecanismos sensibles a los PDU sets para llevar a cabo una gestión de QoS más eficiente. Los autores de [9] sostienen que el diseño intercapas sensible a tráfico XR en 5G-Advanced representa un paso fundamental hacia un scheduling eficiente de estos servicios, lo que permite obtener mejoras iniciales en términos de rendimiento y latencia. Además, se propone un scheduler PF sensible al retardo, y se demuestra sus mejoras de rendimiento en comparación con un scheduler PF tradicional, de referencia. Por otro lado, [10] propone un scheduler que aprovecha la información de los PDU sets. En este caso, en lugar de centrarse solo en los retardos individuales de cada PDU, el scheduler prioriza la transmisión de PDUs que pertenecen al mismo PDU set. En ambos casos, el rendimiento se evalúa mediante simulaciones a nivel de sistema, respaldadas por un conjunto exhaustivo de modelos matemáticos.

Como se introduce en nuestro trabajo previo [11], el marco de scheduling propuesto, basado en la Teoría de Lyapunov, considera la mayoría de las métricas tradicionalmente utilizadas por trabajos relacionados para abordar el problema de scheduling, incluyendo la calidad del canal, la estabilidad del tráfico, el throughput (rendimiento) y el retardo, al tiempo que promueve una utilización eficiente de los recursos radio. En este trabajo, se incluye además la sensibilidad a los PDU sets en el algoritmo. Este enfoque solo se ha considerado en el scheduling MAC por dos trabajos previos (mencionados anteriormente), pero con un alcance más limitado. Ambos se centran principalmente en las características de los PDU sets, mientras que la contribución descrita en este artículo ofrece una solución más integral, al incorporar una gama más amplia de métricas de QoS y adoptar un enfoque novedoso para aprovechar la información que aporta la aplicación XR. Así, este trabajo representa un avance claro en el campo, proporcionando una solución de scheduling más holística y eficiente para aplicaciones XR en tiempo real con estrictos requisitos de QoS.

IV. Modelo de sistema

Esta sección presenta el diseño del scheduler propuesto y detalla su integración con los conceptos específicos de 5G introducidos en la Sección II. La solución gestiona múltiples colas de tráfico, cada una asociada con un 5QI específico y, potencialmente, sujeta a un requisito GFBR concreto e información del PDU set. Se define el conjunto de colas como $\mathcal{N}$. El objetivo de la solución propuesta es asegurar la estabilidad de las colas, cumplir con las restricciones de QoS y, opcionalmente, minimizar la utilización total de recursos. Para simplificar, se asume que cada cola corresponde a un único UE. La modificación para considerar una formulación más general, que mapee colas a Logical Channels (LCs), se podría realizar de manera sencilla.

Se asume un tiempo ranurado, donde $Q_n(t)$ denota el tamaño de la cola $n \in \mathcal{N}$ en t, mientras que $\alpha_n(t)$ representa la decisión de scheduling. El sistema está influenciado por factores estocásticos, como la Signal to Interference plus Noise Ratio (SINR), $\gamma_n(t)$, y la llegada de tráfico a cada cola, $a_n(t)$. La dinámica de cada cola considera tanto las llegadas como las salidas, $b_n(t)$, según establece la siguiente expresión:

$$Q_n(t+1) = \max\{Q_n(t) - b_n(t), 0\} + a_n(t) \quad (1)$$

Además, se representa el throughput instantáneo percibido por el usuario n-ésimo en el instante t como $\rho_n(t)$. Como resultado, los tráficos entrante y saliente por cola, junto con el throughput percibido por el usuario, pueden representarse como funciones generales de las decisiones de scheduling y de eventos estocásticos:

$$a_n(t) = \hat{a}_n(\gamma_n(t)) \quad (2)$$

$$b_n(t) = \hat{b}_n(\alpha_n(t), \gamma_n(t)) \quad (3)$$

$$\rho_n(t) = \hat{\rho}_n(\alpha_n(t), \gamma_n(t)) \quad (4)$$

El objetivo que se persigue es asegurar que se cumpla, en promedio, un requisito de throughput a lo largo del tiempo, Γ_n, para el usuario n-ésimo. Por tanto, se considera la esperanza del throughput promediada en el tiempo $\overline{\rho_n} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\{\rho_n(t)\}$. La definición de Γ_n varía dependiendo del tipo de servicio, como se puede ver en la Ec. (5). Para un flujo XR con información disponible del PDU set, el requisito se calcula como el throughput mínimo promedio en el tiempo durante el resto del PSDB, lo que implica excluir el Head of Line (HOL), necesario

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

para asegurar que el PDU set se entregue a tiempo. Para aquellos flujos de QoS sin información de los PDU sets, como información de *pose* o de UEs que no son XR, el objetivo de throughput corresponde al GFBR, si está definido.

$$\Gamma_n = \begin{cases} \frac{\text{PSSize}_n}{(\text{PSDB}_n\text{-HOL}_n)} & \text{si es XR con PDU set QoS y PSDB}_n > \text{HOL}_n \\ \text{GFBR} & \text{si es GBR o DC-GBR} \\ 0 & \text{si PSDB}_n \leq \text{HOL}_n \text{ o no es GBR} \end{cases} \quad (5)$$

Además, se busca minimizar la cantidad de recursos asignados, siendo $f(t) = \sum_{n\in\mathcal{N}} \alpha_n(t)$ la cantidad de recursos (es decir, Resource Blocks (RBs)) asignados en el instante t. Dado que la minimización de los recursos (sin tener en cuenta otros parámetros) podría llegar a perjudicar el throughput, se incluye, como objetivo, la minimización de los recursos asignados en promedio, $\overline{f} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\{f(t)\}$.

Se aplica la toma de decisiones sobre un espacio de acciones admisibles que varía con el tiempo, $\alpha(t) \in \mathrm{A}(t)$, asegurando que el número de RBs asignados no exceda la cantidad disponible en cada slot. Además, se define el déficit de throughput como $D_n(t) = \Gamma_n - \rho_n(t)$, el cual cumple que $\overline{D_n} = \mathbb{E}\{\Gamma_n - \rho_n(t)\} = \Gamma_n - \overline{\rho_n}$. En general, el scheduler aborda el problema de optimización estocástica formulado en el Problema 1, donde la restricción en la Ec. (7) garantiza que se cumplan los requisitos de throughput. Al mismo tiempo, el scheduler busca mantener la estabilidad en media de las colas de tráfico, $Q_n(t)$, para todo $n \in \mathcal{N}$, de modo que $\lim_{t\to\infty} \frac{\mathbb{E}\,|Q_n(t)|}{t} = 0$.

Problem 1:

$$\min_{\alpha(t)} \quad \overline{f} \quad (6)$$

$$\text{s.a.} \quad \overline{D_n} \leq 0 \quad \forall n \in \mathcal{N} \quad (7)$$

$$\alpha(t) \in \mathrm{A}(t) \quad (8)$$

$$\text{Las colas } Q_n(t) \text{ son estables en media } \forall n \in \mathcal{N} \quad (9)$$

Para abordar el Problema 1, se adopta el marco descrito en [2], que extiende el algoritmo max-weight. Inicialmente, se define una cola virtual $G_n(t)$ en el instante t, y su valor se actualiza para el siguiente instante $t+1$ como sigue:

$$G_n(t+1) = \max\{G_n(t) + D_n(t), 0\} \quad (10)$$

El problema anterior puede resolverse aplicando el algoritmo *mínimo de deriva más penalización* (*min drift-plus-penalty*) [2, Capítulo 4], que se reduce a resolver, en cada slot, el siguiente:

Tabla I
CONFIGURACIÓN DEL ESCENARIO DE EVALUACIÓN.

Parametro	**Valor**
Tiempo de simulación	60 seg.
Escenario	InH-OfficeOpen
Numerología (μ)	4
Frecuencia portadora	30 GHz
Ancho de banda	100 MHz
Potencia Tx gNB/UE	41 dBm / 23 dBm
Array antenas gNB	1 TXRU: 4 × 8 (elementos 3GPP)
Array antenas UE	1 TXRU: 1 × 1 (elemento isotrópico)
Ganancia elemento antena gNB/UE	8 dBi / 0 dBi
Figura de ruido gNB/UE	5 dB / 7 dB
Densidad de potencia espectral de ruido	-174 dBm/Hz
Duplexado	TDD

Problem 2:

$$\min_{\alpha(t)} \quad V \sum_{n\in\mathcal{N}} \alpha_n(t) + \omega_Q \sum_{n\in\mathcal{N}} Q_n(t)[a_n(t) - b_n(t)] \quad (11)$$

$$+ \omega_G \sum_{n\in\mathcal{N}} G_n(t) D_n(t)$$

$$\text{s.a.} \quad \alpha(t) \in \mathrm{A}(t) \quad (12)$$

$$b_n(t) \leq Q_n(t) \quad \forall n \in 1, \ldots, N \quad (13)$$

donde la función objetivo consta de tres términos distintos. El primero, conocido como *penalty*, busca minimizar la asignación de recursos. El segundo y tercer términos corresponden con el *drift* de Lyapunov, que se encarga de estabilizar las colas de tráfico, y las colas virtuales asociadas a los requisitos de throughput. El *penalty* está escalado por un parámetro de diseño, V, que regula el compromiso entre minimizar el uso de recursos y reducir el tamaño de las colas. Además, las colas de tráfico y virtuales están ponderadas por coeficientes, denotados por w_Q y w_G, respectivamente, para equilibrar su influencia en la decisión de scheduling.

V. EVALUACIÓN

En esta sección se analiza el rendimiento del scheduler propuesto. El objetivo principal de esta evaluación es medir las mejoras obtenidas al incorporar la información del PDU set en el proceso de scheduling. Se compara la solución propuesta con otros schedulers de la literatura sensibles a PDU sets [9], [10], el scheduler basado en GFBR [11], y una solución de base, el comúnmente utilizado PF.

El rendimiento de todos los schedulers se evalúa en una topología simple, que consiste en un gNB y varios UEs, como se detalla en la Tabla I. En este escenario, todos los UEs generan tráfico XR, específicamente algunos generan flujos Virtual Reality (VR) y otros Cloud Gaming (CG), cuyas características se describen en la Tabla II. Como se puede ver, el tráfico VR está configurado con un perfil de QoS más estricto (5QI=88), correspondiente a un tipo de recurso DC-GBR, mientras que el tráfico CG utiliza el tipo de recurso GBR con un valor de 5QI más bajo (5QI=3).

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla II
CONFIGURACIÓN DEL TRÁFICO XR.

Tráfico VR		**Tráfico CG**	
5QI	88	5QI	3
Tasa promedio	45 Mbps	Tasa promedio	30 Mbps
FPS	120	FPS	120
GFBR	45 Mbps	GFBR	30 Mbps
PSDB	10 ms	PSDB	50 ms

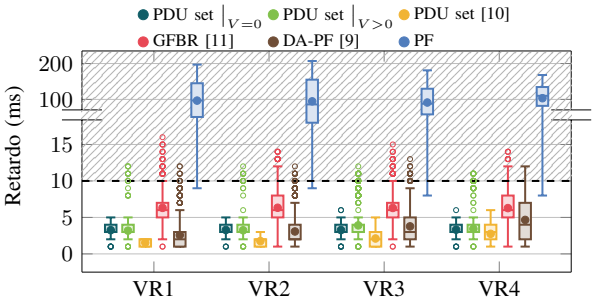


Fig. 3. Retardo experimentado por cada PDU set de los UEs VR bajo diferentes schedulers. Las líneas discontinuas indican el PSDB.

Las Figuras 3 y 4 muestran el retardo del PDU set experimentado por los UEs VR y CG, respectivamente, en un escenario con 8 UEs XR en una celda. El retardo se representa mediante boxplots, donde los bordes inferior y superior de cada caja corresponden con los percentiles 25 y 75, respectivamente. Los bigotes se extienden hasta 1.5 veces el Inter Quartile Range (IQR), cubriendo aproximadamente desde el percentil 2 hasta el 98. Los círculos rellenos dentro de cada boxplot representan el retardo promedio, mientras que los círculos vacíos son valores atípicos observados en los diferentes experimentos.

Cada scheduler se representa con un color distinto, dando lugar a un boxplot por UE para cada scheduler. El scheduler propuesto aparece dos veces, correspondiente a dos configuraciones diferentes: verde oscuro para $V = 0$, lo que significa que no se aplica minimización de recursos, y verde claro para $V > 0$, donde el scheduler intenta minimizar el uso de recursos mientras cumple con los demás objetivos. En ambos casos, $\omega_Q = \omega_G = 1$.

La línea discontinua indica el PSDB, establecido en 10 y 50 ms para VR y CG, respectivamente. Como se muestra en la Figura 4, todos los schedulers cumplen con el PSDB para el tráfico CG, excepto algunos PDU sets fuera de

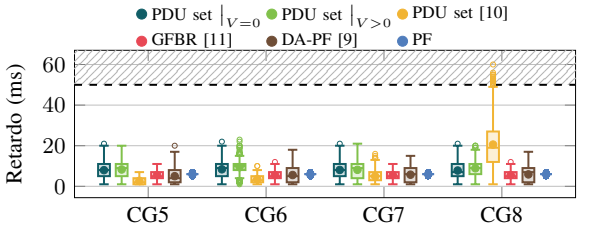


Fig. 4. Retardo experimentado por cada PDU set de los UEs CG bajo diferentes schedulers. Las líneas discontinuas indican el PSDB.

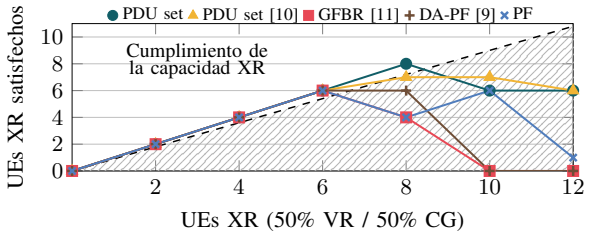


Fig. 5. Capacidad XR como función del número de UEs en la celda bajo diferentes schedulers.

tiempo (2%) en el caso del scheduler sensible al PDU set de [10]. Sin embargo, esto no ocurre con el tráfico VR. La Figura3 muestra que los schedulers sensibles al PDU set ofrecen un rendimiento notablemente mejor, manteniendo de manera constante el retardo del PDU set por debajo del PSDB. Cabe destacar que, cuando la minimización de recursos está activa ($V > 0$), algunos PDU sets superan el límite de 10 ms, aunque representan menos del 1% del total. Entre las demás soluciones, el scheduler sensible al GFBR ofrece un buen comportamiento, debido a la priorización que realiza del tráfico VR, basado en su mayor valor de GFBR. Por otro lado, PF intenta equilibrar el rendimiento y el equilibrio sin considerar explícitamente las restricciones de QoS, lo que resulta en un peor rendimiento para el tráfico VR. En consecuencia, la solución de base (tradicional) no logra cumplir con los estrictos requisitos de retardo impuestos por el PSDB, alcanzando retardos de hasta 200 ms, 20 veces superiores al objetivo de 10 ms.

El comportamiento observado con las distintas soluciones es especialmente relevante al considerar la "capacidad XR", que es el principal Key Performance Indicator (KPI) definido en los estudios del Release 17 del 3GPP para servicios XR [5]. Se define como el número máximo de UEs XR por celda para el cual al menos el 90% de los UEs se consideran satisfechos, lo que sucede cuando más del 99% de sus PDUs se entregan exitosamente dentro de su PSDB. Según los resultados de la simulación, la solución basada en el GFBR y PF solo logran satisfacer a los UEs CG (es decir, el 50% de los UEs XR en la celda), quedando por debajo del objetivo del 90%. Los schedulers sensibles al PDU set presentes en la literatura aumentan el número de UEs satisfechos, con un 62.5% y 87.5% para DA-PF [9] y el scheduler de [10], respectivamente. Por otro lado, el scheduler propuesto en este trabajo, sensible al PDU set, en sus dos configuraciones, alcanza un 100% de UEs satisfechos, siendo la única solución que cumple con la capacidad XR de 8 UEs. En este sentido, la Figura 5 muestra el número de UEs XR satisfechos en función del número total de UEs en la celda, que varía de 0 a 12. Una línea discontinua indica el umbral por encima del cual se cumple la capacidad XR definida por el 3GPP.

Finalmente, la Figura 6 muestra boxplots del rendimiento promedio por ventana de 2 segundos, de acuerdo con el 5QI estandarizado. Los resultados se presentan para 8, 10 y 12 UEs en la celda. Para mayor

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

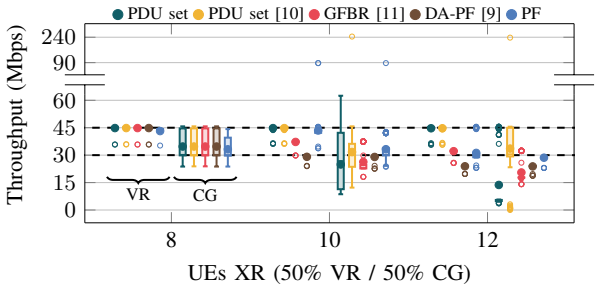


Fig. 6. Throughput promedio en cada ventana de 2 segundos experimentado por los UEs XR con diferentes schedulers. Las líneas discontinuas indican los GFBRs.

claridad, los UEs se agrupan según el tipo de tráfico (VR o CG). Como puede observarse, en el escenario menos exigente (8 UEs), todos los schedulers alcanzan valores similares de rendimiento, muy próximos al GFBR (líneas discontinuas). Esto demuestra que la mejora en retardo lograda por el scheduler sensibles al PDU set no se produce a costa de incumplir el requisito de GFBR. En los escenarios con 10 y 12 UEs, el rendimiento promedio por ventana de 2 segundos por UE disminuye, debido al aumento del tráfico total. Sin embargo, los schedulers sensibles al PDU set logran mantener un rendimiento de 45 Mbps para los UEs VR, debido a su bajo PSDB en comparación con CG. Esto no ocurre con DA-PF, que no logra sostener el mismo nivel de rendimiento para el tráfico VR bajo mayor carga. Esto se debe en parte a su configuración: el exponente de la métrica PSDB [9] no es lo suficientemente alto como para contrarrestar la influencia de la métrica PF, que tiende, en general, a igualar el throughput entre todos los UEs, como se observa para el scheduler PF. También cabe destacar que, al incorporar la ocupación del búfer RLC en la decisión de scheduling, las soluciones sensibles al PDU set y GFBR evitan desbordamientos de búfer, algo que los demás schedulers no consideran. Esto explica los valores atípicos observados: 240 Mbps para el scheduler sensible del PDU set de [10] y 90 Mbps para PF. Además, al no considerar el estado del búfer, podrían causar una degradación de la equidad entre UEs.

VI. Conclusiones

En este artículo, se ha presentado un nuevo scheduler MAC multiobjetivo, sensible a la información de QoS de PDU sets propios de tráfico XR. Se trata de uno de los primeros trabajos que explotan esta información, a nivel de PDU set, para guiar las decisiones de scheduling. Al considerar explícitamente los tiempos de entrega de cada PDU set, el scheduler propuesto garantiza que se cumplan las restricciones temporales de los servicios XR, al mismo tiempo que mantiene la estabilidad de las colas y promueve un uso eficiente de los recursos radio.

La solución propuesta se ha implementado e integrado en ns-3 5G-LENA. A diferencia de otros schedulers de la literatura, que no consideran los requisitos de QoS de XR, el enfoque descrito demuestra, a través de las extensas simulaciones que se han llevado a cabo, una capacidad constante para cumplir con las restricciones de retardo impuestas por los servicios XR. Además, permite una reducción en el uso de recursos, sin comprometer el cumplimiento general de los requisitos de QoS.

En futuros trabajos, se continuará aprovechando los últimos avances en gestión de QoS que se hayan introducido en las releases más recientes del 3GPP (tecnologías 5G y 6G). En este sentido, se ampliarán las políticas del scheduler, optimizando sus capacidades de configuración para incorporar parámetros adicionales de QoS y restricciones específicas de nuevos servicios, mejorando así la adaptabilidad y eficacia de la solución propuesta en una gama más amplia de escenarios de tráfico.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el programa de Horizonte Europa Smart Networks and Services Joint Undertaking (SNS JU) bajo el acuerdo de subvención #101139282 (6G-SENSES). También ha recibido fondos del Ministerio de Asuntos Económicos y Transformación Digital a través del proyecto 6GBLUR/JOINT y 6GBLUR/SMART (TSI-063000-2021-56/57), y por el Gobierno Regional de Cantabria a través de los programas TCNIC (2023/TCN/002) y Concepción Arenal (UC-23-40).

Referencias

[1] 3GPP, "System architecture for the 5G System (5GS) (3GPP TS 23.501 version 18.9.0 Release 18)," 3GPP, Tech. Rep. TS 23.501 V18.9.0 (2025-04), 2025.

[2] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, 2010.

[3] N. Patriciello, S. Lagén, B. Bojović, and L. Giupponi, "An E2E Simulator for 5G NR Networks," *Simulation Modelling Practice and Theory*, vol. 96, p. 101933, 2019.

[4] 3GPP, "5G Real-time Media Transport Protocol Configurations (3GPP TS 26.522 version 19.0.0 Release 18)," 3GPP, Tech. Rep. TS 126 522 V19.0.0 (2025-03), 2025.

[5] ——, "Study on XR (Extended Reality) Evaluations for NR (Release 17)," 3GPP, Tech. Rep. TR 38.838 V17.0.0 (2021-12), 2021.

[6] A. M. Nor, O. Fratu, S. Halunga, A. Alyosef, Z. D. Zaharis, S. Rizou, and P. I. Lazaridis, "Demand based Proportional Fairness Scheduling for 5G eMBB Services," in *2022 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2022.

[7] K. Koutlia, B. Bojovic, S. Lagén, X. Zhang, P. Wang, and J. Liu, "System analysis of QoS schedulers for XR traffic in 5G NR," *Simulation Modelling Practice and Theory*, vol. 125, p. 102745, 2023.

[8] P. R. Priyadarshini, J. Bapat, and A. Mishra, "Channel and Delay Weighted Response Ratio Scheduler for 5G and Beyond Systems," *IEEE Communications Letters*, vol. 28, no. 7, pp. 1658–1662, 2024.

[9] R.-A. Stoica, H. Bagheri, J. Löhr, and D. Karampatsis, "XR-Aware Next Generation Networks," *IEEE Communications Magazine*, vol. 62, no. 8, pp. 62–68, 2024.

[10] P. Paymard, S. Paris, A. Amiri, T. E. Kolding, F. S. Moya, and K. I. Pedersen, "PDU-set Scheduling Algorithm for XR Traffic in Multi-Service 5G-Advanced Networks," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 758–763.

[11] N. Villegas, A. Larrañaga, L. Diez, K. Koutlia, S. Lagén, and R. Agüero, "Extending QoS-aware scheduling in ns-3 5G-LENA: A Lyapunov based solution," in *Proceedings of the 2024 Workshop on ns-3*, ser. WNS3 '24. New York, NY, USA: Association for Computing Machinery, Jun. 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Laboratorio Abierto Científico-Tecnológico de Investigación en 6G de la UPC

M. Martínez-Morfa*, S. Garcia-Cantón*, C. Cervelló-Pastor*, J.A. Lázaro**, A. Umbert**
*Dpto. de Ingeniería Telemática, **Dpto. de Teoría de Señal y Comunicaciones,
Universidad Politécnica de Cataluña
*Esteve Terrades, 7, 08860, Castelldefels. **Jordi Girona, 1-3, 08034, Barcelona.
{mario.jose.martinez, sergi.garcia.canton, cristina.cervello, jose.antonio.lazaro, anna.umbert}@upc.edu

Este artículo presenta el despliegue y operación de una infraestructura experimental abierta de la Universitat Politècnica de Catalunya destinada a la investigación avanzada en redes 5G+ y 6G. La plataforma integra tecnologías de red celular, inalámbrica, óptica, *edge computing*, sensórica y automatización mediante inteligencia artificial. Distribuida entre dos campus interconectados, permite el desarrollo y validación de aplicaciones y servicios en escenarios reales. Se describe la arquitectura, las especificaciones funcionales de sus subsistemas y los servicios habilitados, consolidándose como un banco de pruebas de referencia para redes móviles de próxima generación y aplicaciones inteligentes.

***Palabras Clave*—5G+/6G, *Machine Learning*, Internet, Infraestructura de Investigación, Plataforma Multi-tecnología.**

I. INTRODUCCIÓN

El proyecto 6G-OpenLab [1] de la Universitat Politècnica de Catalunya (UPC) tiene como objetivo desarrollar una infraestructura de investigación abierta y avanzada en tecnologías 5G+ y 6G. Esta plataforma está diseñada para ofrecer a grupos de investigación, empresas y administraciones un entorno colaborativo que integra redes celulares de nueva generación, redes ópticas de alta capacidad, computación en el extremo, sensores y herramientas de medida de alta precisión, todo ello gestionado mediante redes IP inteligentes soportadas por inteligencia artificial y aprendizaje automático (ML, *Machine Learning*).

La motivación central del proyecto radica en impulsar la innovación y la generación de conocimiento en el ámbito 6G, fomentando la transferencia tecnológica y la internacionalización de la investigación mediante la cooperación con infraestructuras científicas europeas. Ha sido complementado mediante el proyecto ELEGANT [2], destinado a expandir las capacidades y permitir el despliegue de una plataforma abierta única.

Esta plataforma integrada, desplegada en el Campus del Baix Llobregat (CBL, en Castelldefels) y el Campus Nord (CN, en Barcelona) de la UPC, se caracteriza por ser:

- Una red celular autónoma y privada de próxima generación (5G+/6G).
- Una infraestructura de computación multi-acceso en el borde (MEC, *Multi-Acces Edge Computing*) que integra redes inalámbricas complementarias (Wi-Fi 6E/7, Radioenlaces de microondas, etc.)
- Un Centro de Procesado de Datos (CPD) con diversos conjuntos de servidores con diferentes configuraciones, implementando Kubernetes (K8s), OpenStack (OS) y computación de altas prestaciones (HPC, *High Performance Computing*).
- Una red de transporte óptica reconfigurable de alta capacidad para la interconexión de ambos campus e instancias de red móvil.
- Una red inalámbrica de sensores (WSN, *Wireless Sensor Network*) y un despliegue de robótica asociados a tecnologías de agricultura inteligente.
- Un sistema de automatización y monitoreo en tiempo real para minimizar la intervención humana y optimizar el funcionamiento global del sistema.
- Una plataforma de ML abierta y configurable apta para desarrollar algoritmos en la nube y en el borde.

Este entorno constituye un banco de pruebas de escenarios reales para el desarrollo y la validación de dispositivos, soluciones verticales y aplicaciones 5G+/6G con fuerte componente de ML. Su despliegue ha sido fruto del trabajo conjunto entre los comités técnico y de gestión de ambos proyectos, definiendo la arquitectura, la topología, la ubicación y la integración con los recursos existentes.

El presente artículo prosigue con la sección II, dedicada a la presentación de la arquitectura de la infraestructura, en la que se incluye la descripción de la topología física y lógica, la distribución entre campus y los componentes interconectados. La sección III detalla las especificaciones técnicas y operativas de cada subsistema. Posteriormente,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

en la sección IV se presenta una revisión de los principales catálogos de servicios habilitados y los posibles casos de uso asociados. Finalmente, el artículo concluye con una sección dedicada a la síntesis de los avances realizados y líneas futuras de investigación y explotación.

II. ARQUITECTURA DE LA INFRAESTRUCTURA

La arquitectura de la plataforma combina de manera integrada una estructura física distribuida con plataformas virtuales inteligentes, capaz de soportar experimentación avanzada y servicios 5G+/6G en entornos reales.

A. *Distribución física*

La infraestructura está distribuida entre dos campus de la UPC: por un lado el CBL en Castelldefels, y por otro lado el CN en Barcelona, con una separación geográfica de aproximadamente 25 km. Ambos campus albergan nodos esenciales interconectados mediante enlaces de fibra óptica de alta capacidad.

En el CBL se encuentra un CPD principal ubicado en la Escola d'Enginyeria de Telecomunicació i Aeroespacial de Castelldefels (EETAC). Además, se cuenta con infraestructuras complementarias, como el DronLab y la Agrópolis. El primero corresponde a una gran infraestructura de investigación, conformada por una estructura protegida para el vuelo de drones, mientras que el segundo está conformado por un invernadero dotado de sensores ambientales y un campo de cultivo provisto de una WSN destinada a aplicaciones de agricultura inteligente. El despliegue de red fija e inalámbrica en el CBL permite la interconexión del CPD con las infraestructuras complementarias y el despliegue de nodos de edge computing en espacios exteriores y calles del campus.

Por otra parte, en CN se encuentran las oficinas del Centre de Comunicacions Avançades de Banda Ampla (CCABA), compuesto por otro CPD como nodo central del campus. Además, se cuenta con una serie de laboratorios interconectados mediante una red experimental de fibra óptica dedicados a la investigación avanzada.

Ambos nodos principales constituyen el centro de una topología en estrella implementada en cada campus, la cual se emplea para interconectar sus respectivas infraestructuras complementarias. La interconexión entre estos puntos de presencia principales se realiza mediante una red de transporte óptico (OTN, *Optical Transport Network*) con un enlace de fibra óptica de aproximadamente 34 km, que utiliza la tecnología de multiplexado denso por división en longitudes de onda (DWDM, *Dense Wavelength Division Multiplexing*) para multiplexar diversos canales de 100 Gbps, garantizando una latencia inferior a 1 ms. De esta manera, la infraestructura de ambos campus funciona como una plataforma única e interconectada.

B. *Topología lógica y planos de funcionamiento*

La infraestructura se sustenta en dos planos de funcionamiento principales que coexisten sobre la capa física común: el plano de datos y el plano de control, operación y mantenimiento (también conocido como plano de gestión).

1) Plano de Datos: El plano de datos se construye bajo los principios de microservicios y virtualización de funciones de la red (NFV, *Network Function Virtualization*) que se sustentarán sobre los elementos físicos y virtuales de la capa física. Este plano, a su vez, está conformado por distintas dimensiones que se corresponden a las redes de las distintas tecnologías integradas.

La Fig. 1 ilustra la relación entre las diversas dimensiones del plano de datos. Como puede observarse, la red fija constituye el nexo de interconexión de las demás tecnologías, dado que está integrada por equipos preparados para conmutar el tráfico de un subsistema a otro sin afectar sus requerimientos específicos.

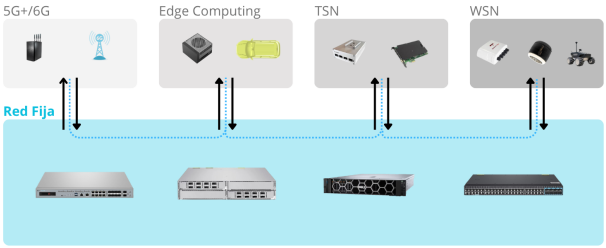


Fig. 1. Plano de datos con sus diferentes subsistemas de red.

La red fija, que se compone de elementos de seguridad, conmutación IP y transporte (tanto óptico como inalámbrico), está presente en ambos campus. La Fig. 2 muestra la topología de red fija de la infraestructura, donde se pueden apreciar los distintos elementos y la interconexión con los subsistemas complementarios.

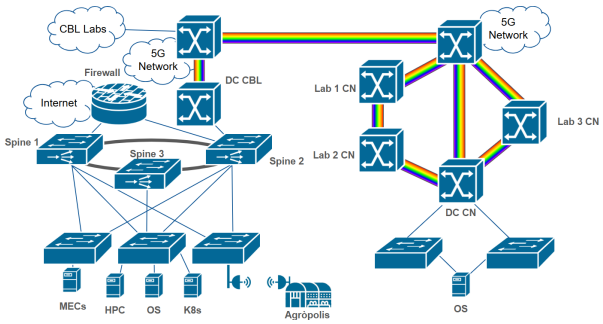


Fig. 2. Topología de red fija.

En el caso del CBL, se observa la presencia de un cortafuegos que actúa como barrera de seguridad frente al dominio externo de Internet. Además, se identifican los elementos de conmutación presentes en el CPD, que se interconectan formando una topología *leaf-spine* con enlaces de 200 Gbps y 400 Gbps para garantizar la comunicación entre los subsistemas y grupos de servidores, así como el manejo del tráfico a muy baja latencia.

En el extremo de CN, se ilustran los nodos OTN que conforman el anillo de laboratorios de investigación y el acceso a su correspondiente CPD.

2) Plano de Control: Las acciones de control, operación y mantenimiento se realizan por medio de una red de gestión totalmente paralela a la infraestructura que

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

compone el plano de datos. Esta red cuenta con equipos de conmutación independientes que permiten acceder a las interfaces de gestión de los equipos de red fija desde un punto de vista global.

La representación gráfica de la estructura de la red se muestra en la Fig. 3. Se dispone de un servidor de gestión destinado al despliegue de *frameworks* abiertos fundamentados en los principios de redes definidas por software (SDN, *Software Defined Networks*) y *network slicing*, con el propósito de gestionar de manera automatizada y con la mínima intervención humana los equipos de la red. Esta red cuenta también con servidores que ejecutan plataformas de gestión propias de subsistemas como la red OTN y la red 5G+/6G, lo que permite obtener, de manera más eficiente, distintas métricas en tiempo real a través de interfaces de programación de aplicaciones (APIs, *Application Programming Interfaces*).

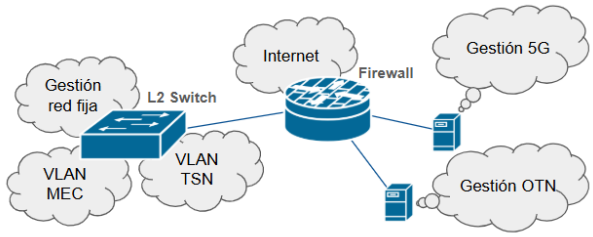


Fig. 3. Topología de red de gestión.

III. ESPECIFICACIONES FUNCIONALES DE LOS SUBSISTEMAS Y ELEMENTOS

Esta sección se dedica a describir las especificaciones técnicas y funcionales de los distintos subsistemas y elementos de red que componen la infraestructura del proyecto.

A. Subsistema de red fija y servidores de cómputo

1) Red IP: Para lograr la interconectividad entre los dispositivos, sistemas y servicios implementados en el marco de este proyecto, fue imperativo desplegar una red IP con la capacidad de gestionar volúmenes significativos de datos de manera escalable y con baja latencia. Para ello, se implementaron diversos conmutadores de altas prestaciones, los cuales se utilizan como nodos troncales en los CPD de ambos campus.

Como se ha mencionado anteriormente, esta red posee una configuración de tipo *leaf-spine*, que contribuye a que las redes de CPDs reduzcan la latencia y el número de saltos, lo que a su vez se traduce en una mayor eficiencia.

En el nivel de *spines* se cuenta con tres conmutadores de última generación interconectados en forma de anillo con enlaces de 400 Gbps. Específicamente, son dos equipos H3C S9855-32D [3] y un FS NC8200-4TD [4]. Cada *spine* se conecta con un enlace de 200 Gbps a cada uno de los *leaves*, formando así una conexión de malla completa entre ambos niveles.

Cada conmutador *leaf* corresponde a un FS5860-48SC [5] y posee una gran cantidad de puertos de 10 Gbps, lo que permite insertar servicios y subsistemas complementarios. La latencia de conmutación del tráfico entre estos subsistemas se garantiza por debajo de 1 μs en cada uno de los elementos de red.

2) Red de transporte óptica: El equipamiento óptico desplegado incluye dos nodos troncales OTN con tecnología de multiplexación óptica *add-drop* reconfigurable (ROADM, *Reconfigurable Optical Add-Drop Multiplexer*) C/F (*Colorless/Flexgrid*) de Nokia [6], con capacidad de expansión de hasta 9 grados, ubicados en cada uno de los campus. La conexión entre estos nodos se realiza mediante enlaces de fibra óptica, garantizando una latencia inferior a 1 ms. Además, se han instalado nodos OTN adicionales: cuatro en el CN y uno en el CBL, junto con *routers* para la interconexión con la red troncal de la UPC.

Para la coordinación y supervisión de toda la infraestructura óptica, se han implementado dos routers (uno en cada campus) y un sistema de reflectometría óptica de dominio temporal (OTDR, *Optical Time Domain Reflectometry*) ubicado en CN, que permite monitorizar en tiempo real el estado de las fibras ópticas y detectar posibles incidencias de transmisión.

La solución se ha basado en la implementación de una red óptica de última generación utilizando los equipos de transporte de la serie Nokia PSS 1830PSS-8 y PSS-16II [7], gestionados de forma centralizada a través del sistema *Wavesuite Network Operations Center* (WS-NOC). Este gestor de red facilita la configuración, operación y supervisión integral de la red de transporte extremo a extremo. La Fig. 4 muestra la topología de esta red de transporte con sus nodos y principales canales de transmisión definidos.

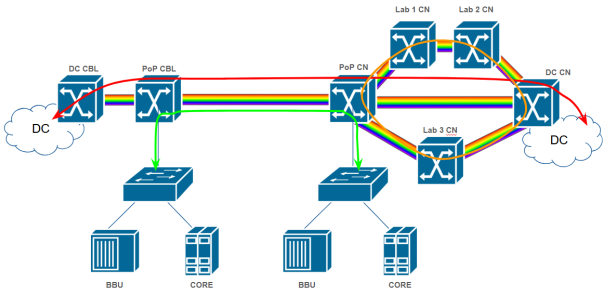


Fig. 4. Topología de la red de transporte óptico.

3) Radioenlace de microondas: Otra de las tecnologías integradas en la infraestructura son las redes de transmisión inalámbrica, en este caso los radioenlaces de microondas.

Con el objetivo de establecer una conexión entre la red troncal y la Agrópolis, localizada en Viladecans (Barcelona) y donde se ha desplegado la red de sensores y actuadores, se han desplegado dos radioenlaces digitales de ondas milimétricas. Uno de ellos se encuentra ubicado en el CBL, mientras que el otro se halla en la Agrópolis, estableciendo una conexión con una distancia de 5 km. El equipo instalado corresponde al modelo airFiber 60 Xtreme-Range, modelo AF60-XR de Ubiquiti [8], el cual opera en la banda de uso común de 57 a 71 GHz, de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

acuerdo con las normas UN-164 [9] del Cuadro Nacional de Atribuciones de Frecuencias (CNAF) y la norma ETSI EN 302 567 [10]. La frecuencia máxima de transmisión utilizada es de 68,04 GHz, con un ancho de banda de 2160 MHz, lo que, para una distancia de 5 km, le otorga una velocidad máxima de transmisión de datos de 2,7 Gbps en condiciones óptimas de transmisión libres de atenuación por lluvia. Dispone además de un enlace de *backup* en la banda de 5,15 GHz a 5,875 GHz, con un ancho de banda seleccionable entre 20, 40 o 80 MHz, que permite agregar fiabilidad al enlace.

4) Clústeres de Openstack: La infraestructura tecnológica está dotada de dos grupos (o clústeres) de servidores que implementan la plataforma de Openstack [11]. Esta integración habilita la implementación de computación en la nube para proporcionar una infraestructura como servicio (IaaS, *Infrastructure as a Service*).

Cada campus está equipado con un clúster desplegado que ofrece servicios de virtualización para diversos proyectos de investigación. En CN, se cuenta con 4 nodos Dell PowerEdge R6525 [12], lo que proporciona recursos totales de 1,5 TB de RAM y 168 TB de almacenamiento. Por otra parte, el CPD del CBL cuenta con ocho nodos del modelos Dell PowerEdge R440 [13], con un total de 768 GB de RAM y 11 TB de almacenamiento.

5) Clúster de Kubernetes: Con el objetivo de automatizar el despliegue, el ajuste de escala y el manejo de aplicaciones en contenedores, se ha desplegado la plataforma de Kubernetes [14] en otro de los clústeres de servidores presentes en el CBL. Este clúster está compuesto por cinco nodos, dos de ellos configurados como controladores (*masters*) y tres como nodos de cómputo (*workers*).

Cada nodo (Dell Poweredge R740xd [15]) cuenta con *hardware* diseñado para ejecutar cargas de trabajo que requieran una potencia considerable. Por lo tanto, cada uno ha sido equipado con dos procesadores de la última generación Intel, 256 GB de RAM, 2 TB de almacenamiento, y una GPU NVIDIA A100 [16], que cuenta con 40 GB de memoria. Esta última ofrece un rendimiento excepcional en aceleración a todos los niveles, y puede dividirse en hasta siete instancias de GPU para ajustarse dinámicamente a las demandas del usuario.

6) Clúster HPC: Para abordar cargas que requieran una alta potencia de cómputo, se ha desplegado un clúster de servidores para computación de alto rendimiento (HPC, *High-Performance Computing*).

Este clúster de supercomputadores emplea técnicas de paralelismo para gestionar grandes volúmenes de datos a altas velocidades. El sistema se encuentra ubicado en el CPD del CBL y cuenta con cuatro nodos (Gigabyte R283-SF1-AAL1 [17]), de los cuales uno es un controlador (nodo *head*) y tres son nodos de cómputo (nodos *compute*). Los nodos de cómputo cuentan con una alta capacidad de recursos para gestionar cargas de trabajo exigentes, con un total de 1,5 TB de RAM y 24 TB de almacenamiento. Además, cuentan con 6 GPUs de última generación (dos NVIDIA H200 y cuatro NVIDIA H100 [18], [19]), que poseen mayor capacidad de memoria (141 GB y 96 GB respectivamente) y mayor velocidad, lo que permite el impulso de la aceleración de la Inteligencia Artificial (IA) generativa y los modelos de lenguaje de gran tamaño (LLM, *Large Language Model*).

B. Subsistema de Red 5G+/6G y elementos asociados

Este subsistema representa un elemento clave en la infraestructura descrita, pues integra diversos componentes que permiten la implementación de una red 5G+/6G totalmente privada y autónoma, así como los medios para la realización de emulaciones y mediciones experimentales en escenarios de escala real y controlada.

1) Red de radio 5G+/6G: Como parte del proyecto se ha desplegado una infraestructura 5G autónoma o *standalone* (SA) redundada con tecnología Nokia, compuesta por dos núcleos de red ubicados en los campus CN y CBL, estaciones de radio 5G operando en las bandas n40 y n258, más de 100 tarjetas SIM, y software de gestión de Nokia con funciones avanzadas de orquestación SDN, interfaces estandarizadas y mecanismos de control seguro.

Se han instalado, en cada campus, un total de tres unidades radio en la banda n40 y dos unidades radio de bandas milimétricas. La ubicación de dichas unidades radio con sus correspondientes antenas se ha seleccionado estratégicamente para favorecer diversos entornos de investigación. La Fig. 5 muestra la ubicación de las unidades radio en cada uno de los campus.

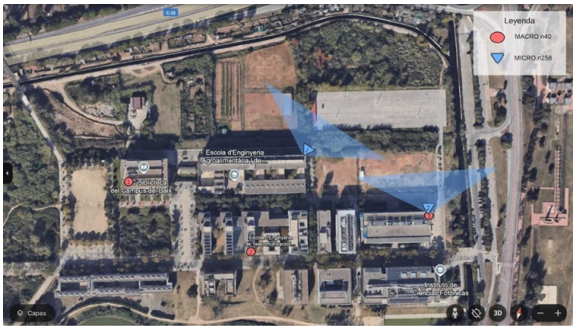
Cada CPD cuenta con una unidad de banda base (BBU, *Baseband Unit*) que se encuentra conectada a las unidades de radio de su respectivo campus. Cada BBU se compone de dos bloques: el bloque rack y el bloque de tarjetas de capacidad. El bloque rack actúa como concentrador de las tarjetas de capacidad y como punto de agregación del tráfico entre las radios y el núcleo de red, en ambas direcciones. Las tarjetas de capacidad se encargan de las funciones de enrutamiento y conmutación. Además, la arquitectura implementada permite una futura expansión mediante la incorporación de nuevas tarjetas de capacidad.

El sistema de red móvil 5G+/6G está equipado con un sistema de sincronización independiente para garantizar un preciso sincronismo de la señal en toda la infraestructura, el cual resulta de vital importancia para la configuración del núcleo 5G (*core*) redundante que se ha implementado en este subsistema.

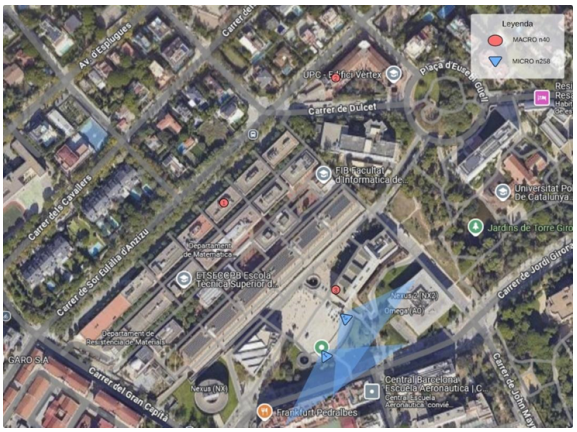
El núcleo 5G implementado en esta arquitectura opera en modo SA, lo que implica que funciona de manera totalmente independiente, sin depender de redes 4G (LTE, *Long-Term Evolution*), y sobre una arquitectura basada en servicios (SBA, *Service-Based Architecture)*. Este *core*, del mode-lo Nokia Compact Mobility Unit (CMU) 24.2 [20], también ofrece capacidades avanzadas como la creación y gestión de *network slices*, es decir, la posibilidad de crear múltiples redes virtuales (lógicas) sobre una misma infraestructura física, adaptadas a distintos casos de uso o requisitos específicos de servicio.

Cada campus posee una instancia de *core* 5G, interconectadas mediante la red de transporte óptica con un canal óptico de 50 Gbps reservado exclusivamente para

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)



(**a**)



(**b**)

Fig. 5. Mapa con las ubicaciones de las antenas de la red radio 5G+/6G en CBL (**a**) y CN (**b**).

este servicio. Esta configuración garantiza que los terminales móviles que se conectan en cada campus forman parte de una misma red, lo que optimiza la gestión de la conectividad y asegura una experiencia de usuario óptima. Además, proporciona resiliencia y robustez al sistema. Por ejemplo, en caso de que una de las instancias de *core* 5G dejase de funcionar, la otra estaría capacitada para gestionar la red móvil en su totalidad. De la misma manera, si el enlace que interconecta las instancias del *core* fallara, cada instancia estaría preparada para trabajar de manera independiente con los elementos que tiene directamente conectados.

2) Emulador LTE/5G+ Narrow Band IoT: Se cuenta con un equipo de emulación de redes LTE/5G+ destinado a la investigación en escenarios controlados, que es independiente del subsistema de red móvil 5G+/6G.

El emulador instalado, denominado *AMARI Callbox Advanced* [21], posibilita la realización de pruebas extremo a extremo de una red celular, el testeo de diferentes elementos de red y la validación del desarrollo de módulos de arquitectura celular basada en interfaces estándar. Además, extiende sus capacidades hasta la banda de milimétricas con la incorporación de un paquete de *hardware* dedicado.

3) Sistema de software radio: Para llevar a cabo casos de uso relacionados con el paradigma de radio definida por software (SDR, *Software-Defined Radio*), se cuenta con diversas plataformas que están diseñadas para aplicaciones que requieran alto rendimiento a nivel de procesamiento y flexibilidad en el desarrollo de sistemas 5G+/6G.

Estas unidades, modelos Ettus USRP X410 y Ettus USRP X440 [22], [23], son plataformas de vanguardia para la emulación de redes 5G+/6G mediante el despliegue de arquitecturas abiertas como srsRAN y *Open Air Interface* [24], [25]. Además, resultan particularmente idóneas para diseñar y evaluar estrategias de *beamforming* en sistemas de antenas masivas (*Massive MIMO*) y para el desarrollo de algoritmos de predistorsión digital (DPD, *Digital Pre-Distorsion*) destinados a optimizar el rendimiento de transmisores de alta eficiencia.

Estas plataformas, en combinación con estaciones de trabajo, permiten la realización de experimentos que abarcan desde la capa física hasta las capas superiores de la red, incluyendo el control de recursos radio, la agregación de portadoras, el *beamforming* adaptativo y la evaluación de la coexistencia de diferentes tecnologías en entornos complejos.

4) Kits de evaluación 5G RedCap: La plataforma cuenta con dos kits de evaluación (Quectel 5GM2EVB-KIT [26]), junto con sus respectivos módulos de dispositivo IoT 5G *Reduced Capabilities* (RedCap). Estos dispositivos se comportan como usuarios terminales (UEs, *User Equipments*) y tienen la capacidad de conectarse a redes móviles 5G con soporte de RedCap.

La tarjeta de evaluación constituye una herramienta que facilita el desarrollo de aplicaciones y la evaluación de funcionalidades básicas para los módulos soportados. El módulo RedCap es compatible con una amplia gama de protocolos de Internet e interfaces estandarizadas.

5) Generadores de señales y analizadores 5G+/6G: Con el objetivo de abrir líneas de investigación relacionadas con el análisis y la mejora de los sistemas 5G actuales y de los futuros sistemas 6G, se ha adquirido una serie de equipos que han implementado una plataforma para linealizar amplificadores de radiofrecuencia.

La plataforma interconecta varios dispositivos:

- Un osciloscopio-digitalizador de alta velocidad del modelo UXR010B de Keysight Technologies capaz de digitalizar señales de hasta 33 GHz en un ancho de banda de 5 GHz [27].
- Un generador arbitrario de alta velocidad del modelo M8190A de Keysight Technologies que es capaz de generar digitalmente señales banda base y pasobanda (IF/RF) hasta 5 GHz [28].
- Convertidores de frecuencia UD Box 5G de Tmytek que permiten convertir señales del rango de frecuencias 1 de 5G (FR1) al rango 2 (FR2) y viceversa [29].
- Conformadores de haz del modelo BBox One 5G de Tmytek capaces de transmitir y recibir inalámbricamente las señales de FR2 [30].

C. Subsistema de computación en el extremo de la red

El subsistema de computación en el extremo (*Edge Computing*) consiste en el despliegue de 24 nodos distribuidos estratégicamente en ambos campus, formando

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

una red que permite implementar casos de uso y verticales como los vehículos autónomos. Estas unidades multitecnología actúan como estaciones base o pico celdas, dotando de una cobertura casi total a las calles del campus. La Fig. 6 muestra la disposición de los nodos en el CBL.

Fig. 6. Disposición de los nodos de Edge Computing en el CBL

El dispositivo utilizado en el despliegue de cada nodo es el Kit de desarrollo NVIDIA Jetson AGX ORIN (64 GB) [31]. Estos dispositivos poseen gran capacidad de cómputo, con 64 GB de RAM y una GPU de arquitectura Ampere. Además, cada nodo es capaz de proporcionar conectividad con las tecnologías de 5G y de WiFi 6E. En el caso de 5G, los dispositivos incluyen un *dongle* USB con una tarjeta SIM para ejercer como clientes de la red 5G+/6G desplegada. Cada nodo del subsistema se conecta a la red fija mediante enlaces por cable Ethernet o radioenlaces inalámbricos para comunicarse con los restantes subsistemas de la plataforma.

D. Subsistema de red sensible al tiempo

Dada a su importancia en el ámbito de la Industria 4.0 y en las futuras aplicaciones de 6G, se ha desplegado una red sensible al tiempo (TSN, *Time-Sensitive Network*), que opera paralelamente a los demás subsistemas de red de la plataforma. Específicamente, la red cuenta con conmutadores modelo RELY-TSN4 de Soc-E [32] en un laboratorio de investigación del CBL, así como en el CPD y en el DronLab. Este subsistema ofrece servicios de red fiables y robustos en diversos entornos utilizándose como banco de pruebas de servicios y prototipos. De igual modo, se dispone de una serie de tarjetas de red que son compatibles con los estándares de TSN. Esto permite la incorporación de ordenadores de sobremesa a esta red, la cual cuenta con una inserción en los planos de datos y gestión de la red fija. A través de estos, es posible establecer una comunicación con los restantes subsistemas.

E. Subsistema de sensórica, robótica y agricultura inteligente

En Agrópolis, instalación de la UPC cercana al CBL, se ha desplegado una red de sensores para poder ejecutar diferentes medidas del entorno agrícola. Estos nodos están conectados con la red troncal a través del radioenlace de microondas previamente descrito.

La red de sensores está formada por diferentes dispositivos y tecnologías de comunicación. Consta de la pasarela IoT inalámbrica de bajo consumo Libelium One [33], con sensores de temperatura y humedad de aire, radiación solar, conductividad y humedad de suelo y hoja.

Adicionalmente, el subsistema cuenta con el robot móvil autónomo RB-SUMMIT [34], que está diseñado para desarrollar una amplia gama de aplicaciones de I+D en entornos interiores y especialmente exteriores. Su versatilidad, tamaño y alta movilidad hacen del RB-SUMMIT un recurso valioso para una amplia gama de aplicaciones, incluyendo la logística, el transporte y principalmente la agricultura, con una capacidad de carga máxima de 50 kg. Esta plataforma móvil multipropósito, robusta y modular está diseñada para navegar de manera autónoma o ser teleoperada mediante una cámara *Pan-Tilt-Zoom* que transmite el vídeo en tiempo real. RB-SUMMIT utiliza una arquitectura de control abierta y modular basada en ROS 2, lo que permite su programación abierta y el desarrollo de nuevas aplicaciones y formas de control más allá del software de control y programación disponible.

IV. DESCRIPCIÓN DE ALGUNOS SERVICIOS Y CASOS DE USO

La infraestructura integra múltiples elementos y servicios basados en tecnologías heterogéneas. Cada uno de estos servicios estará disponible tanto para la comunidad investigadora, así como para empresas privadas y administraciones públicas interesadas en experimentar, validar o desarrollar soluciones sobre estas tecnologías emergentes.

El acceso a la infraestructura se controla mediante redes virtuales privadas (VPN, *Virtual Private Network*) y servidores de autenticación, autorización y contabilidad (AAA, *Authentication, Authorization y Accounting*), aplicando en cada caso la tarifa correspondiente definida por la UPC.

En la presente sección se exponen algunos de los servicios ofertados en la plataforma, incluyendo sus posibles casos de uso y aplicaciones, los dispositivos y componentes involucrados.

A. Subsistema de red fija y servidores de cómputo

1) Red de transporte óptica: Además de ser utilizada por la propia plataforma, la red de transporte óptica ofrece servicios tanto a agentes externos como internos de la propia universidad.

Entre los servicios que ofrece esta red se encuentra la inyección de tráfico de distintas tecnologías. Este proceso permite multiplexar de manera transparente flujos de tráfico de diversas tecnologías en una misma trama OTN, compartiendo un enlace físico común. Por ejemplo, es posible transportar servicios de diversas tecnologías (Ethernet, SDH, entre otros) desde CBL a CN y viceversa, así como realizar la demultiplexación de estos en distintos nodos de la red óptica.

2) Clúster de Kubernetes: El clúster de Kubernetes ubicado en el CPD de CBL ofrece una extensa gama de servicios, que incluyen desde la virtualización de recursos hasta la implementación de aplicaciones de diversa naturaleza mediante contenedores.

Se brinda el servicio de despliegue e inferencia de modelos de ML, aprovechando la escalabilidad de Kubernetes y su red interna para desplegar modelos ML pre-

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

entrenados. Estos son utilizados como APIs en un entorno de contenedores, garantizando una implementación eficiente y segura.

Asimismo, en lo que respecta al entrenamiento de modelos de ML, se ofrecen servicios de aprendizaje distribuido con Kubeflow, una plataforma de ML diseñada para Kubernetes, con el fin de llevar a cabo la formación distribuida de modelos de ML utilizando marcos como TensorFlow, PyTorch o MXNet. Kubeflow facilita la ejecución de extensos conjuntos de datos y la escalabilidad a través de múltiples contenedores.

Además, en combinación con el subsistema de *edge computing*, se ofrece la posibilidad de agregar nodos de *edge computing* como *workers* en el clúster de Kubernetes, tal como se muestra en el esquema de la Fig. 7.

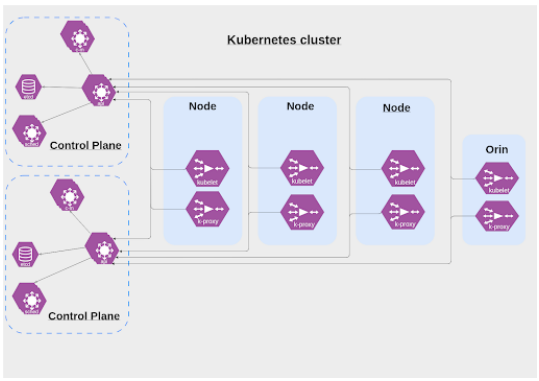


Fig. 7. Visualización del entorno con un nodo de *edge computing* como nodo adicional.

De esta manera, se amplían los dominios del clúster de Kubernetes hasta el subsistema de edge computing para aprovechar la flexibilidad y la robustez de Kubernetes y utilizar la potencia de cálculo de los nodos Jetson AGX Orin en aplicaciones de vehículos autónomos en combinación con algoritmos de aprendizaje federado o distribuido. Kubernetes ofrece ventajas en la automatización y el escalado de contenedores, lo que permite implementar técnicas de migración o traspaso (*handover*) de aplicaciones entre nodos de *edge computing*.

B. Subsistema de red 5G+/6G

Este subsistema unificado y redundante entre campus abre la puerta a un amplio abanico de casos de uso experimentales y servicios que podrían desplegarse.

Dado el alto nivel de interconexión de los subsistemas que constituyen la plataforma, la red 5G+/6G proporciona servicios de monitoreo y control remoto de drones y vehículos autónomos. En este caso, y gracias a la baja latencia que proporciona la infraestructura, es posible controlar un dron haciendo uso de la red móvil, desde cualquier punto de la infraestructura.

De igual manera, es posible conducir o monitorear vehículos autónomos mediante la utilización de la red de *edge computing*. Estos nodos, además de actuar como clientes de la red móvil, cuentan con algoritmos de *handover* vertical entre las tecnologías de radio que soportan.

La red móvil también ofrece servicios de *network slicing*, mediante la configuración de *network slices* lógicos de extremo a extremo. Esta implementación se lleva a cabo en combinación con aplicaciones basadas en ML que se despliegan en el clúster de Kubernetes y que, a su vez, extraen parámetros radio en tiempo real desde la plataforma de gestión de la red para ejecutar acciones de configuración con el objetivo de crear los *network slices* y mantener la calidad del servicio durante su uso.

C. Sistema de computación en el extremo de la red

Las unidades desplegadas en este subsistema actúan como nodos de comunicación y cómputo con GPU en el extremo de la red, lo que permite su utilización en aplicaciones "críticas" que requieren baja latencia y gran ancho de banda, como una comunicación ultrarrápida de baja latencia (uRLLC, *Ultra-High Reliability & Low Latency*). Esto abre paso a la implementación de aplicaciones de exteriores en el entorno del campus, interiores e híbridas con traspasos de los modos de comunicación, tanto verticales como horizontales.

Este entorno es apto para desarrollar casos de uso de conducción asistida, descarga inteligente de contenidos multimedia de gran ancho de banda en tiempo real, emplazamiento dinámico de los recursos en el extremo de la red, comunicaciones vehículo-vehículo-infraestructura (V2V-V2I), *network slicing* de acceso, o la gestión automatizada de la red, entre otros.

Además, se ofrece el servicio de automatización de aplicaciones basadas en algoritmos de ML que se diseñen, desarrollen y validen sobre la propia infraestructura:

- En modo distribuido, donde cooperen todos los agentes ejecutándose cada uno de ellos en una unidad.
- En modo federado, donde un agente desplegado en el clúster de Kubernetes intercambie información con el resto de los agentes localizados en cada unidad.
- En modo centralizado, ejecutándose el modelo de ML desplegado en el clúster de Kubernetes.

V. CONCLUSIONES

La plataforma constituida a partir de los proyectos 6G-OpenLab y ELEGANT de la UPC representa un paso significativo hacia la creación de una infraestructura de investigación abierta, avanzada y flexible, destinada al estudio, validación y despliegue de tecnologías emergentes en el contexto de las redes móviles 5G+ y 6G. El diseño distribuido de la plataforma, que abarca dos campus interconectados mediante enlaces ópticos de alta capacidad y baja latencia, permite combinar entornos reales y controlados, habilitando una gran diversidad de casos de uso y aplicaciones experimentales.

La infraestructura se ha construido sobre una arquitectura robusta y escalable, organizada en planos funcionales bien diferenciados, con una clara separación entre el plano de datos y el plano de gestión. Esta organización permite una administración eficiente de los recursos físicos y virtuales, así como el despliegue automatizado de servicios mediante técnicas de virtualización, redes definidas por software y aprendizaje automático.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Entre sus principales aportaciones se destacan:

- La integración de redes móviles privadas 5G+/6G redundadas y configurables mediante *network slicing*.
- El despliegue de plataformas heterogéneas de cómputo (Openstack, Kubernetes, HPC, *edge computing*) orientadas al soporte de aplicaciones distribuidas y ML.
- La integración de tecnologías de redes ópticas, radioenlaces de microondas, redes TSN, y plataformas de sensórica y robótica orientadas a la agricultura inteligente.
- Un enfoque modular y abierto que permite incorporar nuevas tecnologías, extender la infraestructura y facilitar la colaboración con la industria, administración y otras instituciones de investigación.

AGRADECIMIENTOS

Este trabajo ha sido realizado en el proyecto PID2022-137329OB-C41 financiado por MCIN/AEI/10.13039/501100011033 y FEDER, y en el marco de los proyectos UNICO I+D 6G, Subprograma de infraestructuras de investigación y equipamiento científico-técnico, Convocatoria 2022 TSI-064100-2022-006 y Convocatoria 2023 TSI-064100-2023-0009, financiados por el Ministerio de Asuntos Económicos y Transformación Digital de España y por el Fondo de Recuperación, Tranformación y Resilencia de la Unión Europea - NextGenerationEU.

REFERENCIAS

[1] 6G-OpenLab - Laboratori Obert Científic-Tecnològic d'Investigació en 6G de la UPC. Accedido: Jun. 13, 2025. [Online]. Disponible: https://6g-openlab.upc.edu/en

[2] ELEGANT, Evolution of infrastructure (ultra low Latency) and Expanded mobile system (mmWave) towards 6G-AI NeTworks. Accedido: Jun. 13, 2025. [Online]. Disponible: https://elegant.upc.edu/en

[3] "H3C S9855 DataCenter Switches," H3C. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.h3c.com/en/Products_and_Solutions/InterConnect/Switches/Products/Data_Center/Aggregation/S9800/H3C_S9855/

[4] "Switch de chasis de centro de datos, NC8200-4TD," FS.com. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.fs.com/es/products/110482.html

[5] "Switch 10G S5860-48SC | 48x 10Gb SFP+, 8x 100Gb QSFP28, apilable," FS.com. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.fs.com/es/products/115385.html

[6] "1830 PSS/PSI-L Optical Line Systems," Nokia.com. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nokia.com/optical-networks/1830-pss-psi-l-optical-line-systems/

[7] "1830 Photonic Service Switch (PSS)," Nokia.com. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nokia.com/optical-networks/1830-photonic-service-switch/

[8] Ubiquiti Store Europe, "airFiber 60 Xtreme-Range," Accedido: Jun. 13, 2025. [Online]. Disponible: https://eu.store.ui.com/eu/en/products/af60-xr

[9] Ministerio de Asuntos Económicos y Transformación Digital, "Orden ETD/1449/2021, de 16 de diciembre, por la que se aprueba el Cuadro Nacional de Atribución de Frecuencias", vol. BOE-A-2021-21346. 2021, pp. 162420–162777. Accedido: Sep. 03, 2025. [Online]. Disponible: https://www.boe.es/eli/es/o/2021/12/16/etd1449

[10] ETSI EN 302 567, "Multiple-Gigabit/s radio equipment operating in the 60 GHz band; Harmonised Standard for access to radio spectrum", V2.2.1. 2021. Accedido: Sep. 03, 2025. [Online]. Disponible: https://www.etsi.org/deliver/etsi_en/302500_302599/302567/02.02.01_30/en_302567v020201v.pdf

[11] "Open Source Cloud Computing Infrast. - OpenStack." Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.openstack.org/

[12] "Dell EMC PowerEdge R6525 Especificaciones técnicas," Dell. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.dell.com/support/manuals/en-us/poweredge-r6525/r6525_ts_pub/especificaciones-t%C3%A9cnicas?guid=guid-ecb5cfaa-cf2b-46f6-ab1e-1839d086bc6a&lang=es-mx

[13] "Dell EMC PowerEdge R440 Especificaciones técnicas," Dell. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.dell.com/support/manuals/es-es/poweredge-r440/per440_ts_pub/especificaciones-t%C3%A9cnicas?guid=guid-eed6f152-01c7-4dc2-89e6-d9c6ae9eb785&lang=es-mx

[14] "Orquestación de contenedores para producción," Kubernetes. Accedido: Jun. 13, 2025. [Online]. Disponible: https://kubernetes.io/es/

[15] "Dell EMC PowerEdge R740xd Especificaciones técnicas," Dell. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.dell.com/support/manuals/es-es/poweredge-r740xd/per740xd_techspecs_pub/especificaciones-t%C3%A9cnicas?guid=guid-538fea06-fb40-4286-850b-52bd2df0d5ad&lang=es-mx

[16] "Las GPU NVIDIA A100 impulsan el centro de datos moderno," NVIDIA. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nvidia.com/es-es/data-center/a100/

[17] "R283-SF1-AAL1 | Rack-Server - GIGABYTE Spain," GIGABYTE. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.gigabyte.com/es/Enterprise/Rack-Server/R283-SF1-AAL1

[18] "GPU NVIDIA H200 Tensor Core," NVIDIA. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nvidia.com/es-es/data-center/h200/

[19] "GPU NVIDIA H100 con núcleo Tensor," NVIDIA. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nvidia.com/es-es/data-center/h100/

[20] "Compact Mobility Unit," Nokia. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.dac.nokia.com/connectivity-solutions/compact-mobility-unit-cmu/

[21] "AMARI Callbox Advanced." Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.amarisoft.com/test-and-measurement/device-testing/device-products/amari-callbox-advanced

[22] "NI Ettus USRP X410," Ettus Research. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.ettus.com/all-products/usrp-x410/

[23] "NI Ettus USRP X440," Ettus Research. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.ettus.com/all-products/usrp-x440/

[24] "srsRAN Project - Open Source RAN." Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.srsran.com/

[25] "OpenAirInterface – 5G software alliance for democratising wireless innovation." Accedido: Jun. 13, 2025. [Online]. Disponible: https://openairinterface.org/

[26] "5G-M2 EVB kit," Quectel. Accedido: Jun. 13, 2025. [Online]. Disponible:https://www.quectel.com/product/5g-m2-evb-kit/

[27] "Infiniium UXR-B Series Oscilloscopes | High-Precision Measurements," Keysight. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.keysight.com/us/en/products/oscilloscopes/infiniium-real-time-oscilloscopes/infiniium-uxr-b-series-oscilloscopes.html

[28] "M8190A 12 GSa/s Arbitrary Waveform Generator," Keysight. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.keysight.com/us/en/product/M8190A/12-gsa-s-arbitrary-waveform-generator.html

[29] "UD Box 5G | Up/down converter, frequency converter | Up to 44 GHz | TMYTEK," Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.tmytek.com//products/frequency-converters/udbox5g

[30] "BBox 5G | mmWave Beamformer for Antenna & Protocol Designers | TMYTEK." Accedido: Jun. 13, 2025. [Online]. Disponible: https://tmytek.com/products/beamformers/bbox

[31] "NVIDIA Jetson AGX Orin," NVIDIA. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.nvidia.com/es-es/autonomous-machines/embedded-systems/jetson-orin/

[32] "RELY-TSN4,"SOC-E. Accedido: Jun. 13, 2025. [Online]. Disponible: https://soc-e.com/products/rely-tsn4/

[33] "Libelium One," Libelium. Accedido: Jun. 13, 2025. [Online]. Disponible: https://www.libelium.com/es/productos-iot/libelium-one/

[34] "RB-SUMMIT: Robot para I+D en interiores y exteriores," Robotnik. Accedido: Jun. 13, 2025. [Online]. Disponible: https://robotnik.eu/es/productos/robots-moviles/rb-summit/

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Despliegue de actualizaciones y comunicaciones seguras en redes IoT

Alfonso Sarró, Juan de la Sotilla, Laura Rami, Arnau Cermeron, Nil de Creus, Alan Briones, Julia Sánchez, Guiomar Corral
Blended Network Architectures (BNA) Research Line

La Salle – Universidad Ramon Llull

08022

alfonso.sarro@salle.url.edu, juan.delasotilla@salle.url.edu, laura.rami@salle.url.edu, arnau.cermeron@students.salle.url.edu, nil.decreus@students.salle.url.edu, alan.briones@salle.url.edu, j.sanchez@salle.url.edu, guiomar.corral@salle.url.edu

La extensa utilización de los sistemas *Internet of Things* en la industria implica que es necesario tener un sistema robusto y seguro que proteja los datos y las comunicaciones que transmite. Para lograrlo, este artículo propone una arquitectura centralizada segura para la gestión remota de dispositivos, orientada a entornos donde la integridad y confidencialidad de los datos son prioritarias, como en aplicaciones industriales o médicas. La solución se basa en un servidor central que orquesta la comunicación y actualización de múltiples dispositivos finales, empleando Mender como plataforma para la distribución de firmware y configuraciones. Para proteger las comunicaciones, se implementa una red privada virtual basada en WireGuard, y se establece un sistema de autenticación mediante certificados digitales generados con OpenSSL. El sistema garantiza actualizaciones seguras, escalabilidad y capacidad de recuperación ante fallos. Esta propuesta se valida mediante un caso de uso funcional que demuestra la integración efectiva de todos los componentes.

***Palabras Clave*- IoT, ciberseguridad, certificados, actualizaciones centralizadas, VPN**

I. Introducción

El despliegue seguro y eficiente de software en entornos de dispositivos conectados es uno de los principales desafíos en la evolución del *Internet of Things* (IoT).[1] Con el aumento del uso de arquitecturas que integran desde sensores industriales hasta dispositivos médicos (IoMT, Internet of Medical Things), la necesidad de mantener estos sistemas actualizados, operativos y protegidos frente a amenazas externas es más crítica que nunca. En este contexto, la automatización del ciclo de vida de los dispositivos, junto con la aplicación de medidas criptográficas de protección, se convierte en un componente esencial para garantizar la robustez y fiabilidad de la infraestructura.

Este documento presenta una propuesta de arquitectura basada en un servidor central que actúa como nodo coordinador, permitiendo la administración remota de múltiples dispositivos finales. La solución planteada aborda tres pilares fundamentales: la distribución centralizada de actualizaciones de firmware, la protección de las comunicaciones mediante VPN cifradas con WireGuard y la autenticación de dispositivos mediante certificados digitales generados con OpenSSL. A lo largo del documento se describen en detalle las tecnologías utilizadas, los procedimientos de configuración de los distintos componentes, y el flujo de trabajo necesario para establecer una infraestructura segura, escalable y automatizable.

En primer lugar, se introduce el contexto tecnológico y se analizan los requisitos de seguridad específicos en redes IoT e IoMT. A continuación, se describen los elementos clave de la arquitectura propuesta, incluyendo la elección de Mender como plataforma de actualización remota, las opciones de servidores web y servicios VPN disponibles, y la generación y gestión de certificados digitales.

Finalmente, se expone un caso de uso práctico donde se integran todos los componentes descritos, evidenciando la viabilidad del sistema propuesto en escenarios reales.

II. Background

Los sistemas IoT son el conjunto de dispositivos físicos interconectados que tienen la capacidad de recopilar, procesar y compartir datos a través de la red. Estos dispositivos, normalmente equipados con mecanismos de toma de datos, como sensores o programas específicos, operan de manera autónoma, colaborando entre ellos, con el objetivo de automatizar procesos, incrementar la eficiencia operativa y habilitar nuevas funcionalidades sin la necesidad de intervención humana constante.[2]

Este sistema se ha extendido a múltiples sectores, incluyendo el ámbito industrial, doméstico y urbano. Una de sus derivaciones más relevantes es el Internet de las Cosas Médicas (IoMT), que engloba aquellos dispositivos IoT

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

diseñados específicamente para el entorno médico. En este contexto, los dispositivos recopilan, almacenan y transmiten información clínica sensible relativa al estado de salud de los pacientes, lo cual exige el cumplimiento de estrictos criterios de confidencialidad, integridad y disponibilidad de los datos.[3]

La gestión de estos dispositivos requiere de mecanismos que garanticen tanto la autenticación mutua como la comunicación cifrada entre dispositivos, evitando así posibles accesos no autorizados o ataques de interceptación de datos.

Por otro lado, también es necesario contar con un sistema de gestión centralizada de actualizaciones, que permita aplicar de forma remota parches de seguridad, mejoras funcionales o cambios en la configuración de los dispositivos desplegados. Este enfoque reduce significativamente la exposición a vulnerabilidades y facilita el cumplimiento de políticas de seguridad dinámicas en tiempo real.

Una actualización centralizada hace referencia a un enfoque específico en el que las actualizaciones de software se aplican desde un único punto de control, en lugar de requerir instalaciones manuales en cada equipo. Cuando se lanza una nueva versión o parche, éste se implementa desde el servidor central y los dispositivos cliente lo descargan automáticamente desde dicho servidor. Este enfoque permite lograr una alta eficiencia, uniformidad y control en redes corporativas o de dispositivos distribuidos.

Una de las opciones para llevar a cabo estas actualizaciones es realizando una actualización completa del firmware. Este sistema implica reemplazar completamente la imagen del software o sistema operativo de cada uno de los dispositivos finales. Es normalmente utilizado en entornos homogéneos y que requieren alta confiabilidad, y en el que el hardware y software están muy estandarizados.[4][5]

Otra de las opciones es realizar actualizaciones modulares. En este sistema solo se actualizan ciertos módulos del software, como un protocolo de red o una aplicación específica. Este enfoque está ganando popularidad, ya que tiene un menor impacto energético y de red, y permite realizar actualizaciones más pequeñas y frecuentes.[4]

Un tercer sistema posible es la actualización bajo demanda o por acceso remoto. En este sistema, el software actualizado reside en el servidor central y se descarga según la necesidad. Sin embargo, este sistema no permite actualizaciones de Sistema Operativo, solo de aplicaciones y software de cliente. Este sistema se utiliza en infraestructura de acceso público y en aplicaciones en la nube o escritorios virtuales.[4]

Por último, otra opción es llevar a cabo actualizaciones programadas y automatizadas. Este enfoque se basa en políticas predefinidas y distribuye automáticamente las actualizaciones desde el servidor central en base a esas políticas. Este enfoque se suele utilizar en empresas con miles de estaciones de trabajo homogéneas.[4]

En la Tabla I se muestra las diferentes características de los diferentes tipos de actualizaciones mencionados anteriormente.

Tabla I

COMPARACIÓN ENTRE TIPOS DE ACTUALIZACIONES

Tipo de actualización	Alcance	Consumo de red	Personaliza-ción	Requiere reinicio
Completa	Total	Alto	Muy alta	Sí
Modular	Selectivo	Bajo	Alta	No
Bajo demanda	Centralizado	Medio	Media	No
Programada	Variable	Medio	Baja	No

En este artículo se plantea un sistema centralizado y seguro en el que un dispositivo central se comunica periódicamente con múltiples dispositivos finales. Para poder tener un mejor control sobre los dispositivos finales, se opta por realizar actualizaciones completas donde la imagen entera es actualizada siempre que se distribuya una actualización. Esta arquitectura es común en entornos como el IoT industrial, infraestructura crítica, o dispositivos médicos remotos, donde mantener la integridad, autenticidad y privacidad de los datos, así como la gestión segura del ciclo de vida de los dispositivos, es indispensable.

La topología planteada es una topología en estrella, donde un servidor central actúa como nodo coordinador, y es responsable de recibir los datos de los dispositivos finales, enviar actualizaciones y gestionar la seguridad del entorno.

III. TECNOLOGÍAS UTILIZADAS

A. Servidor web

En los dispositivos finales se ofrecen varios servicios para los clientes, que terminan generando una gran cantidad de datos. Estos datos pueden ser tanto medidas tomadas por sensores como fotografías o modelos 3D generados en los dispositivos finales. Estos dispositivos han de ser capaces de enviar toda la información al servidor central, donde se almacenará en la base de datos interna.

Para poder llevar a cabo la transmisión de información, los protocolos como MQTT, que son ampliamente utilizados en el entorno IoT, no permitirían el envío de archivos grandes como fotografías o modelos 3D, así que sería necesario utilizar otro protocolo distinto. El protocolo más habitual para la transferencia de ficheros es SFTP, pero en este entorno no sería el óptimo. Esto es debido a que los dispositivos finales, una vez suben los datos al servidor central, eliminan los datos almacenados localmente, permitiendo así que tengan un almacenamiento reducido. Cuando se necesite acceder a esta información, el objetivo es que se pueda visualizar de manera remota y que no sea necesario volver a transmitir los archivos del servidor central al dispositivo para reducir el estrés en las máquinas y la congestión de la red. Para poder llevar a cabo esta interacción, se ha utilizado el protocolo HTTPS que, aún siendo menos eficiente transmitiendo ficheros, permite realizar ambas funcionalidades a través del mismo protocolo. De esta manera hay menos servicios activos y, por lo tanto, un ecosistema más seguro con menos posibles puntos de entrada.

Uno de los posibles servidores a utilizar es Apache, caracterizado por ser uno de los servidores web más utilizados. Es un servidor flexible, con soporte para múltiples módulos, e ideal para aplicaciones complejas.[6]

Otra opción es Nginx, un servidor ligero y eficiente en la manipulación de múltiples conexiones concurrentes. Es mayoritariamente utilizado por su gran rendimiento y alta escabilidad.[7]

Una alternativa para entornos con recursos limitados es Lighttpd, un servidor rápido y seguro optimizado para tener un alto rendimiento en cualquier dispositivo.

Por último, Caddy es un servidor moderno con HTTPS automático por defecto, utilizando *Let's Encrypt* para generar certificados de manera simple y automática.[8]

En este caso, se ha utilizado Nginx, ya que tiene una mejor escalabilidad y rendimiento, de manera que si el sistema acaba incorporando grandes cantidades de dispositivos finales, será capaz de utilizar múltiples *threads* y mantener su gran rendimiento. Toda la gestión de certificados se realizará a través del servidor central.

B. Servicio de actualizaciones remotas

Para una buena seguridad y un buen funcionamiento del sistema, es necesario que todos los dispositivos estén siempre actualizados a la última versión. Estas actualizaciones han de ser tanto de los servicios que proporciona el servidor como del Sistema Operativo y el software base de éste. De esta manera, se pueden eliminar la gran mayoría de las vulnerabilidades, debido a que las empresas productoras del software respectivo suelen proporcionar las medidas de seguridad necesarias a través de sus actualizaciones.

Para poder actualizar estos dispositivos, se puede realizar de dos maneras distintas: actualizaciones *pull-based* y actualizaciones *push-based*.

Las actualizaciones *pull-based* son gestionadas por los propios dispositivos finales. Cuando el dispositivo detecta que hay una actualización disponible, el propio dispositivo inicia la actualización. La ventaja de este tipo de actualizaciones es que suelen ser detectadas e instaladas más rápidamente, pero en el caso de que no se detecte correctamente, habrá dispositivos en la red que están funcionando en una versión mientras que otros dispositivos estarán funcionando con otra.[5]

Las actualizaciones *push-based Over The Air (OTA)* nos permiten lidiar con este último problema. Este método de actualización funciona a través de un servidor central que descarga la actualización y la envía a todos los dispositivos conectados. De esta manera, se puede asegurar que todos los dispositivos conectados reciben dicha actualización.[5]

Uno de los servicios que implementa este último tipo de actualizaciones es Mender. Mender ofrece firma criptográfica de los archivos de actualización, garantizando la procedencia de una fuente de confianza. Las actualizaciones pueden realizarse de forma progresiva por fases, ajustando que agrupaciones de dispositivos se actualizarán asignándolos al grupo *canary*. Este sistema también implementa un sistema de *rollback*, permitiendo la reversión automática a la versión anterior en caso de una actualización fallida. Este sistema permite crear distintos tipos de actualizaciones, adaptadas a las necesidades específicas del dispositivo, y mantiene un registro de qué se actualiza, cuándo y en qué dispositivo. Mender está diseñado para grandes grupos de dispositivos o servidores que ejecutan Linux.

Otro servicio que implementa este tipo de actualizaciones es el de Amazon Web Services (AWS), concretamente AWS IoT OTA[9], el cual está estrechamente integrado con el núcleo de AWS IoT y aprovecha su modelo de seguridad, incluida la administración del acceso, la autenticación de dispositivos y el cifrado de datos. Esto la convierte en una buena solución para implementaciones en las que los dispositivos se administran a través de AWS. Sin embargo, esto también disminuye su utilidad en entornos no AWS, introduciendo una portabilidad limitada a través de diferentes infraestructuras.

BalenaCloud[10] también ofrece un sistema OTA, basado en contenedores con Docker. Proporciona una interfaz web para gestionar dispositivos, versiones de aplicaciones y estados de actualización. Sin embargo, para casos de uso como la distribución de actualizaciones en máquinas dedicadas, puede resultar excesivo al tener todo en contenedores, lo que lo hace innecesario y sobredimensionado.

SWUpdate[11] es un marco ligero y versátil utilizado habitualmente en sistemas Linux industriales. Soporta una amplia variedad de formatos de actualización y es altamente personalizable mediante scripts. Su diseño modular permite la integración con flujos de trabajo personalizado. Sin embargo, no ofrece un panel de gestión. Como resultado, la implementación de un flujo de trabajo completo requiere un trabajo extra considerable.

Robust Auto-Update Controller (RAUC)[12] es un sistema OTA centrado en la seguridad y la integridad, ampliamente adoptado por los sectores de la automoción y la industria. Hace hincapié en las actualizaciones de particiones A/B, unas particiones duales que permiten actualizaciones sin interrupción del sistema. También se centra en la verificación de la integridad de los archivos mediante hashes y la conmutación por error en caso de fallo de una actualización. RAUC se integra bien con herramientas como OSTree[13] para la gestión del estado del sistema y la coherencia del sistema de archivos. Sin embargo, tiene una curva de aprendizaje pronunciada en comparación con otros sistemas OTA y carece de utilidades de gestión de alto nivel.[14]

Teniendo en cuenta las opciones anteriores, la elección de Mender como plataforma OTA refleja varias ventajas operativas y de seguridad que se alinean con los objetivos de la gestión centralizada de actualizaciones. En primer lugar, Mender permite realizar actualizaciones a nivel de Sistema Operativo mientras proporciona sólidas garantías de seguridad mediante el uso de criptografía para firmar digitalmente los artefactos distribuidos, asegurando que sólo los artefactos legítimos y verificados llegan a los dispositivos. Además, utiliza canales de comunicación seguros basados en Transport Layer Security (TLS), protegiendo las actualizaciones de accesos no autorizados o manipulaciones.

Otra razón clave para su selección es la completa capacidad de automatización de Mender, que permite un proceso automatizado de distribución y renovación de certificados sin necesidad de intervención manual frecuente.

La escalabilidad de Mender lo hace ideal para entornos que requieren la actualización de grandes distribuciones de dispositivos.[15]

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

C. Servicio VPN

Para poder realizar estas comunicaciones entre el servidor central y los dispositivos finales, es posible que algunos de estos dispositivos se conecten a través de Internet y no estén conectados de manera local. Esto supone una amenaza a la confidencialidad, integridad y autenticidad de la información transmitida. Sin una protección adecuada, los datos pueden ser interceptados y manipulados mediante ataques como el *Man-in-the-Middle* (MitM)[16] o el *Cross-site scripting* (XSS).[17]

Ante estas amenazas, se han desarrollado las *Virtual Private Networks* (VPN) para proteger las comunicaciones. Una VPN es una tecnología que permite establecer conexiones seguras entre dos dispositivos a través de redes públicas. El objetivo principal de una VPN es proteger la privacidad de las comunicaciones a través del cifrado de los datos, el enmascaramiento de direcciones IP, y el acceso a redes corporativas desde ubicaciones remotas.

Estas conexiones están cifradas de extremo a extremo, de manera que los paquetes quedan protegidos desde que salen del dispositivo origen hasta que llegan a su destino. Además, el enmascaramiento de direcciones se realiza de manera que todo el tráfico se enruta a través del servidor VPN, ocultando la dirección IP real del usuario.

Las VPN se pueden clasificar según la arquitectura o según su protocolo.

a) VPNs según la arquitectura:

Las VPN de acceso remoto permiten a los usuarios conectarse de forma segura a una red privada a través de Internet. El tráfico entre el usuario y el servidor VPN se cifra completamente, protegiendo así la información. Este tipo de VPN es normalmente utilizado por empleados que trabajan de forma remota, dejando la posibilidad de acceder a recursos internos como si estuvieran conectados físicamente a la red corporativa.

Otro tipo de VPN es la VPN de sitio a sitio. Este tipo conecta redes completas entre dos o más ubicaciones físicas. En lugar de conectar dispositivos individuales, conecta routers o gateways, lo que permite que todos los dispositivos de una red local accedan a la otra red sin requerir software adicional.

Un tercer tipo es la VPN móvil, que está diseñada para dispositivos que cambian de red frecuentemente. Esta VPN mantiene una sesión persistente, incluso cuando se pierde la conexión o cambia alguna dirección IP. Este tipo se utiliza en la seguridad pública o el transporte.

Por último, la VPN basada en la nube proporciona una conexión segura entre redes locales y los servicios en la nube. Se utiliza mayoritariamente en entornos de infraestructura como servicio, como AWS o Azure.[18]

b) VPNs según el protocolo:

Uno de los primeros protocolos VPN fue el *Point-to-Point Tunneling Procotocol* (PPTP). Es fácil y rápido de configurar, pero actualmente se considera inseguro debido a vulnerabilidades conocidas.[19]

Otro protocolo es *Internet Protocol Security* (IPSec). Este es el protocolo estándar para proteger el tráfico, utilizando autenticación y cifrado.

El tercer protocolo VPN es OpenVPN, un protocolo de código abierto que utiliza la biblioteca OpenSSL para realizar el cifrado. Es altamente configurable y seguro, y es compatible con una gran cantidad de plataformas. Utiliza tanto UDP como TCP como protocolo de transporte.

Otra alternativa es utilizar los protocolos SSL o TLS para establecer conexiones VPN. Este método no requiere instalación de software adicional, lo que lo hace ideal para accesos temporales.

Por último, WireGuard es un protocolo moderno que se ha popularizado por su eficiencia, simplicidad y alta seguridad. Está diseñado para ser más ligero que IPSec o OpenVPN, y utiliza criptografía moderna como Curve25519 y ChaCha20, además de facilitar las auditorías de seguridad debido a su código fuente reducido.[20]

Debido a las múltiples ventajas, tanto en procesamiento como en seguridad, se ha decidido utilizar WireGuard como el protocolo utilizado en la VPN. La arquitectura propuesta es la VPN de acceso remoto, ya que las funcionalidades proporcionadas por las otras arquitecturas no son útiles en este escenario.

D. Gestión de certificados

Para poder alcanzar un nivel de seguridad óptimo y para el correcto funcionamiento de Mender, es necesario utilizar el protocolo TLS, un protocolo criptográfico que proporciona confidencialidad, integridad y autenticidad de los datos transmitidos entre dispositivos. Este protocolo utiliza un sistema de certificados digitales para garantizar la seguridad de las comunicaciones.

Estos certificados pueden ser emitidos por autoridades públicas de certificación reconocidas, o pueden ser autofirmados. En este caso, se utilizarán certificados autofirmados, ya que ofrecen grandes ventajas en entornos propios. Sin embargo, si no se gestionan correctamente, pueden provocar caducidades involuntarias o problemas de seguridad importantes. Para abordar esta solución de generación y renovación de certificados autofirmados, existen varias herramientas que proporcionan estos servicios.

Una de las herramientas es OpenSSL, una herramienta de código abierto que utiliza los protocolos SSL y TLS. Esta herramienta permite la generación de claves privadas, creación de certificados autofirmados, conversión entre diferentes formatos y un sistema de inspección y validación de certificados.[21]

Otra herramienta disponible es GnuTLS, una biblioteca alternativa a OpenSSL. Esta alternativa tiene una mayor integración con proyectos GNU y entornos libres, pero tiene una menor adaptación en proyectos empresariales o software propio.[22]

Una tercera herramienta es Java Keytool, una utilidad incluida en el *Java Development Kit* (JDK) que permite crear y gestionar certificados dentro de un archivo seguro llamado *keystore*. Esta herramienta es ideal para entornos en Java, pero necesita realizar varias conversiones para funcionar fuera de estos entornos.[23]

Por último, otra herramienta es mkcert, una herramienta de código abierto que facilita la creación de certificados autofirmados confiables de manera local. Esta solución es óptima en redes locales, pero no está pensada para entornos de producción.[24]

Analizando todas las herramientas disponibles, OpenSSL es la herramienta que más se adecúa al entorno propuesto, ya que este permite un mayor control y flexibilidad sobre el entorno, y tiene todas las funcionalidades necesarias sin necesidad de herramientas de terceros.[25]

IV. Topología

En el diseño de la arquitectura propuesta, la interacción entre el servidor central y cada dispositivo final se ha estructurado en torno a dos canales de comunicación paralelos, como se puede observar en la Figura 1.

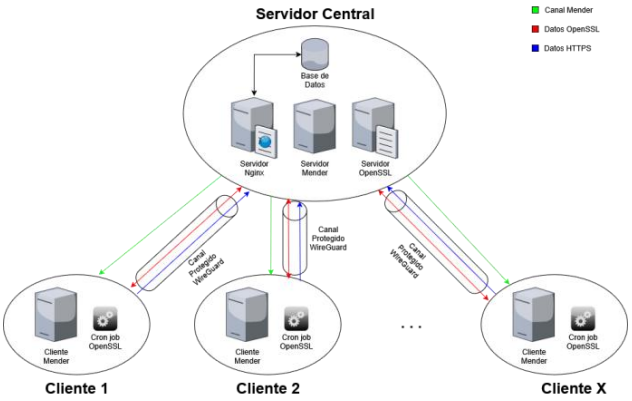


Figura 1: *Topología de la red desplegada.*

El primer canal de comunicación está constituido por la conexión entre el servidor Mender y el cliente Mender instalado en cada dispositivo final. Esta conexión está gestionada directamente por el servidor Mender, y su función principal es permitir la ejecución de actualizaciones remotas del sistema operativo o ciertas aplicaciones, así como el envío de artefactos de configuración, paquetes de software y otros elementos necesarios para mantener actualizado el entorno del cliente.

El segundo canal establece una vía de comunicación segura dedicada para la transmisión de datos recopilados localmente por el dispositivo final hacia el servidor central y la gestión de renovación de los certificados. Este intercambio puede incluir métricas de operación, registros de actividad, fotografías, modelos 3D, información de diagnóstico u otros datos relevantes que deban ser almacenados o procesados de forma centralizada. En este canal, la primera conexión se realiza a través de un servicio web alojado en el servidor central, cuya función es recibir y gestionar las solicitudes entrantes de los clientes. Para la segunda conexión dentro de este canal, tanto el servidor central como el dispositivo final puede enviar y recibir cierta información para la configuración inicial y la renovación de los certificados.

Para garantizar la confidencialidad e integridad de la información transmitida en este segundo canal, se ha implementado un sistema de cifrado punto a punto mediante WireGuard. Además del cifrado a nivel de red, se ha incorporado un sistema de autenticación mediante certificados digitales, emitidos por una autoridad certificadora (CA) local, que permite validar la identidad tanto del cliente como del servidor. En la Figura 2 se puede observar los diferentes servicios activos tanto en el servidor como en cada cliente.

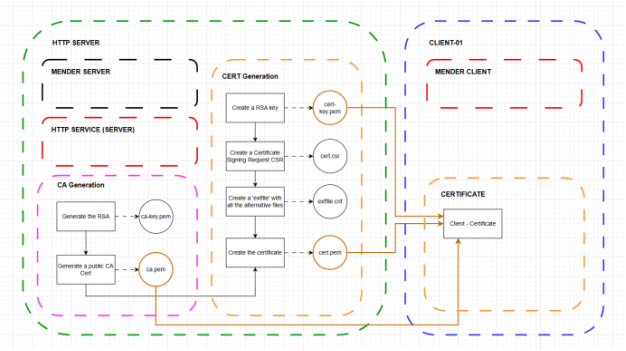


Figura 2*: Procesos y servicios activos en el servidor y el cliente*

V. Despliegue

Para el despliegue de esta arquitectura, se ha utilizado como sistema operativo base para el servidor central la distribución Ubuntu 22.04 LTS, en su versión con entorno de escritorio incluido, debido a la necesidad de compatibilidad con las herramientas necesarias para la configuración del entorno del servidor central, específicamente Mender, que necesita ser configurado a través de una interfaz web.

De la misma manera, en los dispositivos cliente, la necesidad de un entorno gráfico se mantiene, ya que para poder enviar información a través de HTTPS será necesario interactuar con los servicios web a través de un navegador. Por tanto, estos dispositivos utilizan la misma distribución, Ubuntu 22.04 LTS.

A. Servidor central

El primer paso realizado para configurar el servidor central ha sido la instalación y configuración del servidor Mender. Para llevar a cabo esta tarea de forma eficiente y reproducible, se ha optado por utilizar contenedores Docker que correrán los servicios del servidor Mender. Como parte del proceso de inicialización, se ha creado un usuario con privilegios de administrador para acceder a la interfaz de gestión web del sistema. Este usuario tiene acceso completo a todas las funcionalidades que ofrece el sistema, incluyendo la gestión de dispositivos, la supervisión de despliegues y la configuración de artefactos.

Seguidamente, se ha implementado un proceso de generación y distribución de certificados digitales con el objetivo de garantizar comunicaciones seguras entre el servidor central y los dispositivos cliente. Este procedimiento se basa en el uso de criptografía asimétrica, empleando la herramienta OpenSSL para la creación y firma de certificados.

El servidor central actuará como Autoridad Certificadora (*Certificate Authority*, CA), que será la encargada de emitir y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

firmar los certificados utilizados por los dispositivos de la red. Para ello, se generó una clave privada RSA de 4096 bits de longitud para la CA, utilizando el algoritmo de cifrado AES-256 para su encriptación.

Con la clave privada de la CA se ha creado el certificado digital del propio servidor central. Este certificado es autofirmado, ya que es el propio servidor central quien genera su certificado. Este certificado raíz será posteriormente utilizado para verificar la autenticidad de los certificados generados para los dispositivos finales.

La emisión de certificados para los dispositivos cliente se ha realizado en una segunda fase del proceso, en la cual se genera una nueva clave privada RSA, independiente de la clave de la CA, destinada a ser utilizada exclusivamente por el dispositivo solicitante. A partir de esta clave, el dispositivo genera un *Certificate Signing Request* (CSR), que constituye una petición formal dirigida a la CA. Este CSR incluye tanto la clave pública del dispositivo solicitante como la información de identidad asociada, y se utiliza para solicitar la emisión de un certificado válido y firmado por la CA.

Una vez recibido el CSR, el servidor central, actuando como entidad certificadora, verifica la información proporcionada y procede a firmar la solicitud utilizando su clave privada. El resultado de este proceso es un certificado digital que asocia la clave pública del dispositivo solicitante con su identidad, y que ha sido validado por la CA. Este certificado puede ser instalado en el dispositivo final y utilizado durante el establecimiento de conexiones seguras, permitiendo la verificación de identidad.

Este procedimiento debe repetirse de forma individual para cada dispositivo cliente que se pretenda incorporar a la red gestionada. De esta forma, se garantiza que cada entidad dentro del sistema dispone de una identidad criptográfica única y verificable, reforzando así la seguridad global del entorno y evitando posibles ataques basados en suplantación o reutilización de certificados.

Tras completar el proceso de generación de certificados digitales y habilitar el mecanismo de autenticación criptográfica, se procedió a configurar el servicio web en el servidor central con el objetivo de permitir la comunicación bidireccional con los dispositivos cliente.

La solución adoptada para este propósito fue la instalación del servidor Nginx, aplicando las configuraciones por defecto.

No obstante, durante el proceso de configuración se detectó un conflicto en el uso del puerto de escucha por defecto. Tanto Nginx como el servicio web de configuración de Mender utilizan el puerto 80 para servir sus respectivas interfaces web. Debido a esta colisión el puerto utilizado para acceder al servicio web de Nginx se redirigió al puerto 8080.

Si bien la habilitación del servidor web permite el intercambio de información entre el servidor central y los dispositivos cliente, es fundamental garantizar que dicha comunicación se realice de manera segura, protegiendo los datos frente a posibles interceptaciones o accesos no autorizados. Para cumplir con estos requisitos de confidencialidad e integridad, se ha implementado WireGuard como solución de VPN orientada al establecimiento de túneles cifrados entre los distintos dispositivos.

Para poder llevar a cabo la comunicación privada a través de WireGuard, se han generado de un par de claves criptográficas para el servidor: una clave pública y su correspondiente clave privada, las cuales serán utilizadas para establecer los canales cifrados con los dispositivos cliente.

Con las claves generadas, se ha creado un archivo de configuración para WireGuard. Este archivo está estructurado en dos secciones principales que corresponden a los roles de cada dispositivo dentro de la red.

La primera sección, encabezada con *[Interface]*, define los parámetros propios de la máquina donde se ejecuta el servicio, en este caso, el servidor central. En esta sección se encuentra la dirección IP asignada dentro del rango privado de la VPN, la clave privada del servidor, el puerto de escucha a través del cual se reciben las conexiones entrantes, y las reglas del firewall requeridas para permitir y proteger dichas conexiones.

La segunda sección, identificada con el encabezado *[Peer]*, especifica la configuración correspondiente a cada dispositivo cliente que participará en la red privada virtual. Esta sección incluye la clave pública del dispositivo cliente, así como su dirección IP dentro de la subred VPN. Dado que cada dispositivo posee su propio par de claves y una dirección única en la red privada, esta sección se debe replicar para cada cliente que se desee autorizar en la infraestructura. De este modo, el servidor central puede identificar de forma única a cada dispositivo y establecer un canal cifrado dedicado con cada uno de ellos.

B. Dispositivo final

Una vez completada la configuración del servidor central, se ha utilizado una máquina auxiliar, dedicada para la generación de la imagen base de los dispositivos finales, que se distribuirá a todos estos a través del servidor central. Esta imagen, conocida como la *golden image*, actúa como plantilla estandarizada para asegurar la homogeneidad en todos los sistemas desplegados.

El proceso de mantenimiento de esta imagen implica realizar las actualizaciones necesarias, como nuevas versiones del sistema operativo o del kernel, directamente sobre esta máquina auxiliar. Una vez finalizadas dichas modificaciones, se genera una nueva imagen a partir del estado actualizado del sistema, que es posteriormente transferida al servidor central para su redistribución automatizada hacia todos los dispositivos finales compatibles.

La *golden image* incorpora todas las configuraciones necesarias para el correcto funcionamiento de las herramientas integradas, como WireGuard y Mender. No obstante, la gestión de claves criptográficas y certificados se realiza a través del propio servidor central, de manera que cada dispositivo cliente utilizará sus propias claves y certificados únicos, garantizando así una configuración segura e individualizada.

Una vez generada la *golden image*, se ha llevado a cabo la instalación y configuración del cliente Mender en ésta. Una vez instalado el cliente, es necesario definir la dirección IP del servidor Mender para establecer la conexión con el servidor central.

Cabe destacar que cada dispositivo cliente se identifica de forma única a través de su dirección MAC, lo que permite al servidor distinguir entre los diferentes dispositivos.

Con el fin de establecer un canal de comunicación seguro entre los dispositivos cliente y el servidor central, es necesario habilitar la funcionalidad de conexión cifrada mediante el protocolo WireGuard en cada uno de los dispositivos finales.

Para ello, durante el proceso de preparación de la *golden image*, se ha instalado el paquete wireguard-tools, que proporciona las utilidades esenciales para el funcionamiento de WireGuard. Esta instalación se realiza de forma previa a la distribución de la imagen, asegurando que todos los dispositivos derivados cuenten con la infraestructura necesaria para la posterior configuración del túnel VPN.

A diferencia del servidor, donde se define explícitamente el archivo de configuración con todos los parámetros necesarios, en el caso de los dispositivos cliente este archivo no se genera de forma inmediata. Esto se debe a que la información requerida para su configuración, como la clave pública del servidor, la IP asignada dentro de la VPN y otras variables de red, será proporcionada desde el servidor Mender, una vez que el dispositivo se haya registrado correctamente y esté operativo dentro de la infraestructura.

Dado este enfoque diferido de configuración, se ha implementado un script automatizado que se ejecuta de forma periódica en segundo plano. La función de este script es monitorizar continuamente el estado del cliente y verificar si ya se ha recibido la información necesaria desde el servidor. En el momento en que dicha información esté disponible, el script se encargará de generar de forma automática el archivo de configuración con la información corespondiente y activará la interfaz WireGuard, estableciendo así el túnel seguro de comunicaciones.

Adicionalmente, se ha diseñado un segundo script con el propósito de gestionar la validez de los certificados utilizados en el proceso de autenticación. Este componente verifica regularmente la fecha de expiración de los certificados emitidos por la CA del servidor. En caso de que se detecte una proximidad a la fecha de caducidad, según un umbral predefinido, el script iniciará un proceso automatizado de renovación, generando una nueva solicitud de firma de certificado (CSR) y transmitiéndola al servidor para su validación y actualización.

Una vez realizadas estas configuraciones, la *golden image* está lista para utilizarse y ser distribuida a todos los dipositivos finales.

VI. Conclusiones

El caso de uso realizado en este artículo demuestra que es posible construir un entorno IoT seguro, escalable y totalmente automatizado mediante la integración de tecnologías consolidadas como Mender, WireGuard y OpenSSL. Esta combinación permite no solo la actualización remota y controlada del sistema operativo y las aplicaciones en dispositivos distribuidos, sino también la protección de las comunicaciones mediante túneles cifrados y la autenticación mutua mediante certificados digitales.

La utilización de una *golden image* estandarizada facilita la homogeneidad en la configuración inicial de los dispositivos finales, a la vez que permite su personalización segura a través de la gestión dinámica de claves y certificados. Por otro lado, la separación de funciones entre el canal de actualizaciones OTA y el canal de transmisión de datos asegura la especialización de servicios y una mayor robustez operativa.

En conjunto, la solución propuesta ofrece una alternativa viable para entornos críticos que requieren alta disponibilidad, alta seguridad y mínima intervención manual. Su implementación permite sentar las bases para futuros desarrollos orientados a la gestión inteligente de dispositivos IoT en sectores como el industrial, médico o de infraestructuras inteligentes.

Agradecimientos

This work has been partially funded by the European Cybersecurity Competence Centre (ECCC) under grant agreement No 101145874

Este trabajo ha sido parcialmente financiado por la *European Cybersecurity Competence Centre* (ECCC) bajo el acuerdo No 101145874

Referencias

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, y M. Ayyash, «Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications», *IEEE Communications Surveys and Tutorials*, vol. 17, n.º 4, pp. 2347-2376, oct. 2015, doi: 10.1109/COMST.2015.2444095.

[2] A. A. Laghari, K. Wu, R. A. Laghari, M. Ali, y A. A. Khan, «A Review and State of Art of Internet of Things (IoT)», 1 de mayo de 2022, *Springer Science and Business Media B.V.* doi: 10.1007/s11831-021-09622-6.

[3] S. Razdan y S. Sharma, «Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies», 2022, *Taylor and Francis Ltd.* doi: 10.1080/02564602.2021.1927863.

[4] J. Bauwens, P. Ruckebusch, S. Giannoulis, I. Moerman, y E. De Poorter, «Over-the-Air Software Updates in the Internet-of-Things: An Overview of Key Principles».

[5] D. Schuff y R. St Louis, «CENTRALIZATION VS. DECENTRALIZATION OF APPLICATION SOFTWARE», 2001.

[6] «Welcome! - The Apache HTTP Server Project». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://httpd.apache.org/

[7] N. Palma Pérez y * Autor, «Eficiencia de los servidores web Apache 2 y Nginx: un estudio de caso Efficiency of Apache 2 and Nginx web servers: a case study», 2020. [En línea]. Disponible en: http://publicaciones.uci.cu

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[8] «Caddy - The Ultimate Server with Automatic HTTPS». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://caddyserver.com/

[9] «AWS IoT Over-the-air Update: Design». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://aws.github.io/ota-for-aws-iot-embedded-sdk/main/ota_design.html

[10] «balena - Powerful IoT device management made simple». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://www.balena.io/

[11] «SWUpdate – Your OTA for Embedded Linux and IOT». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://swupdate.org/

[12] «Safe and Secure OTA Updates for Embedded Linux | RAUC». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://rauc.io/

[13] «OSTree Overview | ostreedev/ostree». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://ostreedev.github.io/ostree/introduction/

[14] «Comprehensive Embedded OTA Guide for IoT Solutions». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://witekio.com/blog/ota-update-solutions-the-ultimate-guide/

[15] «Mender Docs | Mender documentation». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://docs.mender.io/

[16] S. Gangan, «A Review of Man-in-the-Middle Attacks».

[17] S. J. Y. Weamie, «Cross-Site Scripting Attacks and Defensive Techniques: A Comprehensive Survey», *International Journal of Communications, Network and System Sciences*, vol. 15, n.º 08, pp. 126-148, 2022, doi: 10.4236/ijcns.2022.158010.

[18] «Tipos de conexión de VPN | Microsoft Learn». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://learn.microsoft.com/es-es/windows/security/operating-system-security/network-security/vpn/vpn-connection-type

[19] «PPTP VPN Security Risks - MyWorkDrive». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://www.myworkdrive.com/vpn-alternative/pptp-vpn-security-risks/

[20] «Protocolos VPN: tipos y en qué se diferencian | NordVPN». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://nordvpn.com/es/blog/los-mejores-protocolos-vpn/

[21] «OpenSSL». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://www.openssl.org/

[22] «GnuTLS». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://www.gnutls.org/

[23] «Java Development Kit - Documentación de IBM». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://www.ibm.com/docs/es/i/7.4.0?topic=platform-java-development-kit

[24] «mkcert.org». Accedido: 13 de junio de 2025. [En línea]. Disponible en: https://mkcert.org/

[25] E. V Bonet Esteban, «Creación y administración de certificados de seguridad mediante OpenSSL». [En línea]. Disponible en: http://www.openssl.org.

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Modelado e Implementación de un problema de Virtual Network Embedding en redes O-RAN

Daniel Jaramillo*†, Luis Diez†, Ramón Agüero†
*Departamento de Ingeniería de Telcomunicaciones. Universidad Estatal Península de Santa Elena.
Santa Elena, Ecuador. Email: djaramillo@upse.edu.ec
†Departamento de Ingeniería de Comunicaciones. Universidad de Cantabria.
Santander, España. Email: daniel-armando.jaramillo@alumnos.unican.es {ldiez, ramon}@tlmat.unican.es

La evolución de las nuevas generaciones de redes de acceso 5G/6G se está apoyando en soluciones tecnológicas que modifican tanto su despliegue como su operación y gestión. Entre ellas se encuentra la virtualización de funciones de red, que permite una mayor desagregación de los elementos de acceso, paradigma fomentado por organismos de estandarización como el 3GPP y por alianzas como O-RAN. Estas nuevas arquitecturas dan lugar a problemas de tipo Virtual Network Embedding (VNE), aplicados a la instanciación de estaciones base. En este trabajo en progreso se describe un primer modelado del problema VNE aplicado a redes de acceso desagregadas y se presentan unos resultados iniciales. El problema resultante es de tipo programación lineal entera (ILP) y se resuelve haciendo uso de la librería OR-Tools SCIP dentro del entorno VIRNE. Esta herramienta permite automatizar y evaluar diferentes tipos de algoritmos para problemas VNE. Los resultados iniciales que se presentan en el trabajo validan la implementación acometida.

***Palabras Clave*—VNE, O-RAN, desagregación, 5G**

I. INTRODUCCIÓN

En las arquitecturas tradicionales de redes de acceso móvil, las funciones de red se alojan en equipos tradicionales, normalmente en equipos hardware monolíticos, que además son rígidos y propietarios. Además, su ubicación y dimensionamiento están predefinidos. Sin embargo, en las últimas generaciones de tecnologías celulares, y muy especialmente en 5G, se ha visto una tendencia hacia la virtualización y desagregación de elementos de red de acceso. En los últimos años esta tendencia se ha acentuado con la aparición de la arquitectura O-RAN. De esta forma las implementaciones monolíticas de estaciones base se dividen entre la unidad radio (*Radio Unit* - RU) que aloja los elementos radiantes y dos entidades que implementan las funciones de banda base: *Distributed Unit* (DU) y *Central Unit* (CU). Tanto la CU como la DU son instancias software independientes, que se ejecutan sobre servidores de propósito general, ubicándose en distintos puntos de la infraestructura (edge, cloud, data center). Aunque esta capacidad de virtualización ofrece ventajas, también introduce algunos retos operativos.

Cuando las funciones están desacopladas del *hardware*, hay libertad para llevar a cabo su despliegue en diferentes ubicaciones físicas. Cada ubicación implica distintos costes de transporte de datos. No basta con saber qué función debe ejecutarse, sino dónde y cuándo, para garantizar que los flujos de los segmentos de red *fronthaul* (RU-DU), *midhaul* (DU-CU) y *backhaul* no degraden el servicio ni saturen enlaces críticos. Por otro lado, cada instancia virtual consume recursos de CPU, memoria y capacidad de comunicaciones, y su uso más intensivo puede elevar los gastos operativos (OPEX).

Por lo tanto se precisan algoritmos de localización para este tipo de arquitecturas. Este problema se puede ver como un caso particular del Virtual Network Embedding (VNE), que mapea nodos y enlaces virtuales sobre una infraestructura física. En este trabajo se presenta un modelado inicial de una red de acceso radio desagregada, dando lugar a un problema de Programación Lineal Entera (ILP) que lo resuelve. Haciendo uso del entorno VIRNE [1] se ha evaluado la solución sobre varias topologías.

El artículo se estructura como sigue. La Sección II proporciona una descripción de los antecedentes técnicos sobre los que se plantea este trabajo, en particular describiendo O-RAN y VNE. Posteriormente, la Sección III introduce el modelado que se plantea para resolver el problema, mientras que en la Sección IV se muestran unos resultados preliminares. Por último, la Sección V muestra las conclusiones del análisis e identifia las principales líneas futuras de investigación.

II. ANTECEDENTES

A. Desagregación de RAN

En O-RAN, la desagregación comprende tres componentes principales: la Unidad Central (O-CU), la Unidad

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

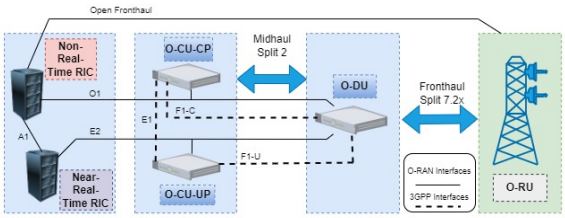


Fig. 1. Desagregación e interfaces de RAN

Distribuida (O-DU) y la Unidad de Radio (O-RU). Esta estrategia promueve la virtualización de funciones de red, permitiendo su implementación en entornos virtualizados y basados en software, lo que proporciona mayor flexibilidad y escalabilidad a los operadores [2].

- **O-DU:** Se encarga de funciones como la gestión de recursos de radio en una escala más localizada, planificación y asignación de las frecuencias y coordinación de la interfaz radio.
- **O-CU:** Aborda la gestión de recursos en una escala más amplia, optimizando la planificación, control y coordinación de la red. En la Figura 1 se muestra la interacción de estas entidades.
- **O-RU:** Atiende funciones de la interfaz radio, de capa física, como modulación y demodulación de señales, conversión analógico-digital y digital-analógico, y transmisión y recepción de señales [2].

Los splits definen los puntos de separación funcional entre la RU, DU y la CU. En 3GPP Rel-15 se evaluaron varias opciones; los más relevantes aquí son el split 2 (PDCP/RLC) y el split 7 (PHY), que condicionan latencias y cargas de cómputo entre DU y CU. La CU ejecuta las funciones por encima del split 2, la DU gestiona las funciones entre split 2 y split 7, y la RU se encarga de las tareas por debajo del split 7 y del procesamiento RF. Por su parte, la O-RAN Alliance definió además una interfaz de fronthaul multi-fabricante RU–DU basada en un split 7 modificado (split 7.2x) y mejoró la interfaz de midhaul CU–DU para entornos multi-fabricante, permitiendo tanto despliegues completamente distribuidos como la centralización de CU (y opcionalmente DU) en centros de datos con virtualización completa de funciones RAN [3], [4].

B. Virtual Network Embbedding VNE

Se define como el problema de asignar de manera eficiente redes virtuales a infraestructuras físicas. Específicamente, VNE se refiere al proceso de mapear nodos virtuales y enlaces virtuales de una red virtual a nodos físicos y enlaces físicos de una infraestructura de red subyacente (substrato físico). Los algoritmos y técnicas de VNE tienen como objetivo optimizar esta asignación [5]. Actualmente se tiene diversas soluciones que han sido propuestas para resolver este problema entre los que se puede destacar: métodos exactos, heurísticos [6], metaheurísticos y de aprendizaje por refuerzo [7].

El segmento RAN es esencial para la conectividad inalámbrica en sistemas móviles, y las técnicas de Virtual Network Embedding (VNE) se adaptan para optimizar la virtualización de recursos en este entorno inalámbrico. Además, la implementación de Network Slicing en la RAN permite crear segmentos de red virtuales que satisfacen requisitos específicos de servicios 5G/6G, y enfoques como el Algoritmo de Algas Artificiales (AAA) ofrecen una asignación eficiente de estas redes virtuales sobre la infraestructura física, respetando restricciones y demandas de cada segmento [8].

III. Modelado

El problema de VNE en el contexto de O-RAN se centra en la asignación eficiente de recursos para una red virtual. En este caso, las funciones de la CU y DU sobre una infraestructura física de servidores y enlaces fronthaul/midhaul. Este reto es intrínsecamente complejo, ya que el número de posibles asignaciones crece exponencialmente con el tamaño de las redes virtual y física, lo que conduce a problemas *NP-hard* de optimización.

Para optimizar los recursos se debe mapear cada función virtual (CU, DU) a un nodo físico de la red, garantizando que se satisfagan sus demandas de cómputo y capacidad de transmisión, sin exceder las capacidades disponibles. Al mismo tiempo, se precisa minimizar el consumo de recursos en la infraestructura física cuando sea posible, fomentando un balance de carga adecuado.

A. Modelado de la red

A continuación se detalla la formulación del problema basado en teoría de grafos para la asignación de nodos virtuales. Aunque el problema completo implica tanto el mapeo de nodos como el establecimiento de rutas, se propone abordarlos de forma separada. El modelado propuesto se centra en la asignación de nodos, dando lugar a un problema que pertenece a la categoría de Programación Lineal Entera (ILP). Para el mapeo de enlaces virtuales se propone Dijkstra como algoritmo base por su simplicidad, eficiencia y amplio uso en problemas de encaminamiento de redes, lo que lo convierte en un punto de partida válido antes de evaluar métodos más complejos en trabajos futuros

La red física se define como un grafo

$$G_P(N_P, L_P)$$

, donde N_P son los nodos físicos (servidores multifuncionales) y L_P son los enlaces físicos que interconectan dichos nodos.

Cada nodo $n \in N_P$ representa un servidor físico al cual se asocia in atributo $cpu_n \in \mathbb{R}^+$, que indica su capacidad de cómputo. Por su parte, cada enlace $(n, m) \in L_P$ representa un canal físico de comunicación que tiene asociada una capacidad de comunicaciones $bw_{n,m} \in \mathbb{R}^+$.

Cada una de las redes virtuales se define igualmente como un grafo

$$G_V(N_V, L_V)$$

, donde N_V son los nodos virtuales (que pueden representar funciones de red como CU y DU) y L_V son los enlaces virtuales que conectan estos nodos. De forma similar a la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

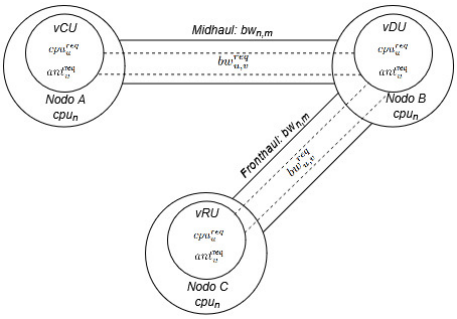


Fig. 2. Diagrama canónico O-RAN

infraestructura física, cada nodo $u \in N_V$ tiene asociado un atributo $cpu_u^{req} \in \mathbb{R}^+$, que indica su demanda de capacidad de cómputo. Asimismo, cada enlace $(u, v) \in L_V$ se define por un parámetro $bw_{u,v}^{req} \in \mathbb{R}^+$, que indica la capacidad (ancho de banda) que necesita.

La red virtual considera estos parámetros para integrarse con la red física, asegurando que los requisitos de recursos de los nodos y enlaces virtuales se alineen con las capacidades de la infraestructura física. Esto incluye la evaluación de la disponibilidad de recursos en los nodos físicos y la optimización de la asignación de recursos para minimizar la fragmentación y maximizar la eficiencia. Cabe indicar que, en este caso, una red virtual corresponde a una única estación base, por lo que estará compuesta por 3 nodos, con una topología en línea (CU - DU - RU); este modelado está representado de manera básica en la Figura 2. Por otro lado, en la infraestructura física deberá haber nodos con elementos radiantes (antenas) que se puedan vincular a las RU de la red virtual. En el caso más sencillo, los nodos físicos que representan una torre de comunicaciones contendrán una única antena, que se corresponderá al requisito de la RU y establecerá las posibles ubicaciones de las RU en la red física. Por otro lado, se podrían considerar esquemas en los que una DU gestione varias RUs, o una CU varias DUs, lo que daría lugar a topologías en forma de árbol. En este trabajo se resuelve únicamente el mapeo de nodos; el mapeo de enlaces se realiza a posteriori con Dijkstra, verificando la capacidad de los enlaces físicos.

B. Variables de decisión

Estas variables permiten modelar las decisiones que el algoritmo debe tomar para optimizar la asignación de recursos. La primera de ellas, $x_{u,n}$ indica la asignación de cada nodo virtual $u \in N_V$ y cada nodo físico $n \in N_P$ tal como sigue:

$$x_{u,n} = \begin{cases} 1, & \text{si } u \text{ se asigna a } n, \\ 0, & \text{en otro caso.} \end{cases} \tag{1}$$

Asimismo se define una variable relacionada con la anterior, que indica la utilización de nodos físicos. Para cada nodo físico n se define la variable binaria z_n de forma que toma valor 1 si se asigna algún nodo virtual al nodo físico, o 0 en caso contrario.

C. Problema de optimización

A partir de las variables definidas anteriormente se plantea el Problema 1 que tiene como objetivo minimizar el número de nodos físicos utilizados.

Problem 1:

$$\min_{z_n, x_{u,n}} \quad \sum_{n \in N_P} z_n \tag{2}$$

$$\text{s.a.} \quad \sum_{n \in N_P} x_{u,n} = 1, \quad \forall u \in N_V \tag{3}$$

$$\sum_{u \in N_V} \text{cpu}_u^{\text{req}}\, x_{u,n} \leq \text{cpu}_n^{\text{cap}}, \quad \forall n \in N_P \tag{4}$$

$$\sum_{u \in N_V} \text{ant}_u^{\text{req}}\, x_{u,n} \leq \text{ant}_n^{\text{cap}}, \quad \forall n \in N_P \tag{5}$$

$$z_n \geq x_{u,n} \quad \forall n \in N_P, u \in N_V \tag{6}$$

$$x_{u,n}, z_n \in \{0,1\} \forall u \in N_V, \forall n \in n \in N_P \tag{7}$$

Además se define un conjunto de restricciones que garanticen que la asignación respeta las características de la red física y los requisitos de la virtual. La restricción (3) asegura que todos los nodos virtuales se asignan a un único nodo físico. Por su parte, la restricción (4) garantiza que la capacidad de procesado de los nodos virtuales alojados en un nodo físico no excede la capacidad del segundo. Mediante la restricción (5) se asegura que no se excede la capacidad de antenas ubicadas en los nodos.

Finalmente, la restricción (6) permite que la variable z_n tenga valor 1 cuando un nodo físico se utiliza. Para un nodo físico $n \in N_P$, si al menos un $x_{u,n} = 1$ esta restricción obliga a que $z_n = 1$. En caso contrario, si no se asigna ningún nodo virtual al nodo físico n, z_n tendrá valor 0 al utilizarse para minimizar la función objetivo.

IV. RESULTADOS

Se eligió VIRNE por ser una plataforma académica de libre acceso que integra múltiples solvers (exactos, heurísticos y metaheurísticos) diseñados específicamente para problemas VNE, este entorno incluye tanto la definición de escenarios como la generación de resultados. Además, esta herramienta realiza el despliegue de redes virtuales (estaciones base en el escenario contemplado) de forma secuencial, y actualiza el estado de las capacidades, tanto de nodos como enlaces. Más concretamente se ha implementado un nuevo solver dentro de VIRNE que hace uso de la librería OR-TOOLS para resolver el problema resultante.

Así, se busca analizar la respuesta del problema sobre una topología canónica tipo *spine-leaf*, tal como se muestra en la Figura 3. La topología representa una jerarquía de 3 niveles de agregación. La red física se ha configurado de tal forma que los nodos inferiores pueden alojar RUs; concretamente cada nodo inferior tiene un elemento de antena de tal forma que se asegura que cada RU se asocia a un único nodo. Por otro lado, los 4 nodos físicos superiores pueden alojar DUs o CUs. Los parámetros con los que se realizó las pruebas fueron:

$cpu_{RU}^{req} = [5]$, $cpu_{DU}^{req} = [10]$, $cpu_{CU}^{req} = [0]$, $cpu_n^{cap} \in \{0, 5, 10, 15, 20, 25, 30\}$.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

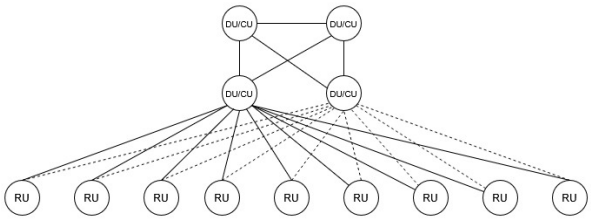


Fig. 3. Topología *spine-leaf*

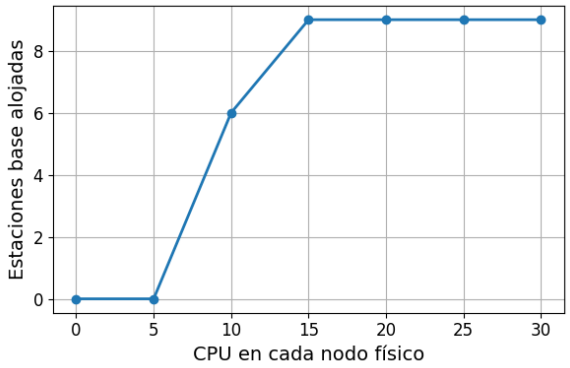


Fig. 4. Experimentos en escenario básico

Respecto a las redes virtuales, cada una de ellas consiste en 3 nodos, de tal forma que la RU requiere un elemento de antena y las DU tuvieran mayores requisitos de CPU que las CU, reflejando que las DU deben ejecutar funciones de las capas bajas, con mayor demanda computacional. En este experimento de validación se generan 9 redes virtuales, y se analiza el comportamiento a medida que se aumenta la capacidad de procesado de los elementos físicos.

Los resultados se pueden observar en la Figura 4, en el eje de abscisa se muestra la capacidad de los nodos que pueden alojar DU/CUs. Como era de esperar, para valores muy bajos de capacidad el total de recursos de procesado es insuficiente para alojar ninguna VNR. Por otro lado, cuando la CPU por nodo alcanza una capacidad de 15, se logran alojar las 9 estaciones base.

El valor de capacidad no corresponde a una métrica física exacta, sino a una abstracción en unidades de CPU dentro del modelo, que permite observar la relación entre capacidad de procesamiento y cantidad de estaciones base embebidas, aportando información preliminar sobre la escalabilidad del esquema y validando la utilización de VIRNE para la solución del problema.

Por tanto, la función objetivo minimiza $\sum_n z_n$ (Ec. 2), y el solver empieza a encontrar soluciones que alojan las 9 VNR a partir de $cpu_n^{cap} = 15$ por nodo físico, reduciendo el número de nodos activos necesarios conforme crece la capacidad.

V. CONCLUSIONES

En este trabajo inicial se han presentado un modelado inicial de un problema VNE para escenarios de redes de acceso virtualizadas y desagregadas. A diferencia de otros enfoques similares, las redes que se despliegan contienen tanto elementos virtuales, cuya ubicación es abierta, como elementos que requieren ciertas características (restricciones) físicas (elementos de antena), cuya ubicación está más restringida, encorsetando más la posible solución.

Este modelado planteado da lugar a un problema de tipo ILP, cuya resolución se acomete dentro del entorno VIRNE. Para validar la implementación se ha analizado el comportamiento del *solver* implementado sobre un escenario sencillo, comprobando que se corresponde con el esperado.

En el futuro se acometerá la extensión del modelo, incluyendo topologías de estaciones base más complejas. Estas incluirán los casos en que varias RU son gestionadas por una misma DU, o hay una CU común para varias DUs. Asimismo, se explorará el modelado cuando se requiere una coordinación estrecha entre diferentes estaciones base. Por otro lado, se pretende explotar las funcionalidades de VIRNE tanto para extender el análisis como para comprar las capacidades de resolución con diferentes algoritmos existentes dentro de la plataforma.

AGRADECIMIENTOS

Este trabajo ha sido financiado parcialmente por el Ministerio de Asuntos Económicos y Transformación Digital de España mediante el proyecto SITED (PID2021-125725OB-I00), y por el Gobierno de Cantabria, Fondo Europeo de Desarrollo Regional (FEDER), a través del programa "Ayudas a proyectos de investigación con alto potencial industrial de agentes tecnológicos de excelencia para la competitividad industrial (TCNIC)" (2023/TCN/002).

REFERENCIAS

[1] "Virne framework," https://github.com/GeminiLight/virne, [Online; accessed: Sep. 12, 2025].
[2] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023.
[3] F. Lorca-Hernando, E. Serna-Santiago, and M. Aparicio-Peña, "Telefónica views on the design, architecture, and technology of 4g/5g open ran networks," Telefónica, Tech. Rep., 2025. [Online]. Available: {https://www.telefonica.com/es/wp-content/uploads/sites/4/2021/08/Whitepaper-OpenRAN-Telefonica.pdf}
[4] L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and learning in o-ran for data-driven nextg cellular networks," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 21–27, 2021.
[5] H. Cao, S. Wu, Y. Hu, Y. Liu, and L. Yang, "A survey of embedding algorithm for virtual network embedding," *China Communications*, vol. 16, no. 12, pp. 1–33, 2019.
[6] M. Minardi, T. X. Vu, L. Lei, C. Politis, and S. Chatzinotas, "Virtual network embedding for ngso systems: Algorithmic solution and sdn-testbed validation," *IEEE Transactions on Network and Service Management*, vol. 20, no. 3, pp. 3523–3535, 2023.
[7] P. Zhang, Y. Hong, X. Pang, and C. Jiang, "Vne-hpso: Virtual network embedding algorithm based on hybrid particle swarm optimization," *IEEE Access*, vol. 8, pp. 213 389–213 400, 2020.
[8] R. Gomes, D. Vieira, M. B. Pereira, and M. F. de Castro, "Artificial algae optimization for virtual network embedding problems in 5g network slicing scenarios," *Expert Syst. Appl.*, vol. 239, no. C, Apr. 2024.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Arquitectura para el Slicing Dinámico en Redes Privadas 5G

Antonio Hidalgo[(1)], David Cortés-Polo[(1)], Jesús Calle-Cancho[(1)], Pablo Bertrand[(2)], Claudina Rattaro[(2)].
ahidalgof@unex.es, dcorpol@unex.es, jesuscalle@unex.es, pbertrand@fing.edu.uy, crattaro@fing.edu.uy
[(1)]Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos. Universidad de Extremadura. Cáceres, España.
[(2)]Instituto de Ingeniería Eléctrica. Universidad de la República. Montevideo, Uruguay

La evolución de las redes 5G adaptándose a los requerimientos de los operadores y la industria, para atender la creciente demanda, ha introducido el concepto de *slicing*, permitiendo múltiples redes lógicas sobre una misma infraestructura física. La introducción de estas capacidades de *slicing* sobre el *Core*, la red de transporte y la red de acceso de la red móvil ha supuesto una revolución en la forma en que se gestionan y ofrecen los servicios de red, siendo una de las líneas de trabajo no sólo en redes comerciales sino en soluciones Open Source para redes privadas. Este trabajo presenta un *testbed* de código abierto para *slicing* en redes 5G *standalone*, que permite la realización de pruebas avanzadas para la compartición de recursos de la red entre diferentes usuarios. Para ello, todo el *software* utilizado en la red privada está liberado como código abierto, además se ha basado el *testbed* en SDN para la gestión de los recursos del *core* de la red. Además, se ha desarrollado un modelo basado en Deep Learning empleando, un modelo LSTM, entrenado con tráfico multimedia en tiempo real, para predecir y ajustar dinámicamente el *throughput* de los *slices*. El modelo toma como entrada el *throughput* de los UEs y optimiza el uso de recursos, garantizando calidad de servicio y reduciendo la congestión. La plataforma es modular, reproducible y útil para el prototipado de políticas de orquestación, incorporando un enfoque inicial de *slicing* impulsado por Deep Learning.

***Palabras Clave*—*slicing*, 5G, *testbed*, SDN, LSTM, Deep Learning**

I. INTRODUCCIÓN

La evolución de las redes 5G ha introducido el *slicing* de red, un paradigma que permite la creación de múltiples instancias de redes virtuales sobre una infraestructura física compartida, cada una de ellas a medida de las necesidades del servicio o aplicación específico. Esta aproximación es esencial para atender las diversas demandas de aplicaciones inalámbricas modernas, particularmente los tres principales casos de uso definidos por la Unión Internacional de Telecomunicaciones (UIT): enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), y Massive Machine-Type Communications (mMTC) [1].

Según el 3rd Generation Partnership Project (3GPP), un *slice* de red se define como una red lógica que provee unas capacidades y características específicas, soportando varias propiedades de servicio para los clientes del *slice* de red. El sistema 5G está diseñado para ofrecer soporte optimizado a una extensa variedad de servicios de comunicación, diferentes cargas de tráfico, y diversas comunidades de usuarios finales. Por ejemplo, algunos servicios que utilizan el *slicing* de red son comunicaciones Vehicle-to-Everything (V2X), servicios eMBB constantes con FMC (Fixed-Mobile Convergence), y conexiones IoT (Internet of Things) masivas. Además, habilita el uso de aplicaciones emergentes como el cloud gaming en tiempo real, streaming de vídeo 8K, experiencias de realidad virtual, aumentada e inmersiva —casos de uso que exigen un ancho de banda alto, una latencia ultra baja, y una consistente calidad de servicio [2].

La gestión y orquestación de los *slices* de red son fundamentales para que las operadoras puedan ofrecer sus servicios de comunicación de manera eficiente. Esta capacidad no solo optimiza el uso de los recursos de red, sino que también permite la creación de nuevos modelos de negocio. Al repartir y monetizar los recursos bajo demanda, se pueden atender las necesidades específicas de industrias verticales y proveedores de servicios de terceros [3].

Uno de los aspectos técnicos más relevantes en la gestión del *network slicing* es la definición y cumplimiento de los Acuerdos de Nivel de Servicio (ANS), que establecen las métricas de rendimiento que deben garantizarse para cada cliente. Para asegurar que estas métricas se cumplan, es necesario implementar mecanismos de monitoreo continuo que evalúen el rendimiento mediante KPIs (Key Performance Indicators). Estos indicadores permiten verificar que cada *slice* funcione conforme a los estándares establecidos.

Además, el análisis de los datos de gestión juega

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

un papel crucial, ya que permite optimizar tanto el rendimiento como la asignación de recursos dentro de cada *slice*. Este análisis se complementa con la garantía de servicios de comunicación de circuito cerrado, que implica la implementación de mecanismos de retroalimentación automatizados para mantener y mejorar la calidad del servicio de forma continua. Otro elemento importante es la gestión de redes no públicas, que se refiere al manejo de redes privadas que utilizan *slicing* para ofrecer servicios especializados a determinados sectores o empresas. Finalmente, la seguridad de cada *slice* debe ser una prioridad, asegurando que se implementen medidas robustas para proteger la integridad y confidencialidad de los datos y servicios que se ejecutan en cada segmento de red.

Para implementar el *slicing* de red de forma eficiente, se han utilizado SDNs (Software Defined Network), que provee un mecanismo de control programable que abstrae y centraliza la gestión de la red. Junto con NFV (Network Function Virtualization, SDN permite desacoplar el *hardware* y el *software* de la red, permitiendo una reserva de los recursos altamente flexible y dinámica. Cada *slice* de red puede ser configurado independientemente para satisfacer los requisitos específicos de rendimiento —como la latencia, ancho de banda, fiabilidad y seguridad— haciéndolo adecuado para un extenso rango de casos de uso [4]. Esta capacidad es particularmente importante en escenarios donde coexisten múltiples servicios y la adaptación de los recursos de la red es necesaria para mantener la calidad de servicio o QoS (Quality of Service).

Aunque el *slicing* de red conforma los cimientos para la diferenciación de servicio, la provisión dinámica de *slices* —especialmente en repuesta a la demanda fluctuante y los requisitos de servicio heterogéneos— permanece como un desafío de investigación abierto. Los mecanismos de reserva de recursos eficientes deben tener en cuenta patrones variables de tráfico, restricciones de latencia y niveles de prioridad en tiempo real. En este contexto, la inteligencia artificial (IA) y el machine learning (ML) han surgido como aproximaciones prometedoras para mejorar la toma de decisiones, automatizar u orquestar, y mejorar la escalabilidad de la gestión del *slicing*.

En este trabajo, presentamos un *testbed* 5G modular y de código abierto que aprovecha el *slicing* de red basado en SDN para reservar de forma dinámica los recursos. Para construir este *testbed* se ha utilizado Proxmox como entorno de virtualización e integrado Open5GS como *core* 5G, Ryu como controlador SDN, Open vSwitch (OVS) para gestionar flujos de tráfico, UERANSIM para simular la RAN (Radio Access Network), y srsRAN para soportar la conectividad entre un gNB físico externo y un nodo móvil o UE (User Equipment) con tráfico real.

Una innovación clave de este *testbed* es la gestión adaptativa del *throughput* del *slice*, controlado por SDN. La estrategia implementada es la siguiente: utilizar una red neuronal LSTM (Long-Short Term Memory) para predecir patrones de tráfico proactivamente ajustando la reserva de recursos. Esta aproximación asegura la eficiente utilización de los recursos y mantiene un aislamiento del rendimiento, garantizando que el *slice* con mayor prioridad reciba suficiente ancho de banda incluso bajo una congestión de la red. La naturaleza abierta y extensible de esta plataforma la hace una valiosa herramienta para investigación avanzada en orquestación de redes dirigidas por IA y experimentación reproducible en 5G.

Este artículo pretende demostrar como el control basado en SDN combinado con componentes de código abierto 5G pueden habilitar una gestión de *slices* dinámica, inteligente y reproducible. Se evaluará el rendimiento de la gestión dinámica de recursos con el LSTM, en términos de eficiencia de recursos y diferenciación de servicios.

El resto del paper está organizado de la siguiente manera: la Sección II habla de trabajos relacionados desde *slicing* de red a reserva de recursos basado en SDN. La Sección III detalla la arquitectura del *testbed*, incluyendo componentes *hardware* y *software*. La Sección IV explica los mecanismos de adaptación del *slice*. La sección V demuestra los resultados obtenidos. Por último, la sección VI expone las conclusiones y trabajos futuros.

II. TRABAJOS RELACIONADOS

Los últimos años han visto un progreso significativo en el desarrollo de testbeds de *slicing* 5G y de mecanismos de *slicing* basados en SDN. Varios grupos de investigación han demostrado la viabilidad de la reserva dinámica de recursos y la garantía de QoS utilizando los paradigmas SDN/NFV. Abajo, destacamos algunos de los más relevantes e iniciativas comparables a nuestro propio trabajo.

Pineda *et al.* [5] exponen uno de los primeros despliegues integrales de código abierto de un *testbed* 5G *standalone* (SA) con capacidades SDN. Su plataforma integra Open5GS para el *core* 5G, UERANSIM y srsRAN para la RAN y UE, OpenDayLight (ODL) como el controlador SDN, y Open vSwitch (OVS) para la redirección del plano de datos. En este trabajo se prueban dos escenarios —uno totalmente virtualizado con UERANSIM y otro con un sistema híbrido utilizando srsRAN sobre *hardware* SDR (Software Defined Radio)— configurando el *slicing* de red, a través de las políticas SDN, para conformar el *throughput* y la latencia. Nuestra arquitectura está directamente inspirada por este trabajo, pero se desvía en algunos aspectos clave: adoptamos Proxmox VE para la virtualización y así simplificar la gestión de bridges y VLANs, dividimos las funciones del plano de control y de usuario en máquinas virtuales separadas para mayor aislamiento, e incrementamos la configuración básica del *slice* con un umbral dinámico —y mecanismos de adaptación basados en machine learning.

Sulaiman et al. [6] presentan vNetRunner, un framework innovador para el modelado de *slices* en redes 5G y más allá. Este trabajo se distingue por su capacidad para predecir el *throughput* (y el retardo medio) a nivel de flujo para funciones de red virtualizadas (VNFs) individuales, que luego se componen para modelar *slices*

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de extremo a extremo. La validación de vNetRunner se realiza utilizando conjuntos de datos de testbeds 5G de código abierto, demostrando una alta precisión en la estimación del *throughput* y destacando la importancia de las características a nivel de flujo para la eficiencia y escalabilidad en la gestión de recursos. Este enfoque de modelado y predicción es directamente relevante para nuestra propuesta, ya que busca optimizar la asignación de recursos basándose en el tráfico previsto, similar a nuestra estrategia de adaptación basada en LSTM. En nuestro caso, tanto el *dataset* como el modelo son más ligeros, permitiendo el uso del LSTM en tiempo real.

En una línea complementaria, Owusu et al. [7] proponen un enfoque de balanceo de carga dinámico en entornos SDN que utiliza algoritmos de aprendizaje automático, específicamente *Transformers* y *Deep Reinforcement Learning*. Aunque su objetivo principal es el balanceo de carga, este estudio es pertinente porque su modelo es capaz de considerar tanto las tendencias históricas como las cargas futuras pronosticadas para tomar decisiones más efectivas en la gestión de recursos. La capacidad de pronosticar la carga futura para optimizar la asignación de recursos tiene un impacto directo en el *throughput* y se alinea con la necesidad de una gestión de red adaptativa en tiempo real en entornos 5G, lo que refuerza la validez de integrar el aprendizaje automático para la adaptación proactiva del *throughput* en nuestro *testbed.*

Estos trabajos subrayan el potencial de enfoques basados en SDN para gestionar el *slicing* 5G de manera efectiva, así como la gestión de los recursos de la red. A diferencia de plataformas existentes, nuestra arquitectura combina de forma única una pila 5G SA de código abierto, virtualización basada en Proxmox VE, SDN gestionado por el controlador Ryu, y una gestión dinámica de los recursos en tiempo real, aportando un entorno flexible y reproducible para la investigación de *slicing* 5G.

III. ARQUITECTURA DEL *TESTBED*

Para evaluar el control dinámico de los *slices* de red en un entorno 5G, hemos construido un *testbed* experimental utilizando *software* de código abierto y mecanismos de control de tráfico basados en SDN. Este *testbed* está diseñado para la configuración, monitorización, y reserva adaptativa del ancho de banda entre los *slices*, así como la diferenciación de servicios, permitiendo a los investigadores prototipar estrategias de gestión de recursos en condiciones controladas. El sistema implementa una red end-to-end con capacidades de *slicing* tanto en el *core* como en las capas de transporte.

El *testbed* integra una mezcla de elementos físicos y virtualizados desplegados sobre el hipervisor Proxmox VE, elegido por su simplicidad y por mantener la heterogeneidad con respecto a instalaciones previas. Su arquitectura se alinea con el framework de *slicing* definido por la 3GPP [8], soportando caminos aislados para flujos de tráficos priorizados y best-effort. Las decisiones de *slicing*, ajustes de *throughput* y diferenciación de servicios se realizan en tiempo real utilizando técnicas SDN y APIs RESTFUL, que permiten que la lógica externa aplique políticas dinámicas basadas en la demanda de tráfico y la intención de nivel de servicio.

A. Componentes hardware y software

El *testbed* al completo está alojado en un solo nodo Proxmox VE, elegido por sus capacidades de hipervisor ligeras, el soporte para bridges nativos compatibles con VLAN (Virtual Local Area Network), y planificación eficiente de recursos. Los entornos host son nueve máquinas virtuales, cada una asignada a un rol en la red específico de acuerdo a la arquitectura 5G.

Open5GS [9] se utiliza como *core* 5G y se divide en tres MVs (Máquina virtual): una que aloja el plano de control, AMF (Access and Mobility Management Function y SMFs (Session Managemente Function) entre otros), y las otras dos se encargan de las funciones del plano de usuario o UPFs (User Plane Function), una por cada *slice* de red. Este esquema es totalmente extensible para soportar *n slices*, simplemente instanciando UPFs adicionales y asociándolos con sus respectivas configuraciones de *slice* específicas. Esta separación nos permite emular el manejo del plano de usuario por *slice*.

UERANSIM [10] provee los gNB y UEs emulados, asegurando una compatibilidad total con Open5GS y exponiendo de forma completa la señalización NAS (Non-Access Stratum). En paralelo, el *testbed* soporta un escenario de acceso de radio real utilizando srsRAN [11] y un SDR Ettus USRP B210, permitiendo la comunicación inalámbrica con un teléfono móvil comercial configurado con una SIM (Subscriber Identity Module) personalizada.

El controlador SDN es Ryu en su versión 4.34, que es ejecutado en una máquina virtual dedicada. Para la comunicación del controlador con el OVSDB se ha hecho uso de una API RESTFUL [12]. El Open vSwitch (OVS), versión 2.13.8, está alojado en otra MV, que sirve como switch del plano de datos. OVS maneja todo el reenvío de capa 2 y aplica el control de *throughput* utilizando colas y reglas QoS basadas en el Linux Hierarchical Token Bucket (HTB) scheduler. Cada MV tiene Ubuntu 20.04, con una configuración personalizada de la red que incluye múltiples tarjetas de red conectadas a bridges Linux etiquetados con VLAN. Esto permite control más detallado de las rutas del tráfico y la interconexión entre los nodos. Todas las MV fueron configuradas con IPs estáticas y reglas de NAT persistentes, además se habilitó SSH para posibilitar una orquestación centralizada. La interfaz de Proxmox se utiliza para gestionar el mapeo de direcciones MAC, asegurando una asignación de interfaces de red consistente tras reinicios.

Se ha prestado especial atención a la integración de los componentes: los puertos OVS están dinámicamente asignados a unas colas específicas, además de añadir reglas de flujo para asociar el tráfico de ingreso de estos puertos con la política QoS apropiada. Las APIs REST expuestas por Ryu se han validado utilizando Postman y posteriormente usadas en scripts de Python para control en tiempo real. El esquema de adaptación está regido

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

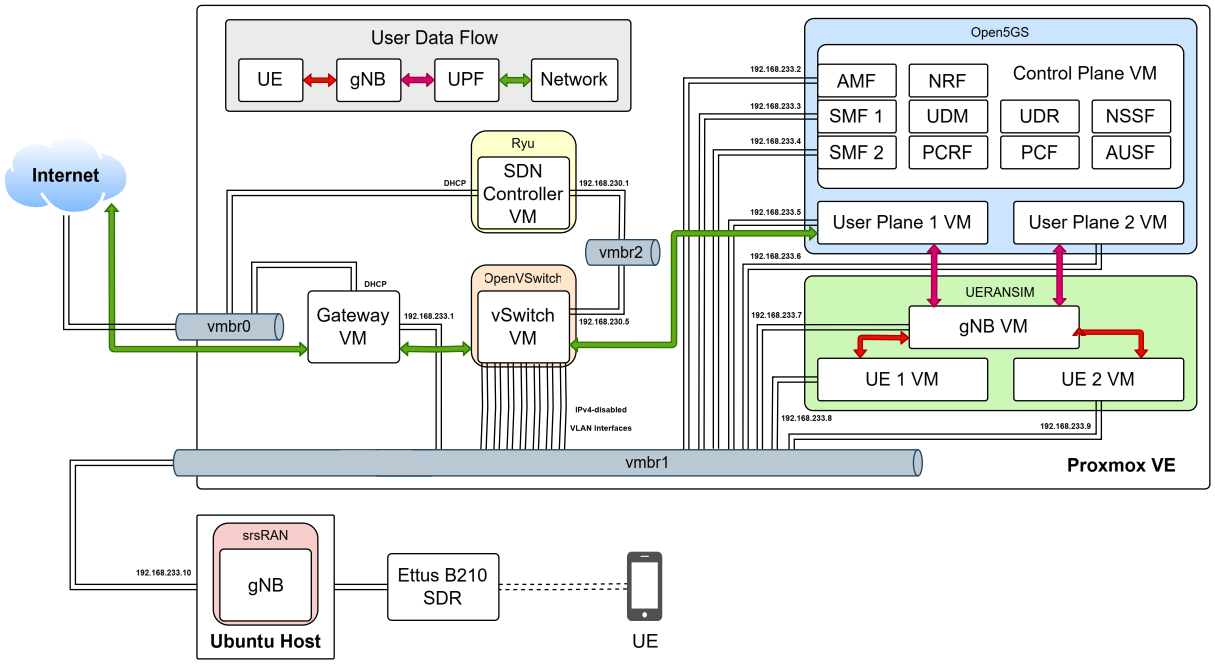


Fig. 1. Arquitectura del *testbed*

por el modelo LSTM, entrenado con tráfico VoD (Video on Demand), que permite predecir la demanda así como identificar el tipo de servicio.

B. Arquitectura

La Figura 1 de la arquitectura ilustra la separación lógica de los *slices*, rutas de tráfico, y elementos de control. Todos los UEs están conectados ya sea al gNB simulado o real, y su tráfico está enrutado hacia el switch OVS, donde es etiquetado y clasificado. Se han definido dos *slices*:

- *Slice* 1: *Slice* con prioridad inicializado con la mitad del ancho de banda. El *throughput* máximo de este *slice* estará definido por las predicciones del modelo LSTM.
- *Slice* 2: *Slice* que sigue un comportamiento *Best-Effort* inicializado con la mitad del ancho de banda.

El tráfico de cada uno de los planos de usuario están dirigidos a un UPF dedicado, permitiendo la aplicación de políticas QoS por *slice*. El switch OVS es el punto de unión para todo el trafico de datos y está configurado con múltiples colas Linux-HTB. Ryu gestiona las colas y su configuración de max-rate a través de las API REST.

La configuración de los mecanismos de adaptación requiere de una coordinación significativa entre el plano de control y de datos, mientras que la creación de las colas y flujos ha sido orquestada para asegurar que los identificadores de puertos, direcciones MAC, y etiquetas VLAN se ajustaran a lo largo de Proxmox, OVS y Ryu. La lógica de adaptación se comunica con el controlador para medir el *throughput* y ajustar dinámicamente la configuración de las colas.

La inclusión de un sistema con gNB real utilizando srsRAN añadió un punto de complejidad, ya que requiere de una sincronización *hardware*, configuración del SDR, y una interfaz de unión entre entornos físicos y virtuales. Este modelo híbrido nos permite validar la funcionalidad del *slicing* con trafico real a la vez que se mantiene la flexibilidad de la simulación.

El resultado es un *testbed* modular y reproducible que puede ser usado para explorar estrategias de reserva dinámica de *slices*, aplicación de QoS, coexistencia de múltiples *slices*, y una adaptación inteligente de tráfico sobre condiciones de carga realistas.

IV. ADAPTACIÓN DEL *SLICE*

Un objetivo clave de este trabajo es demostrar que se puede controlar el *throughput* en un entorno de *slicing* 5G utilizando técnicas SDN. Esta capacidad es particularmente importante para servicios de video bajo demanda, donde mantener un bitrate consistente y minimizar el buffering son requisitos críticos para asegurar una experiencia de usuario de alta calidad. Con ese fin, el *testbed* propuesto incluye mecanismos de control capaces de monitorizar en tiempo real, flujos de tráfico y adaptar el ancho de banda reservado por cada *slice* como corresponda. Este mecanismo está implementado en Python, creado como una aplicación Ryu, que aprovecha la API REST de Ryu para comunicarse con el switch OVS.

La lógica de adaptación opera en los elementos del plano de datos configurados en OVS. Cada *slice* está asociado con una cola dedicada gestionada por OVS utilizando Calidad de Servicio (QoS) basada en Linux HTB. Estas colas están definidas con valores máximos de *throughput* (max-rate) que determinan cuánto ancho de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

banda un *slice* puede utilizar en un momento dado.

Con el fin de apoyar la reactividad del sistema, se ha implementado una red neuronal LSTM que permite predecir la futura demanda de tráfico del *slice* 1. Esto permite que el *testbed* ajuste el ancho de banda de los *slices* según la utilización esperada determinada por el modelo.

Este modelo ha sido entrenado de forma offline utilizando trazas de tráfico reales recogidas utilizando el propio *testbed* como escenario de prueba. El tipo de tráfico elegido para el entrenamiento fue vídeo bajo demanda (VoD) de YouTube, Twitch, Tiktok y Prime Video. Cada uno de ellos presentan patrones diferentes, lo que permite además de predecir la demanda de tráfico, identificar el tipo de servicio que está consumiendo el UE (YouTube, Twitch, Tiktok o Prime Video), una funcionalidad que habilitaría una reserva de recursos aún más detallada. Los patrones presentados por estos servicios son ideales para los modelos LSTM, que son especialmente buenos en la captura de dependencias temporales y en tendencias a largo plazo en series temporales.

Una vez entrenado, el modelo es integrado en la aplicación Ryu. Esta aplicación mantiene una ventana de los últimos valores de *throughput* del *slice*, que es la entrada del modelo, y este devuelve una predicción del valor de *throughput* para el siguiente intervalo y también del tipo de servicio que se está consumiendo. Según esta predicción, el script aplica la siguiente política adaptativa:

- Se inicializan ambos *slices* con un max_rate igual al ancho de banda máximo dividido entre dos.
- Si la predicción es mayor que el max_rate actual del Slice 1, entonces el max-rate de la cola QoS asignada al *slice* 1 pasará a ser, la predicción multiplicada por 1.1, para evitar problemas de buffering. Por otra parte, el max_rate de la cola Qos asignada al *slice* 2 pasará a ser el resto del ancho de banda no utilizado por el *slice* 1.
- En el caso de que el ancho de banda restante no utilizado por el *slice* 1 sea menor a 1 Mbps, entonces el max_rate del *slice* 2 será 5 Mbps, para evitar desconexiones en este *slice*.
- Si la predicción es menor que el max_rate actual del Slice 1, entonces se repartirá equitativamente la capacidad entre ambos *slices*.

Esta política está implementada como se expone en el Algoritmo 1. Con este mecanismo, el *slice* 1 se asegura un servicio continuo y suave, ya que siempre obtendrá los recursos necesarios. El modelo continuará siendo llamado periódicamente con los últimos datos de la topología, permaneciendo relevante mientras los patrones de tráfico evolucionan.

Algorithm 1 Algoritmo de Adaptación de Throughput de *slices*

Require: prediccion_{s1}: Predicción de *throughput* para el Slice 1 (Mbps)
Require: ancho_banda_t: Ancho de banda total disponible (Mbps)

if $\text{prediccion}_{s1} > \text{max_rate_actual}_{s1}$ **then**
  $\text{nuevo_max_rate}_{s1} \leftarrow \text{prediccion}_{s1} \times 1.1$ ▷ Evitar problemas de buffering
  $\text{nuevo_max_rate}_{s2} \leftarrow \text{ancho_banda}_{total} - \text{nuevo_max_rate}_{s1}$
  if $\text{nuevo_max_rate}_{s2} < 5$ Mbps **then**
    $\text{nuevo_max_rate}_{s2} \leftarrow 5$ Mbps
  end if
else
  $\text{nuevo_max_rate}_{s1} \leftarrow \frac{\text{ancho_banda}_t}{2}$ ▷ Reparto equitativo
  $\text{nuevo_max_rate}_{s2} \leftarrow \frac{\text{ancho_banda}_t}{2}$
end if

Este método de adaptación es modular y extensible. Valida la habilidad de los controladores SDN, como Ryu, de servir como orquestadores en entornos de *slicing* 5G, gestionando el tráfico de forma dinámica sin estorbar a flujos existentes.

Por otra parte, para cerrar la interacción entre el script, Ryu y OVS, se ha realizado de siguiente manera:

1) Monitorización del *throughput*: El controlador recibe todos los paquetes provenientes de la función del plano de usuario (UPF) del *slice* 1. De esta manera, se acumula el número de bytes recibidos por cada UE durante 0,5 segundos y se hace el cálculo del *throughput* en Mbps. Este cálculo se realiza de forma independiente para cada UE conectado al *slice*.
2) Ajuste del max-rate: Cuando se toma una decisión, utilizando el algoritmo anteriormente mencionado, para ajustar los recursos de los *slices*, el script envía una petición a la API REST de Ryu para modificar los max-rate de las colas asociados a cada *slice*. Los identificadores de las colas y sus puertos asociados son descubiertos y validados via OVS y la API de Ryu.

En cuanto a la redirección del tráfico hacia el controlador, se optó por esta solución ya que, cuando los paquetes pasan por el switch, están encapsulados bajo la cabecera GTP (GPRS Tunelling Protocol) que tuneliza el tráfico entre el gNB y el UPF. Debido a esta encapsulación, OVS no puede "ver" la IP del UE, sino la del gNB o UPF, los *endpoints* del túnel, y por lo tanto OVS no puede distinguir el tráfico de cada UE por separado.

V. RESULTADOS

En esta sección se validará la efectividad de los mecanismos de adaptación de los *slices* dirigidos por SDN, bajo condiciones de tráfico reales, para demostrar las

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

capacidades de aislamiento de *slices* y de reserva dinámica de recursos del *testbed*.

En la primera evaluación, el tráfico corresponde a un flujo de Youtube, ya que es uno de los servicios VoD más utilizados y presenta un patrón de buffering bastante más claro que los demás, con picos de *throughput* cada cierto tiempo de manera constante. Para comprobar la evaluación, se comprobará que los picos de *throughput* son predichos y se reservan los recursos necesarios para el *slice* 1.

Otra de las evaluaciones, será la de congestionar el *slice* 2 con tráfico, y comprobar que el *slice* 1 tiene los recursos que necesita en todo momento, manteniendo así la Calidad de Servicio. El UE del *slice* 2 se trata de un UE emulado utilizando UERANSIM.

Por último, se mostrarán los resultados de predicción y clasificación del modelo LSTM implementado obtenidos con *datasets* de prueba de los distintos tipos de servicio VoD.

El modelo consta de tres capas LSTM de 100, 50 y 50 unidades ocultas respectivamente, a las que sigue una capa de *dropout*, las dos primeras de 0,2 y la última de 0,5, y una capa densa a la que le sigue una función de activación (ReLU) para mejorar el aprendizaje de la representación no lineal a la salida de las capa densa. Para finalizar, como el modelo realiza dos tareas a la vez, tendrá dos salidas por lo que el resultado de la función ReLU se introduce en una capa densa de una neurona, que va a ser la salida de la predicción del throughput, y también en una capa densa de `n` neuronas, donde `n` es el número de clases a predecir. Las funciones de pérdida empleadas son: Mean Absolute Error (MAE) para la predicción de throughput, y Cross Entropy para la clasificación.

A. Reserva de recursos

Para verificar la reserva de recursos, se ha reproducido un vídeo de YouTube en un smartphone conectado a la red 5G, concretamente al *slice* 1. En la Figura 2, se muestra el *throughput* del UE medido por el controlador, junto con el valor de max_rate de la cola QoS asignada al *slice* 1. Se aprecian claramente los picos de *throughput* característicos del tráfico de vídeo en streaming.

El modelo es capaz de predecir estas subidas, y mediante un script en Python que utiliza la API REST de Ryu, se puede modificar dinámicamente el parámetro max_rate de la cola QoS. De esta forma, es posible reservar los recursos necesarios para garantizar que el UE mantenga una conexión fluida y estable.

B. Calidad de servicio

Otro de los puntos cruciales de esta reserva de recursos es la de asegurar la calidad de servicio al *slice* 1. Como se observaba en el apartado anterior, el ancho de banda reservado para el este *slice* está regido por las predicciones del LSTM.

Para validar esta calidad de servicio, se ha utilizado un UE físico conectado al *slice* 1, y un UE emulado con UERANSIM conectado al *slice* 2. Ambos UEs ejecutan un iperf UDP: el UE físico con un bitrate de 30 Mbps (el ancho de banda máximo) y el segundo con un bitrate de 15 Mbps, el máximo que se le puede asignar al *slice* 2. Por lo tanto, ambos UEs estarán luchando por la cantidad máxima de recursos.

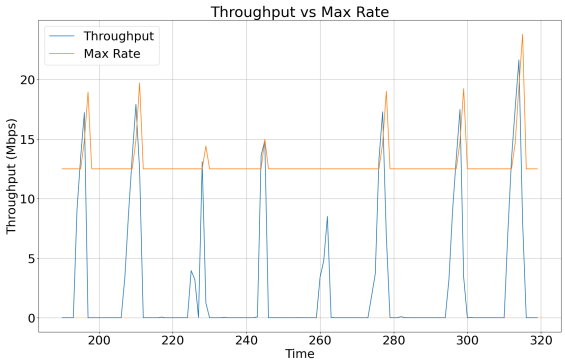


Fig. 2. Reserva de recursos mientras el UE consume un vídeo de YouTube

La Figura 3 ilustra los resultados: en la parte superior se muestra el *throughput* del UE físico junto con el parámetro max_rate de la cola del *slice* 1; en la parte inferior, se presenta la información correspondiente al UE emulado y su respectiva cola en el *slice* 2. Como se puede observar, el *slice* 1 obtiene una prioridad mayor en la asignación de recursos, mientras que el *slice* 2, es limitado por el ancho de banda del *slice* 1.

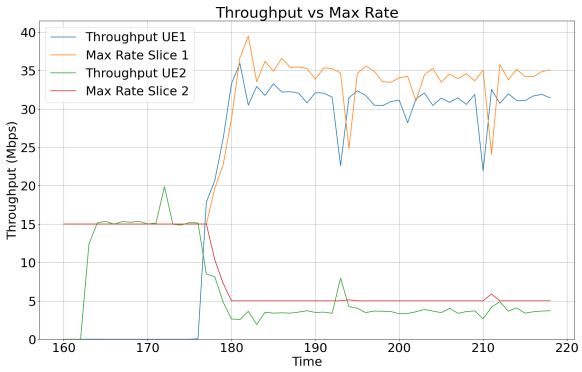


Fig. 3. Comparativa de los recursos de ambos slices y gestión del max-rate cuando el *slice* 1 requiere más recursos

C. Rendimiento del modelo

Como último punto, se ha evaluado el rendimiento del modelo en cuanto a la predicción de *throughput* y a la clasificación de servicios a través de las redes neuronales ya definida. Aunqe la clasificación del servicio no se ha utilizado para la reserva dinámica de recursos, es otra de las capacidades que nuestro modelo puede ofrecer.

Para la evaluación se han utilizado varios *datasets* que capturan el throughput usado de los distintos servicios VoD, los cuales el modelo no ha procesado anteriormente, de forma que se pueda probar la verdadera capacidad de abstracción del mismo. En la siguiente tabla, se exponen

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

los valores medios de RMSE (Root Mean Square Error), para la predicción del *throughput*, y la precisión media de la clasificación:

Tabla I
VALORES DE RMSE Y PRECISIÓN PARA LOS DISTINTOS SERVICIOS.

Servicio	RMSE	Precisión
YouTube	2.35	0.86
Twitch	3.35	0.98
TikTok	2.9	0.55
Prime Video	3.6	0.84

En cuanto a los valores de RMSE, se tratan de unos resultados aceptables, teniendo en cuenta que el rango de valores de los *datasets* va de 0 a unos 40 Mbps, por lo que la predicción del modelo, de media, está bastante cercana al valor real.

Por otra parte, la precisión a la hora de clasificar, obtiene unos valores bastante buenos en servicios como YouTube, Twitch o Prime Video, pero en contrario, Tiktok obtiene una precisión baja. Esto se debe a la propia naturaleza de estas aplicaciones, ya que YouTube, Twitch o Prime son plataformas de vídeos, o directos, largos, en los que el ancho de banda es utilizado de una manera consistente y periódica, ya que el usuario ve un mismo vídeo durante varios minutos, incluso horas. En cambio, TikTok, es una aplicación en la que constantemente se consumen vídeos cortos y muy distintos entre sí, cada pocos segundos, lo que hace que la utilización del ancho de banda sea mucho más aleatoria, y por ello, el modelo no es capaz de identificarlo de una forma más consistente.

VI. CONCLUSIONES Y TRABAJOS FUTUROS

En este artículo, se ha presentado el diseño e implementación de un *testbed open source* basado en SDN para *slicing* de red 5G *standalone*. Creado utilizando Proxmox VE para la virtualización, la plataforma integra Open5GS como *core* 5G, UERANSIM como simulador de RAN y srsRAN para conectividad via SDR, Ryu como controlador SDN, y Open vSwitch para la gestión del tráfico de los *slices*. Además se ha implementado un modelo LSTM capaz de gestionar los recursos de la red en tiempo real, además de poder clasificar el tráfico por distintos servicios VoD.

Nuestro *testbed* tiene la capacidad de:

- Emular *slices* 5G SA end-to-end sobre enlaces RAN físicos y emulados.
- Asignar dinámicamente el ancho de banda entre un *slice* de alta prioridad y un *slice* best-effort a través de Ryu y su API REST.
- Validar que las predicciones del LSTM reducen la congestión, y mejoran la utilización de recursos, así como la garantía de la calidad de servicio.

Por último, este trabajo se trata se un framework modular y reproducible, creado para que investigadores prototipen y comparen distintas políticas de control de *slices* bajo condiciones de tráfico 5G reales.

En cuanto a trabajos futuros, hay varios caminos que destacar: uno de ellos, es el de poder asignar dinámicamente *slices* según el tipo de tráfico que el LSTM haya clasificado, realizando así, un *slicing* aún más detallado, creando colas QoS para cada tipo de servicio. Por otro lado, explorar formas de mejorar el LSTM como introducir más parámetros de entrada y así mejorar la precisión de las predicciones, predecir más valores a futuro o explorar otros modelos distintos de Deep Learning, como GRU o Transformers y realizar una comparativa del rendimiento entre ellos. También sería interesante realizar más mediciones que permitan un mejor análisis del QoS, como la latencia o el *jitter*.

Otro de los factores que se podrían estudiar en un futuro sería la escalabilidad de la topología, así como del modelo LSTM, analizando el comportamiento con un número más elevado de UEs conectados. También cabe destacar la comparación entre tecnologías de contenedores e hipervisores para realizar este tipo de arquitecturas, siendo este último capaz de migrar servicios entre máquinas físicas de manera más rápida, siguiendo el paradigma de la hiperconvergencia.

Por último, se tiene planeado publicar los *datasets* utilizados así como todo el código fuente desarrollado utilizado para entrenar el modelo, para el controlador Ryu, y para la creación de los *datasets*.

AGRADECIMIENTOS

Esta publicación ha sido financiada en parte por el Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación (MICIU/AEI/10.13039/501100011033) a través de la ayuda PID2023-151462OB-I00. La publicación ha sido cofinanciada al 85% por la Unión Europea, Fondo Europeo de Desarrollo Regional, y la Junta de Extremadura. Autoridad de Gestión: Ministerio de Hacienda. Números de expedientes de las ayudas: IB24108 y GR24177. Está financiado también por el Plan de Investigación y Transferencia 2024 de la Universidad de Extremadura a través de la Acción V (Proyectos de Iniciación a la Investigación – AccionV-05).

REFERENCIAS

[1] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5g: Survey and challenges," IEEE Communications Magazine, vol. 55, no. 5, pp. 94–100, 2017.

[2] 3rd Generation Partnership Project (3GPP), "5g; management and orchestration; concepts, use cases and requirements," ETSI, ETSI Technical Specification TS 128 530 V16.3.0, October 2020, 3GPP TS 28.530 version 16.3.0 Release 16. [Online]. Available: https://www.etsi.org/deliver/etsi ts/128500 128599/128530/16.03.00 60/ts128530v160300p.pdf

[3] M. Moussaoui, E. Bertin, and N. Crespi, "Telecom business models for beyond 5g and 6g networks: Towards disaggregation?" in 2022 1st International Conference on 6G Networking (6GNet), Paris, France, 2022, pp. 1–8. [Online]. Available: https://hal.science/hal-04005533

[4] H. O. Otieno, B. Malila, and J. Mwangama, "Deployment and management of intelligent end-to-end network slicing in 5g and beyond 5g networks: A systematic review," IEEE Access, vol. 12, pp. 190 411–190 433, 2024.

[5] D. Pineda, R. Harrilal-Parchment, K. Akkaya, A. Ibrahim, and A. Perez-Pons, "Design and analysis of an open-source sdn-based 5g standalone testbed," in IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2023, pp. 1–6.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[6] M. Sulaiman, B. Sun, M. A. Salahuddin, R. Boutaba, and A. Saleh, "vNetRunner: Per-VNF Slice Modeling for 5G and Beyond Networks," in Proceedings of the IEEE/IFIP Network Operations and Management Symposium, Honolulu, HI, USA, May 2025.
[7] E. T. Owusu et al., "A Transformer-Based Deep Q-Learning Approach For Dynamic Load Balancing in Software-Defined Network," in 2025 IEEE, pp. 1–10
[8] 3GPP, "3gpp network slice management," https://www.3gpp.org/technologies/slice-management, accedido: 2025-04-20.
[9] S. Lee, "Open5GS documentation," https://open5gs.org/open5gs/docs/, accedido: 2025-05-27.
[10] A. Gungor, "Ueransim repository," https://github.com/aligungr/UERANSIM, accedido: 2025-05-27.
[11] S. R. Systems, "srsran documentation," https://docs.srsran.com/, accedido: 2025-05-27.
[12] Ryu SDN Framework Community, "QoS Ryubook," https://osrg.github.io/ryu-book/en/html/rest_qos.html accedido: 2025-05-27.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

5G-Enabled Industrial Mobile Robotics: Integration, Coordination, and Performance

Ignacio Rodriguez[1], Miguel Villanueva-Fernandez[2], Carmen Blason-Tronco[1], Preben Mogensen[2]
[1] Department of Electrical Engineering, University of Oviedo, Gijon 33203, Spain.
[2] Department of Electronic Systems, Aalborg University, Aalborg Øst 9220, Denmark.
irl@uniovi.es, mvf@es.aau.dk, UO277359@uniovi.es, pm@es.aau.dk

Industrial autonomous mobile robotic systems (I-AMRS) integrate mobile platforms with manipulators for advanced, on-the-move interactions within factory environments. This paper introduced the I-AMRS concept and discusses key aspects related to data traffic requirements, and IT/OT integration strategies considering 5G as the core communication enabler. The paper also proposed a real-time coordination method allowing for synchronized interaction between static and mobile robotic elements. Experimental validation was done over a private 5G network in industrial operational conditions, demonstrating reliable closed control-loop performance.

***Keywords*—Industry 4.0, Mobile Robots, 5G, Integration, Synchronization, Performance**

I. INTRODUCTION

The emergence of Industry 4.0 has transformed modern manufacturing and logistics by promoting the convergence of physical and digital systems through automation, data exchange, and real-time connectivity. Central to this transformation is the integration of cyber-physical systems (CPS), Internet-of-Things (IoT), and edge-cloud computing, which enable smarter, more responsive industrial environments [1]. Among the key enablers of this paradigm are mobile robotic systems, based on autonomous mobile robots (AMRs) or automated guided vehicles (AGVs), which are increasingly deployed to handle material transport, inspection, and other mission-critical tasks [2]. These systems require seamless orchestration and high levels of adaptability, which pose significant challenges in traditional industrial communication infrastructures characterized by latency, limited bandwidth, and inflexible architectures.

To address these challenges, 5G wireless communication has emerged as a foundational technology, offering enhanced mobile broadband (eMBB), ultra-reliable low-latency communication (URLLC), and massive machine-type communication (mMTC) capabilities. These features make 5G particularly well-suited for industrial mobile robotics, enabling flexible real-time support for different control data flows, dynamic coordination, and scalable operations. By bridging information technology (IT) and operational technology (OT) domains, 5G facilitates distributed intelligence, edge computing integration, and reliable high-capacity performance [3].

A. Industrial Autonomous Mobile Robotic Systems

Driven by the ongoing digital transformation and evolving industrial demands, traditional AMR/AGV platforms have evolved into more advanced and complex industrial autonomous mobile robotic systems (I-AMRS), by merging and integrating autonomous navigation and robotic manipulation capabilities, forming a versatile solution that enables smart production operations, by adding an extra degree-of-freedom to current standard mobile logistic platforms. Fig. 1 depicts an I-AMRS in operation, composed by an AMR and a robotic manipulator arm, with the aim of handling and transporting materials and/or manufacturing products within the factory.

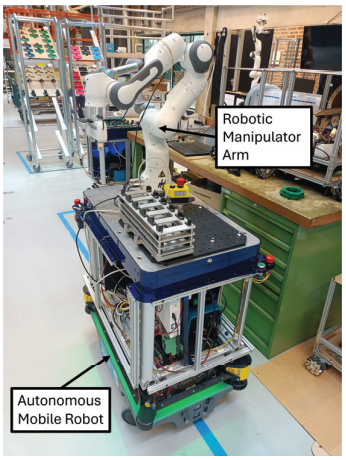


Fig. 1. Example of an advanced I-AMRS composed of an AMR equipped with a robotic manipulator arm.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

While conventional AGV/AMR elements typically follow fixed paths or predefined routes, these novel I-AMRS add a new layer of complexity by incorporating manipulation capabilities that allow interaction with objects and machinery. This enhancement increases flexibility and automation potential but also introduces new challenges across operational planning, and the communication and control domains (main focus of this paper).

From a planning perspective, I-AMRS require more sophisticated task allocation and coordination strategies. Unlike traditional AMRs that primarily handle point-to-point transport, I-AMRS are set to perform complex tasks that involve both mobility and interactions such as picking, placing, or assembling. This requires more advanced motion planning and synchronized mobile platform navigation and arm manipulation in dynamic environments. Additionally, these systems must be aware of and responsive to a broader range of constraints, including safety zones, shared workspaces, and task priorities, which complicates both offline programming and real-time decision-making [4].

Communication and control requirements also scale significantly. The I-AMRS mobile base and robotic manipulator generate distinct types of data traffic with different temporal and reliability quality-of-service (QoS) demands [5]. Further, I-AMRS demand tighter coupling between enterprise-level planning systems and real-time control layers. Traditional systems can function with relatively loose IT/OT alignment, as their workflows are limited and often predefined. However, I-AMRS must continuously exchange data between centralized systems (e.g., manufacturing execution systems, fleet managers) and distributed control units (e.g., motion controllers, sensor systems). This requires robust middleware communication capable of managing diverse protocols, synchronizing states across systems, and dynamically adapting to changes in production orders or floor layouts [6].

Under these constraints, 5G is identified as the optimal communication technology to support I-AMRS because it can simultaneously accommodate the distinct data traffic requirements of the robotic components, while ensuring reliable performance in industrial mobility conditions [7].

B. State of the Art

Few works have been reported analyzing integrated communications for I-AMRS solutions. [8] describes one of the first prototyping efforts involved in developing an AMR with a manipulator for carrying out tasks related to manufacturing. They refer to wireless technology as a key design aspect. [9] emphasizes the importance of decoupling and having independent I-AMRS sub-components, with different communication interfaces, to improve performance. In [5], results from an operational AGV case study are reported, analyzing the main protocols in use and identifying 5G as main candidate technology for operation, and suggesting duplication of wireless interfaces to reduce potential communication disruption in suboptimal deployments. [10] reports continuous mobile manipulator performance results with a focus on control algorithms. In [11], a prototype of a 5G-integrated mobile robotic manipulator was presented, including time-sensitive network (TSN) performance results with operational URLLC latencies of 1-2 ms and synchronization accuracies lower than 8 us. With respect to robotic operation synchronization, [12] presents a 5G-based distributed coordination method for static collaborative robots, and [13] reports a method for time-synchronized actions of mobile manipulators based on 5G latency boundary estimations. In [14] a demonstrator for on-the-move cooperation of a stationary manipulator and a mobile platform operating over 4G was reported, indicating a number of technological shortcomings and unreliable performance.

C. Objectives and Contributions

The main objective of the paper is to provide a reference for the understanding the operational characteristics of 5G-enabled I-AMRS. This paper goes beyond state of the art in I-AMRS, by providing specific insights into the use of 5G as an harmonization layer in IT/OT integration, and quantifying the communication requirements for the different components of I-AMRS. Further, a simple 5G-enabled coordination control method that allows for operational interaction of mobile I-AMRS with static machinery is proposed and validated in real-world industrial scenario.

The rest of the paper is organized as follows. Section II elaborates on the requirements and 5G integration matters. Section III describes the proposed 5G-based operational synchronization scheme. Section IV presents and discusses performance and validation results. Finally, Section V concludes the paper.

II. IT/OT INTEGRATION

Effective deployment of I-AMRS in operational industrial environments requires seamless integration between IT systems, responsible for enterprise-level functions such as production planning, analytics, and cloud services; and OT, which governs real-time control, actuation, and sensing at the shop floor level. Traditionally, IT and OT have operated in separated architectures, hindering end-to-end coordination. Here, the adoption of 5G as a unifying middleware infrastructure allows for the seamless support of multiple robotic control communication interfaces with different end-points and data traffic requirements.

Fig. 2 illustrates the role of a 5G network with edge-cloud capabilities in the integration of the different components of an I-AMRS, by simultaneously interfacing them with cloud-based IT systems and local OT controllers, enabling closed-loop control and high-level coordination within a unified framework. One of the most important elements in this architecture is the 5G gateway (GW) installed at the robotic device side. This device will act as main integration element within the 5G-based I-AMRS framework and should not be a simple 5G modem, but implement network intelligence logic, allowing for simple integration and routing of multiple sub-networks and data traffic flows to and from the edge-cloud [19]. As opposed to single-vendor closed proprietary commercial solutions,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Tabla I
TABLA DE EJEMPLO.

I-AMRS element	Hardware component	Control protocols	Avg. packet size [B]	Avg. data rate [kbps]
Industrial-grade AMR	MiR 200 [15]	TCP/HTTP/REST	4282	34.3
Industrial-grade cobot arm	UR5e [16]	TCP/RTDE	130	10.4
Educational AMR	Elephant Robotics myAGV [17]	TCP/HTTP/ROS	184	22.1
Educational robotic arm	Elephant Robotics myCobot 280 Pi [18]	TCP	60	7.2

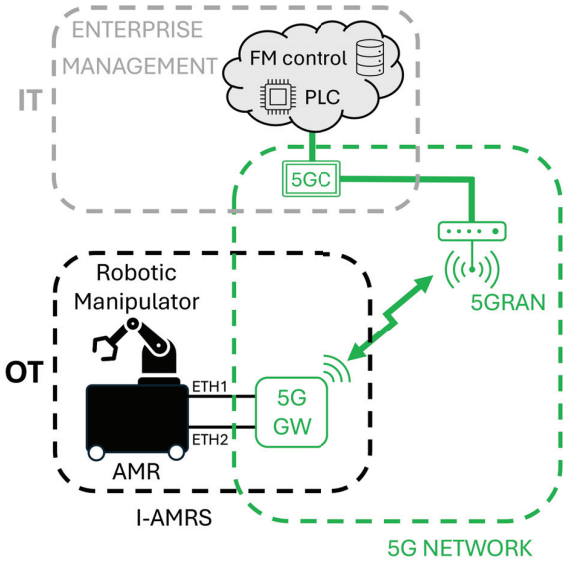


Fig. 2. Network architecture for 5G-based IT/OT integration of I-AMRS.

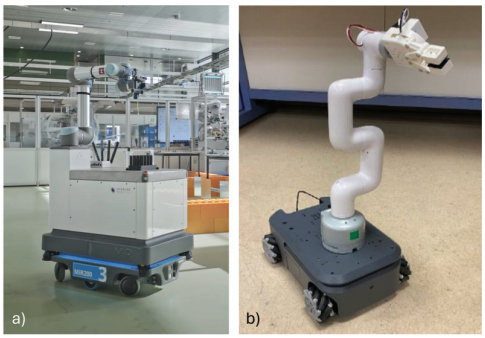


Fig. 3. Picture of the I-AMRS systems considered in this work: a) commercial industrial grade I-AMRS, b) educational I-AMRS.

such open device capabilities guarantee universal compatibility towards any individual I-AMRS elements, favoring the adoption of 5G technology in industrial brownfield deployments and multi-vendor settings [20].

A key challenge in this integration lies in the heterogeneous nature of control data traffic generated by I-AMRS. Navigation systems on the mobile base typically require continuous telemetry and mission updates, while robotic manipulators demand precise motion control. These two flows are essential for I-AMRS and can, therefore, be categorized as critical, requiring reliable low latency capabilities. Simultaneously, non-critical data such as diagnostics, video streaming, and task-level coordination may tolerate higher latencies but require guaranteed throughput. The ability of 5G to support differentiated QoS through edge computing integration makes it uniquely suited for managing these concurrent communication streams. This ensures that each functional layer of the I-AMRS receives the appropriate service level, enabling both high-frequency actuation and enterprise-level orchestration without compromising system integrity or safety.

Table I summarizes reference protocols and data traffic characteristics of both commercial industrial-grade I-AMRS components, and educational equivalents. These conform the integrated I-AMRS depicted in Fig. 3. In the case of the educational components, they are simply reported as a reference to illustrate the fact that, from a robot programming and networking perspective, there is a large degree of similarity with industrial components in specs and supported interfaces and protocols; which implies that industrial communication-oriented can be done in-lab with such equipment in a more cost-efficient manner, generating results that are relevant and generalizable to real-world industrial-grade versions [21]. All values are obtained from data traffic sniffing measurements performed over the communication interfaces of the different robotic elements. These are marked as ETH1 and ETH2 in Fig. 2.

In Section IV, selected performance results are presented, illustrating the control-loop latency (CCL) performance of the industrial-grade I-AMRS components reported in the table over operational 5G networks. For completeness of the work, 5G I-AMRS performance is also compared to that obtained over Wi-Fi 6 systems, which are typically considered as the de-facto wireless choice in industrial settings nowadays [3], [7].

III. COORDINATED REAL-TIME CONTROL

In the following, we propose a novel 5G and edge-cloud-based message exchange method for operational on-the-move coordination between mobile (e.g., I-AMRS) and static (e.g., production module) robotic entities. Differently from those proposed in [11], which relies on 5G Rel. 16 TSN support capabilities; and [13], which relies on statistical boundary estimation of expected 5G latency; our control method leverages on available operational parameters to coordinate robotic actions utilizing 5G and edge-cloud with coarse standard network time protocol (NTP) synchronization.

Since no specific 5G Rel. or network features are needed, our proposed solution guarantees compatibility with any 5G network. This is done by performing action-triggered message exchange at application layer. Fig. 4

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

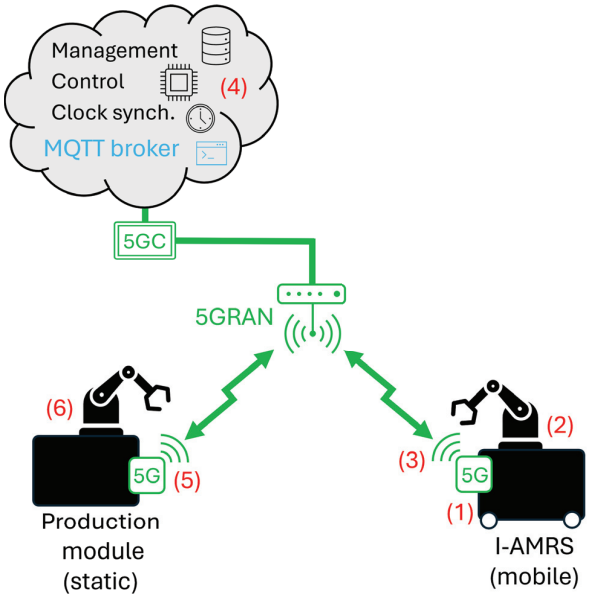


Fig. 4. Proposed network architecture for 5G-based coordination of static and mobile robotic machinery.

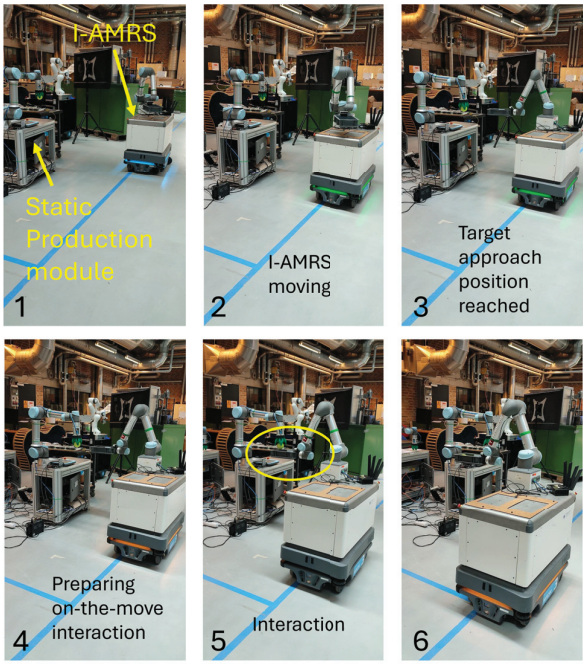


Fig. 5. Sequence of pictures of the functional demo implemented for validation of the proposed coordination method.

illustrates the reference network architecture of the proposed solution, including the main entities involved in the messaging:

- I-AMRS mobile platform (1)
- I-AMRS robotic manipulator (2)
- I-AMRS 5G gateway (3)
- Edge-cloud controller (4)
- Static machinery 5G interface (5)
- Static machinery robotic element (6)

For the implementation, message queuing telemetry transport is selected, given its good 5G performance [22]. The coordinated control is handled by the edge-cloud controller which is equipped with a MQTT broker, centralizing all message exchange between the robotic elements. To perform the actual message transmission, the different end points subscribe to relevant MQTT topics [23].

Although the developed method provides general interaction functionality, we exemplify its application by assuming a industrial use case where a robotic arm from a static robotic cell production module needs to handle a product or material piece on-the-move, in a synchronized manner, to a mobile I-AMRS. Our algorithm monitors, from the edge-cloud control, the relative I-AMRS position with respect to the static robotic element. We rely on 5G network performance for synchronization and reliable latency performance, allowing for timely signaling of the moment where the I-AMRS approached the static module. At this moment, the edge-cloud control releases the interaction movement orders to the two robotic arm end points on the I-AMRs and on the production module to begin their coordinated operation. This process is illustrated in Fig. 5, which captures a relevant sequence of events.

Our proposed messaging exchange sequence to enable coordinated interaction of mobile and static robotic elements is as follows:

1) Initialization:
   (4) → (3): MQTT Pub(AMR control)
   (3) → (1): POST (mission) / HTTP GET (position)
   (1) → (3): HTTP RE(position)
   ...
2) When target approach position reached:
   (3) → (4): MQTT Pub(”position!”)
   (4) → (5): MQTT Pub(”position!”)
   (5) → (6): RTDE Msg(set SDO)
3) Continue movement:
   (3) → (1): HTTP GET (position)
   (1) → (3): HTTP RE(position)
4) When target interaction position reached:
   (3) → (4): MQTT Pub(”interaction!”)
5) Triggering interaction (robot programs):
   (4) → (3): MQTT Pub(”target UR5e joints mobile”)
   (4) → (5): MQTT Pub(”Target UR5e joints static”)
   (3) → (2): RTDE Msg(”setreg:target-mobile”)
   (5) → (6): RTDE Msg(”setreg:target-static”)
   (2) → (3): RTDE Msg(”tmp:currentpos-mobile”)
   (6) → (5): RTDE Msg(”tmp:currentpos-static”)
   (3) → (4): MQTT Pub(”currentpos-mobile”)
   (5) → (4): MQTT Pub(”currentpos-static”)
   ...
6) When interaction position reached at mobile side:
   (3) → (4): MQTT Pub(”ready-mobile”)
7) When interaction position reached at static side:
   (5) → (4): MQTT Pub(”ready-static”)
8) Synchronized interaction is performed automatically.

As the method relies on ordered sequences of messages,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

the different network elements should implement time synchronization to guarantee causal operation. As no 5G-specific time synchronization features are leveraged in order to fulfill with our network-agnostic approach, we rely on simple optimized NTP methods, which are capable of providing accuracies over industrial 5G in the order of ± 0.1 ms at with a reliability of 99.9% [23]. In our implementation, we distribute NTP timing from the edge-cloud, as indicated in Fig. 4.

In order to validate our proposed method, apart from the functional validation demo in operation conditions illustrated in Fig. 5, a performance evaluation was carried out by monitoring operational closed control-loop latency (OCCLL) during different runs of the demo. In this case, the 5G performance was also benchmarked against that achieved over Wi-Fi 6. This is addressed in Section IV.

IV. PERFORMANCE EVALUATION

This section summarizes key 5G performance evaluation results, addressing those relevant aspects highlighted in Sections II and III for I-AMRS in an operational industrial research factory lab. The measurements were performed at the AAU 5G Smart Production Lab in Denmark [3] in a private industrial 5G standalone (SA) network, operating in the 3.7 GHz band (N78), with 100 MHz bandwidth, time division duplex (TDD) mode, and uplink/downlink ratio of 1/4 [24].

All reported measurement results have been done in industrial operational conditions, monitoring the I-AMRS communication during navigation and operation in the different tests. For individual characterization of the performance of the I-AMRS components, navigation CLL communication is monitored in mobile conditions and accounts for the full control loop between edge-cloud controller and mobile platform for telemetry request and/or mission update and the subsequent response from the mobile platform to the edge cloud with the telemetry data. This implies two transmissions over the 5G network (one in downlink and one in uplink). For the I-AMRS robotic manipulator, the CLL pattern is similar, but accounting for the full control loop between edge-cloud and robotic arm with operational instructions, and the reply from the arm to the edge-cloud with status and position update. In this case, the performance is evaluated in static conditions, to have a reference in different operational conditions.

For performance evaluation of the proposed mobile-static coordination method, OCCLL is considered instead. In this case, we monitor the latency experienced by the I-AMRS as a single entity (considering its multiple control flows) in executing the different communications with the edge-cloud controllers. This includes synchronization traffic, navigation control, robotic manipulator control, and all the signaling introduced by the proposed message exchange algorithm. As done for the CLL, for the OCCLL we account for the full control loop latency including requests and responses from the different control flows.

The following CLL performance results are based on over 10.000 sample values. However, the OCCLL ones are computed over approximately 500 samples only, due to operational constraints and limitations.

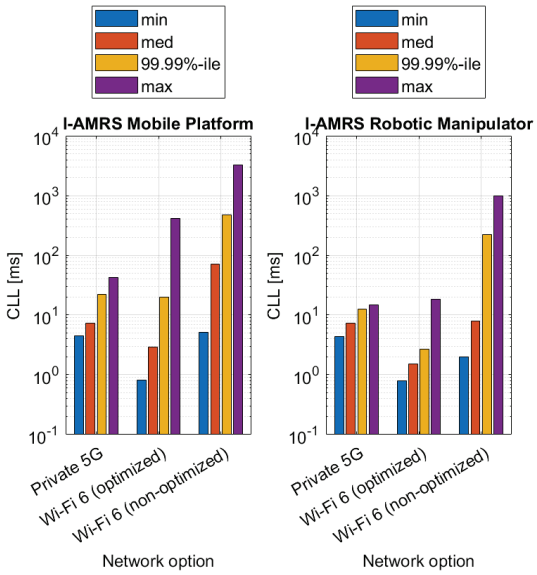


Fig. 6. CLL performance for the different I-AMRS components in mobile and static conditions for the different network configurations explored.

A. 5G Latency Performance for Static and Mobile I-AMRS Robotic Elements

Fig. 6 presents the CLL performance results for the I-AMRS mobile platform in mobility conditions (left), and for the I-AMRS robotic manipulator arm in static conditions (right). Both sets of results include minimum, median, 99.99%-ile, and maximum CLL values not only for private 5G, but also for Wi-Fi 6 in optimized conditions (single user and no interference from overlapping Wi-Fi channels, which is a non-realistic assumption), and non-optimized conditions (multiple users and co-channel interference from neighboring access points, which is close to the settings experienced in real-world industrial deployments nowadays).

For the I-AMRS mobile platform, the CLL performance over private 5G is well bounded below 50 ms 100% of the time, and below 25 ms if the 99.99%-ile is considered. A median CLL of 7 ms was observed. This illustrates the robustness and reliability of 5G when operating under mobility conditions and roaming between access nodes. Wi-Fi 6 offers comparable to private 5G in the optimized case, except for the maximum value, which exceeded 400 ms due to the suboptimal roaming procedures between access points, which clearly affects the link reliability performance in mobility conditions. In non-optimized conditions, Wi-Fi 6 performance is further degraded in mobility conditions, largely exceeding 500 ms above median.

For the I-AMRS robotic manipulator in static conditions, Wi-Fi 6 in optimized ideal conditions offers the best

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

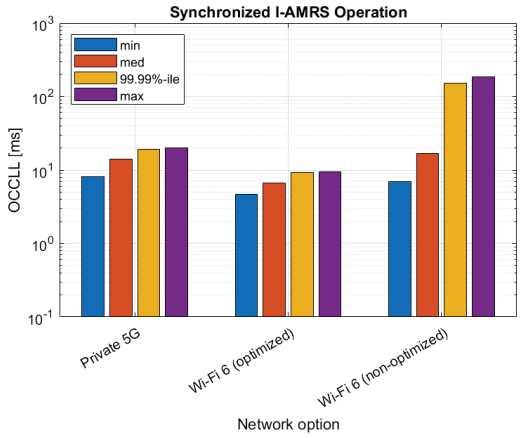


Fig. 7. OCCLL performance experience by the I-AMRS in operational conditions when operating the proposed coordination scheme for the different network configurations considered.

CLL performance with valued consistently bounded below 5 ms in 99.99% of the cases. As explained above, it is not realistic to consider that this performance is achievable in operational industrial scenarios, where the situation is closer to the non-optimized Wi-Fi 6 case. In that case, CLL performance exhibits a median value of 8 ms, which is enlarged to over 225 ms above the 99.99%-ile. In this static case, 5G leads to bounded performance below 15 ms in 100% of the cases, with a median value of 7 ms.

B. Operational Coordinated I-AMRS Performance

The OCCLL performance results obtained during the multiple validation runs of the coordination test are displayed in Fig.7. In this case, the latency performance experienced by the I-AMRS in operational conditions when operated over private 5G ranges from minimum 8 ms to maximum 20 ms, with a median of 19 ms. This can be described as a very robust and reliable performance for the selected use case. In this case, optimized Wi-Fi 6 in ideal conditions allows for the best use case performance with OCCLL values below 10 ms but, as explained earlier, it illustrates a non-realistic deployment option, only valid as benchmarking reference. Realistic non-optimized Wi-Fi 6 performance was observed to be worst than the private 5G one with a median value of 16 ms and a maximum value of 186 ms.

These results clearly illustrates the superior capabilities of private 5G for reliable operation of I-AMRS, handling control communication in a robust and reliable manner with bounded and stable control-loop latencies.

V. CONCLUSIONS

This paper introduced the concept of industrial autonomous mobile robotic systems (I-AMRS), which are advanced robotic entities with integrated navigation and manipulation capabilities. The data traffic requirements for the different I-AMRS components was analyzed and IT/OT integration strategies were discussed in the context of 5G technology as the main integration element. The paper also highlighted the advanced capabilities of 5G-based I-AMRS, capable of performing complex on-the-move interactions with other stationary factory machinery. This was done through a proposed synchronized real-time method leveraging coarse NTP network synchronization and edge-cloud-centralized message exchange. The proposed scheme was validated and evaluated in terms of closed control-loop performance over a private 5G network in operational conditions and benchmarked against that one obtained for Wi-Fi 6 over the exact same settings.

Performance evaluation results indicated that I-AMRS closed control loops over 5G in operational conditions are robust, reliable and well bounded below 20-50 ms. This is not the case for Wi-Fi 6, which struggles to support mobility resulting in observed closed control loop latencies above 186 ms. These results highlight the superior capabilities of private 5G for the reliable support of I-AMRS in operational conditions.

For future work, improved time synchronization methods will be explored, allowing for coordination of more demanding use cases involving industrial machinery with ultra-fast production loops. Further, use case scalability will be analyzed, exploring not only multi-robotic entity interactions, but analyzing 5G-based I-AMRS scalability in scenarios where more computing functionalities are offloaded to the edge-cloud.

ACKNOWLEDGMENT

This work was supported in part by the Spanish Ministry of Science, Innovation, and Universities through project PID2024-160783OA-I00, funded by MICIU/AEI/10.13039/501100011033 and by the European Regional Development Fund (ERDF); in part under Ramon y Cajal Fellowship number RYC-2020-030676-I, funded by MICIU/AEI/10.13039/501100011033 and by the European Social Fund ”Investing in your future”; in part by the Government of Asturias-Sekuens/ERDF under grant IDE-2024-000693; and, in part by Innovation Fund Denmark through the project ”5G-ROBOT - 5G Enabled Autonomous Mobile Robotic Systems”.

REFERENCES

[1] B. Chen et al., ”Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges,” in IEEE Access, vol. 6, pp. 6505-6519, 2018.

[2] E. A. Oyekanlu et al., ”A Review of Recent Advances in Automated Guided Vehicle Technologies: Integration Challenges and Research Areas for 5G-Based Smart Manufacturing Applications,” in IEEE Access, vol. 8, pp. 202312-202353, 2020.

[3] I. Rodriguez et al., ”5G Swarm Production: Advanced Industrial Manufacturing Concepts Enabled by Wireless Automation,” in IEEE Communications Magazine, vol. 59, no. 1, pp. 48-54, January 2021.

[4] S. Valiollahi et al., ”Experimental Evaluation and Modeling of the Accuracy of Real-Time Locating Systems for Industrial Use,” in IEEE Access, vol. 12, pp. 75366-75383, 2024.

[5] A. -L. Kampen et al., ”Case Study of AGV in Industry 4.0 Environments – An Evaluation of Wireless Communication Protocols,” 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 2022, pp. 2049-2055.

[6] M. Wollschlaeger et al., ”The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0,” in IEEE Industrial Electronics Magazine, vol. 11, no. 1, pp. 17-27, March 2017.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[7] P. Mogensen, and I. Rodriguez, "5G for Smart Production," in The Future of Smart Production for SMEs. Springer, Cham, 2022, pp. 327-333.
[8] S. Datta et al., "Development of autonomous mobile robot with manipulator for manufacturing environment," in International Journal of Advanced Manufacturing Technologies, vol. 38, pp. 536–542, 2008.
[9] A. Baldassarri et al., "Design of a Reconfigurable Mobile Collaborative Manipulator for Industrial Applications," in ASME Journal of Computational and Nonlinear Dynamics, vol. 18, no. 9, 091006, September 2023.
[10] O. Aboul-Enein et al., "NIST AMS 100-57: Continuous Mobile Manipulator Performance Measurement Data", National Institute of Standards and Technology, January 2024. Accessed: 10 June 2025 [Online]. Available: https://doi.org/10.6028/NIST.AMS.100-57
[11] P. Kehl et al. "Prototype of 5G Integrated with TSN for Edge-Controlled Mobile Robotics," in Electronics, vol. 11, no. 11, 1666, May 2022.
[12] A. E. Celik et al., "Decentralized System Synchronization among Collaborative Robots via 5G Technology," in Sensors, vol. 24, no. 16, 5382, August 2024.
[13] W. Zhang et al., "Implementation of Wireless Synchronous Robotic Control in 5G and Edge-Cloud Environments," 2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2025, pp. 1-4.
[14] N. Enayati et al., "Using Cellular Connectivity for On-the-move Cooperation of Stationary Manipulator and Mobile Platform," 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vasteras, Sweden, 2021, pp. 1-8.
[15] HowToRobot, "MiR200 Technical Specifications", December 2021. Accessed: 10 June 2025 [Online]. Available: https://howtorobot.com/sites/default/files/2021-12/MiR200_Brochure_en.pdf
[16] Universal Robots, "UR5e", 2025. Accessed: 10 June 2025 [Online]. Available: https://www.universal-robots.com/products/ur5e/
[17] Elephant Robotics, "myAGV 280 Pi 2023", December 2024. Accessed: 10 June 2025 [Online]. Available: https://www.elephantrobotics.com/en/myagv-2023-pi-en/
[18] Elephant Robotics, "myCobot 280 Raspberry Pi 2023", December 2024. Accessed: 10 June 2025 [Online]. Available: https://www.elephantrobotics.com/en/mycobot-280-pi-2023-en/
[19] I. Rodriguez et al., "An Experimental Framework for 5G Wireless System Integration into Industry 4.0 Applications," in Energies, vol. 14, no. 15, 4444, July 2021.
[20] M. Gundall et al., "Introduction of a 5G-Enabled Architecture for the Realization of Industry 4.0 Use Cases," in IEEE Access, vol. 9, pp. 25508-25521, 2021.
[21] C. Blason-Tronco, "Control of Autonomous Mobile Robotic Syystems using 5G", B.Sc. Thesis, University of Oviedo, February 2025.
[22] D. A. -C. Rincon et al., "An Operational 5G Edge Cloud-Controlled Robotic Cell Environment Based on MQTT and OPC UA," 2023 21st International Conference on Advanced Robotics (ICAR), Abu Dhabi, United Arab Emirates, 2023, pp. 7-14.
[23] M. Villanueva-Fernandez, "Wireless Control and Coordination of Industrial Robots using 5G", M.Sc. Thesis, University of Oviedo, July 2024.
[24] S. B. Damsgaard et al., "Commercial 5G NPN and PN Deployment Options for Industrial Manufacturing: An Empirical Study of Performance and Complexity Tradeoffs," 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Toronto, ON, Canada, 2023, pp. 1-7.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

Computación Autonómica para el Uso Eficiente de Energía en Redes de Sensores Destinadas a la Monitorización Ambiental

Miguel Zaragoza-Esquerdo, Sandra Sendra, Alberto Ivars-Palomares, Jaime Lloret.
Instituto de Investigación para la Gestión Integrada de zonas Costeras,
Universitat Politècnica de València,
Camino de vera, s/n. 46022, Valencia (España)
mizaes2@epsg.upv.es, sansenco@upv.es, aivapal@epsg.upv.es, jlloret@dcom.upv.es.

La monitorización de entornos naturales es hoy en día un tópico de investigación candente en tanto, que la detección temprana de una anomalía o la presencia de contaminantes en dicho escenario permite la mitigación de los daños causados en el entorno. En este sentido, para llevar a cabo dichas tareas, hasta ahora, se ha trabajado con nodos que ejecutaban los códigos de manera secuencial, principalmente por su incapacidad de ejecutar varios hilos. Sin embargo, en este artículo proponemos el uso de nodos sensores de manera colaborativa dentro de una red, donde un único nodo sea polivalente y pueda prestar servicio monitorizando diferentes sets de variables para un mismo entorno. Para tal fin proponemos el uso de la computación autonómica, es decir, un modo de autogestión de los nodos que les permite interactuar entre ellos y ejecutar diferentes códigos, según sus necesidades. Tras testear el sistema se ha podido comprobar que este tipo de funcionamiento genera importantes ahorros de energía en cada uno de los nodos integrantes de la red y que finalmente, es posible trabajar con un número menor de nodos, lo que implica un consecuente ahorro económico.

***Palabras Clave*- ahorro de energía, redes de sensores, computación autonómica, monitorización ambiental.**

I. Introducción

Los entornos marino-costeros constituyen ecosistemas de gran valor ecológico, económico y social. Son zonas altamente dinámicas y vulnerables, donde confluyen actividades humanas intensivas como el turismo, la pesca, la navegación y el vertido de aguas residuales tratadas a través de emisarios submarinos. A pesar de los procesos de depuración, estos vertidos pueden contener contaminantes químicos, biológicos y físicos que alteran la calidad del agua y amenazan la biodiversidad marina, provocando fenómenos como la eutrofización, la degradación de hábitats sensibles o la proliferación de especies invasoras [1].

En este contexto, la monitorización continua y multiparamétrica de la calidad del agua en zonas costeras y en las salidas de emisarios se vuelve una herramienta esencial para:

- Detectar a tiempo niveles anómalos de contaminación.
- Evaluar la eficacia de los sistemas de tratamiento.
- Prevenir impactos ecológicos irreversibles.
- Apoyar la toma de decisiones en la gestión ambiental y la planificación costera.

Los métodos tradicionales de muestreo manual y análisis en laboratorio presentan limitaciones importantes, como los altos costes, discontinuidad en el tiempo y la poca adaptabilidad a condiciones dinámicas del medio marino.

Las tareas de monitorización en tiempo real de dichos entornos pueden llevarse a cabo mediante el uso de redes de sensores. Estas redes suelen estar compuestas por múltiples dispositivos distribuidos espacialmente que recopilan datos ambientales como temperatura, humedad, presión atmosférica, calidad del aire o del agua, suelen estar conectados a través de tecnologías inalámbricas como Wi-Fi, Zigbee, LoRa, redes celulares o incluso usando ondas acústicas, si se hacen en entorno subacuáticos [2]. Los datos recogidos por cada nodo se transmiten a una unidad central, como una estación base o servidor en la nube, donde se almacenan, procesan y analizan. Cada uno de estos nodos, se focaliza en la recolección de datos medioambientales, o de calidad del agua, o del suelo, pero resulta complicado tener un dispositivo capaz de hacerlo todo a la vez, según sus necesidades. Por tanto, para monitorizar diferentes escenarios, en un entorno natural se requeriría de diversos nodos, cada uno funcionando de manera individual y comunicándose con sus homólogos. Tener un número elevado de nodos, incrementaría, así mismo, el consumo de energía de cada uno de los nodos y de la red global.

La mayoría de estos dispositivos, como por ejemplo los basados en el integrado ESP32, a pesar de tener varios núcleos [3], están pensados para llevar a cabo una única de monitorización en un entorno concreto dado que únicamente son capaces de ejecutar un hilo en uno de sus núcleos procesadores. El segundo núcleo suele estar ocupado en las tareas relacionadas con la conexión inalámbrica.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Para superar estas limitaciones de funcionamiento de estos tipos de dispositivos nace el concepto de computación autonómica [4]. La computación autonómica [5], inspirada en el sistema nervioso autónomo humano, permite que cada nodo se autogestione con mínima intervención humana, a través de cuatro propiedades clave:

- **Auto-configuración**: integración automática del nodo en la red y adaptación a las condiciones del entorno.
- **Auto-optimización**: ajuste dinámico del consumo energético y de la frecuencia de muestreo para maximizar la eficiencia.
- **Auto-recuperación**: detección y corrección de fallos, o reasignación de funciones entre nodos.
- **Auto-protección**: identificación y mitigación autónoma de amenazas de seguridad.

En este sentido, y basándose en sus principios de funcionamiento, el sistema sensor que daría respuesta a estos problemas sería el desarrollo de un dispositivo hardware capaz de autogestionarse y decidir qué debe monitorizar en cada momento.

Por tanto, en este artículo se propone el desarrollo de una red de sensores inalámbrica conformada de nodos capaces de trabajar en computación autonómica y distribuida para la mejora de la eficiencia energética de la red. Tras diseñar y programar los nodos, estos son evaluados en término de consumo de energía y se compara su funcionamiento con el de un nodo ejecutando un código secuencial. Finalmente se miden los anchos de banda que consumiría un nodo en la red, trabajando en el sistema propuesto.

El resto del articulo se estructura como sigue; la sección II presenta una serie de trabajos relacionados con el tópico de este artículo. La sección III presenta la arquitectura de nuestro sistema y su funcionamiento. En la sección IV se muestran los resultados sobre consumo de energía y el tráfico de red que generaría un nodo de este tipo y, finalmente la sección V muestra las conclusiones y líneas futuras a seguir.

II. Trabajos Relacionados

Esta sección presenta diversos trabajos previos relacionados con el desarrollo de redes de sensores inalámbricas y sistemas de comunicación autonómica.

El concepto de los modelos MAPE-K no son nuevos. De hecho, IBM [6] ya basaba sus procedimientos en este modelo para tener una computación más eficiente. Siguiendo esta idea E- Rutten et al. [7] propone un enfoque semiautomático para la generación de modelos de objetivos utilizando inteligencia artificial generativa, específicamente grandes modelos de lenguaje (LLMs), en combinación con el ciclo MAPE-K. De acuerdo con los resultados, la combinación de ambas tecnologías mejora la definición estructurada de objetivos en sistemas auto adaptativos, abordando problemas comunes como la ambigüedad, la inconsistencia y la subjetividad en la especificación de requisitos. Se realizaron estudios de caso, como el de un sistema de gestión bibliotecaria, donde múltiples expertos simulados por IA colaboran para construir y refinar el modelo. Los resultados muestran que la IA generativa puede producir modelos útiles, aunque más generalistas.

Por otra parte, la idea de realizar una monitorización cooperativa o colaborativa entre nodos presentes en un escenario complejo como puede ser un entorno marino o rural se trata en diferentes trabajos previos. Por ejemplo, G. Fortino et al. [8] proponen el diseño e implementación de sistemas IoT mediante objetos inteligentes cooperativos orientados a agentes. A través de la metodología ACOSO (Agent-based COoperating Smart Objects) y su middleware asociado, los autores presentan una arquitectura que permite a los objetos físicos actuar como agentes autónomos capaces de percibir, razonar, comunicarse y cooperar. Esta propuesta destaca por su capacidad de escalar desde aplicaciones pequeñas, hasta entornos complejos gracias a su flexibilidad y adaptabilidad.

Finalmente, hay que destacar que la elección de un entorno natural, en concreto el entorno marino no es casual, y es que las zonas marítimo-costeras son hoy en día las grandes maltratadas. Por ello, podemos ver como S. Sendra et al [9]. Trabajan de manera continuada en la propuesta de sistemas de monitorización de zonas costeras mediante entornos colaborativos basados en grupos, donde nodos sensores distribuidos trabajan de forma coordinada para recopilar y compartir datos ambientales. El sistema presentado se centra en mejorar la eficiencia de los sistemas de monitorización mediante la cooperación entre nodos, lo que permite una mayor cobertura, redundancia y resiliencia frente a fallos. Como muestran los resultados, el diseño del sistema le permite adaptarse dinámicamente a las condiciones del entorno y a la disponibilidad de recursos, lo que lo hace especialmente útil en contextos marinos, donde las condiciones pueden ser cambiantes y hostiles.

Como se observa, existen referencias donde se tratan los diferentes conceptos que aglutina la propuesta. Sin embargo, no se han encontrados sistemas que integren todos estos conceptos en un solo diseño. Por ello, se cree que la implementación de los modelos MAPE-K en la monitorización de los entornos marítimo-costeros, además de reducir el consumo energético global de toda la red, también mejoraría la precisión en la toma de decisiones para mitigar problemas de contaminación en el agua.

III. Arquitectura Propuesta

Esta sección presenta la arquitectura de red propuesta, así como el funcionamiento de cada uno de los nodos dentro de la red. Finalmente presenta el sistema de comunicación de utiliza nuestra propuesta y cómo se realiza en intercambio de información entre los diferentes módulos.

A. Arquitectura interna de un nodo con capacidad de cómputo autonómica

A diferencia de un nodo convencional, un nodo con capacidad autonómica es capaz de autogestionarse y tomar decisiones por él mismo sin la necesidad de intervención humana. Así mismo, un nodo con capacidad de cómputo autonómica es capaz de gestionar los recursos hardware necesarios para realizar las tareas de monitorización necesarias. Así, en lugar de tener 3 o 4 nodos para monitorizar cada tipo de medida, un único nodo sería capaz de realizar todas las tareas, en función del estado del entorno, lo que, entre otros generaría un importante ahorro de energía.

Para entender cómo funciona internamente un nodo podemos tomar como ejemplo la Fig 1. En ella podemos ver que se compone de varias series de sensores y actuadores que serán invocados para llevar a cabo las tareas de medida o acción sobre el entorno, en función de lo que en él suceda [10].

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

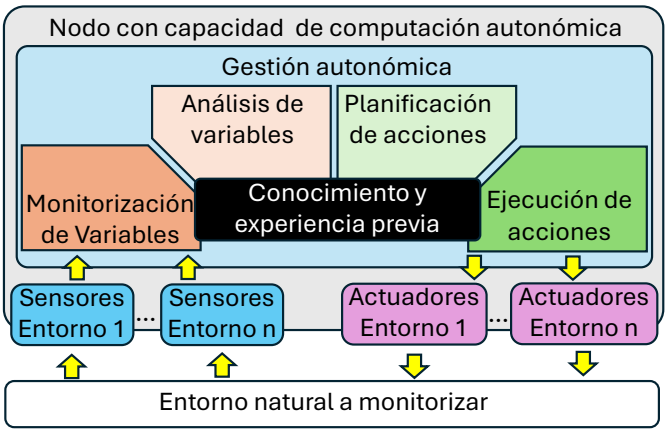


Fig. 1. Diagrama de bloques con capacidad de cómputo autonómica

Un sistema de computación autonómica se basa en un nodo central con capacidad de tomar decisiones de forma autónoma, integrando todos los módulos necesarios para adaptarse dinámicamente a los cambios del entorno. Cada uno de estos módulos, tiene asociado una serie de tareas y responsabilidades y esta manera de organizar las tareas suele recibir el nombre de modelo MAPE-K (Monitor, Analyze, Plan, Execute, Knowledge) [5].

El modelo MAPE-K es una arquitectura fundamental en los sistemas de computación autonómica, diseñada para permitir que un sistema gestione sus propios recursos y se adapte automáticamente a los cambios del entorno. Su implementación práctica implica la integración de sensores, algoritmos de análisis, motores de decisión, actuadores y una base de conocimiento (normalmente una IA previamente entrenada) que guía el comportamiento del sistema.

El primer componente Gestor autonómico, se encarga de la recolección continua de datos relevantes del entorno o del propio sistema. Esto se logra mediante sensores físicos (como temperatura, humedad o movimiento) o sensores lógicos (como uso de CPU o tráfico de red). Una vez recogidos los datos, el módulo Análisis de variables los interpreta para detectar patrones, anomalías o tendencias. Este análisis puede realizarse mediante reglas heurísticas, algoritmos de análisis de datos o incluso modelos de aprendizaje automático. Por ejemplo, en nuestro caso (ver Fig. 2) se puede identificar la presencia de un contaminante o in incremento de temperatura repentino, debido a la presencia de aguas residuales en las zonas circundantes

El siguiente paso es la planificación de acciones, donde el sistema decide qué acciones tomar en respuesta al análisis. Esto puede implicar el uso de motores de reglas como lógica difusa o reglas discriminantes previamente configuradas. Por ejemplo, si se detecta que la presencia mayor concentración de fitoplancton o zooplancton en las zonas cercanas a la salida, el sistema puede planificar una activación del sistema de eliminación de biofouling.

El módulo de ejecución de acciones se encarga de llevar a cabo las acciones planificadas. Esto puede implicar el control de actuadores físicos o la ejecución de scripts automatizados.

Finalmente, el componente de conocimiento y experiencia previa actuaría como base de datos de conocimiento que almacena reglas, políticas, datos históricos y modelos predictivos. Esta información resultaría especialmente importante a medio y largo plazo para mejorar la toma de decisiones frente a diferentes tipos se situaciones, permitiendo que el sistema aprenda de experiencias pasadas y se adapte de forma más inteligente.

Un caso de uso práctico y muy indicado podría ser la monitorización en el mal de la salida de aguas tratadas procedentes de depuradoras (ver Fig. 2). En la salida al mar de un emisario submarino de una planta depuradora, se instalan varios nodos sensores inteligente con capacidad de computación autonómica basado en el modelo MAPE-K. Además, estos nodos (Nodo 1, nodo 2 y nodo 3) son capaces de comunicarse entre ellos y hacerlo de manera cooperativa, de manera que, en función de las anomalías detectadas por sus vecinos, cada nodo habilitarías únicamente una serie de sensores, en lugar de todos ellos. El Nodo 3 detectaría la presencia de actividad en el emisario (lo que indicaría que se están vertiendo aguas al mar) y activaría los sensores de turbidez y temperatura y generaría una señal de aviso a las boyas (Nodo 1 y Nodo 2) más cercanas para que comenzases a monitorizar la presencia de contaminantes en el agua. A través del ciclo MAPE-K, el sistema monitorizaría estas variables, únicamente cuando se detecta actividad en el emisario; analizaría los datos para detectar anomalías, planificaría acciones correctivas como alertas o ajustes en la frecuencia de muestreo, y ejecutaría dichas acciones mediante actuadores o comunicación con las estaciones base. Sin embargo, cuando no existiese actividad en el emisario, los Nodos 1 y 2, podría dedicarse a monitorizar, por ejemplo, variables meteorológicas o estado del mar. Finalmente, toda la información generada y recogida durante este tiempo, quedaría almacenada en una base de conocimiento que permitiría al sistema aprender de experiencias pasadas y mejorar su capacidad de respuesta para futuras acciones, garantizando así una vigilancia ambiental eficiente y adaptativa del ecosistema marino.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

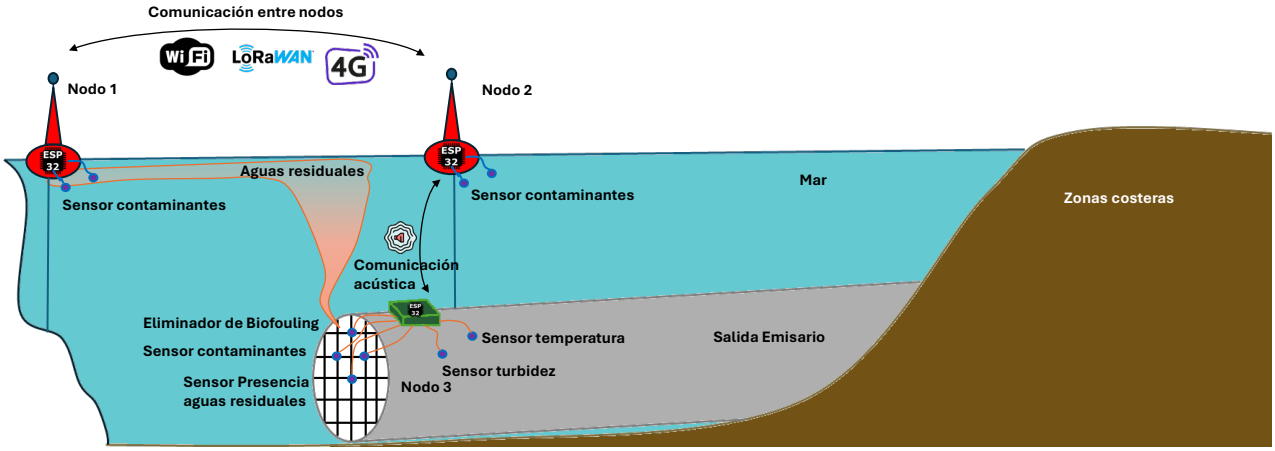


Fig. 2. Esquema de la topología de res desplegada en las zonas circundantes de un emisario marino

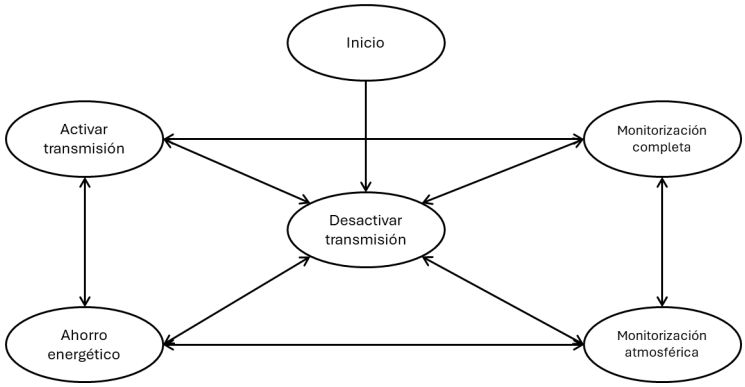


Fig. 3. Diagrama de estados de la computación distribuida

```
INICIAR configuración Conexión_Inalámbrica y MQTT
INICIAR sensores: DHT11, MQ-135, BME280, BH1750, FZ0430, TA12-200, B_LUX_V30B

CONFIGURAR temporizador de 1 segundo con interrupción
ESPERAR mensajes entrantes MQTT con valores de estado (0x0000 a 0x0005)

LOOP PRINCIPAL:
  SI MQTT no está conectado:
    REINTENTAR conexión Wi-Fi y MQTT

  SI la bandera del temporizador está activa:
    LIMPIAR bandera
    EVALUAR estado actual:
      CASO "0x0000": LLAMAR MonitorizacionEntorno()
      CASO "0x0001": LLAMAR ActivarTransmision()
      CASO "0x0002": LLAMAR DesactivarTransmision()
      CASO "0x0003": LLAMAR MonitorizarTransmision()
      CASO "0x0004": LLAMAR Sleep()
      CASO "0x0005": LLAMAR MonitorizacionCompleta()

FUNCIÓN MonitorizacionEntorno():
  LEER temperatura y humedad del DHT11
  LEER presión del BME280
  LEER iluminación del B_LUX_V30B
  LEER concentración de gases del MQ-135
  ENVIAR datos al servidor meteorológico por HTTP POST

FUNCIÓN ActivarTransmision():
  ACTIVAR relé de salida
  REGISTRAR tiempo de inicio
  ESTABLECER bandera de transmisión activa

FUNCIÓN DesactivarTransmision():
  DESACTIVAR relé
  LIMPIAR tiempo de transmisión
  DESACTIVAR bandera de transmisión

FUNCIÓN MonitorizarTransmision():
  LLAMAR ActivarTransmision()
  MEDIR corriente RMS desde TA12-200
  LEER voltaje desde FZ0430
  ENVIAR datos al servidor Tesla por HTTP POST

FUNCIÓN Sleep():
  ENTRAR en modo de suspensión ligera por 1 segundo

FUNCIÓN MonitorizacionCompleta():
  LLAMAR ActivarTransmision()
  EJECUTAR MonitorizacionEntorno()
  MEDIR corriente y voltaje
  ENVIAR datos combinados a ambos servidores

FUNCIÓN medirCorrienteAC():
  LEER múltiples muestras analógicas del TA12-200
  CALCULAR valor RMS eliminando offset DC
  CONVERTIR voltaje RMS a corriente usando la sensibilidad del sensor

FUNCIÓN callback(topic, message):
  EVALUAR mensaje recibido en el tópico MQTT
  ACTUALIZAR estado del sistema según el código recibido
```

Fig. 4. Pseudocódigo del funcionamiento de la computación autonómica

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

B. Funcionamiento de la computación autonómica en el nodo

La computación distribuida en el nodo se basa en que el sistema combine diferentes tipos de entornos de procesamiento para trabajar juntos de forma coordinada. Para ello se propone diferentes estados o tareas que se requiera que haga el nodo y se elija la tarea de computación necesaria para ese momento basándose en una máquina de estados finitos. La elección de los estados y control se produce mediante el módulo gestor, usando como elemento discriminante la base de datos de conocimiento.

En la Fig. 3 se muestra el diseño de los estados, donde cuando se inicia el dispositivo se establece el estado de transmisión desactivada y desde éste, el nodo detector de la anomalía le indique a qué estado debe pasar el nodo de monitorización, como, por ejemplo, monitorizar, activar la transmisión, monitorización de la transmisión, monitorización completa o el estado de ahorro energético. De esta forma, se consigue que la computación sea optima, se reduzca el tiempo de computación y se ahorre energía dado que el código del nodo ya no se ejecutaría de manera secuencial, como de costumbre, sino únicamente una pequeña parte de él. La comunicación de las alertas entre los nodos se realiza mediante instrucciones MQTT.

La Fig. 4 muestra el pseudocódigo, para un nodo meteorológico, que da lugar a que la computación de las tareas o estados se comporte de manera colaborativa y que el comportamiento del dispositivo sea autonómico. En el código, se observar que hay una serie de funciones o tareas que mediante un sistema de banderas se elige qué estado se activa. Estas banderas pueden cambiar cuando se activa el callback de que ha llegado una instrucción MQTT desde el dispositivo que detectó el problema, donde se interrumpe la tarea y se elige el estado nuevo y el dispositivo funciona con el nuevo estado entrante.

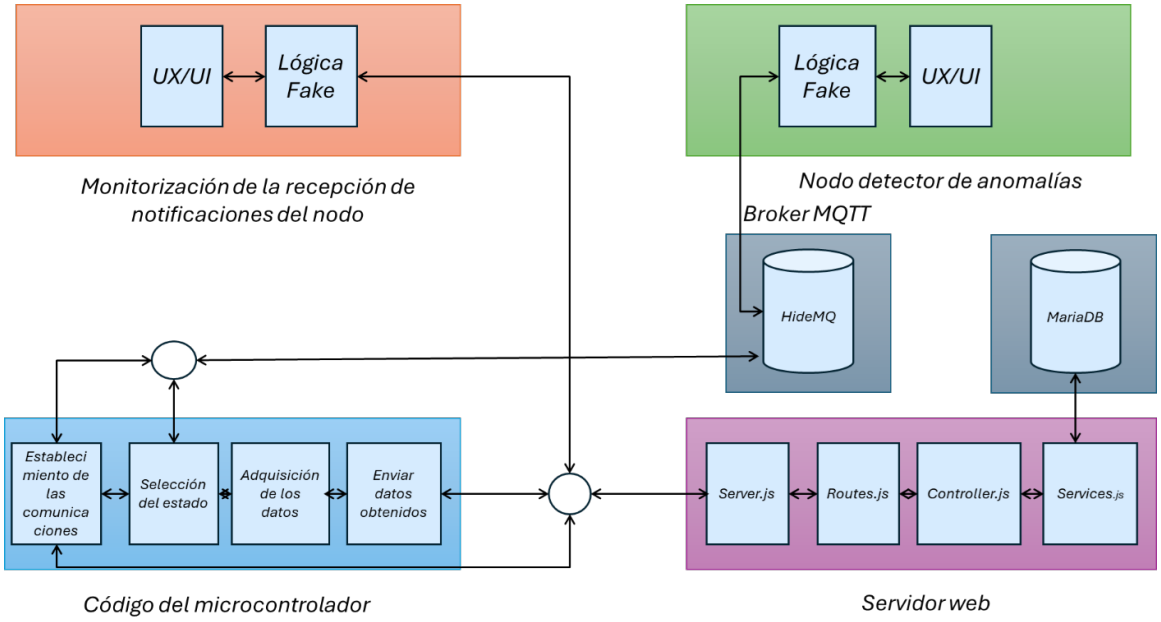


Fig. 5. Diagrama de flujo de los datos entre los módulos del sistema

C. Desarrollo de las comunicaciones de los datos hasta los softwares de visualización y control

Los módulos del sistema se componen de 6 partes: el nodo receptor de avisos para que inicie la monitorización, el código del firmware, el servidor web, el bróker MQTT, la base de datos MySQL y el nodo detector de anomalías. El nodo receptor de avisos tiene dos módulos; la UX/UI encargada de la interacción y la visibilidad de los elementos de la página web y la lógica falsa que se comunica con el servidor web mediante consultas REST de tipo GET. El servidor web se compone de 4 partes siendo estás la configuración y arranque del servidor, las rutas de la API, los controladores de las consultas y la lógica de la base de datos. La base de datos es de tipo MySQL montado sobre un motor de MariaDB conectada con la lógica del servidor. el nodo detector de anomalías se compone de la comunicación con el bróker MQTT y la UX/UI de la aplicación. Por último, el código del firmware de los dispositivos se compone de la comunicación MQTT y HTTP, la máquina de estados, la captura de los datos y el envío y recepción de los datos por las distintas comunicaciones del dispositivo tal como se puede ver el flujo de los datos en la Fig. 5. Como elemento opcional, se puede considerar la presencia de un servidor para el almacenaje de los nodos para futuras acciones o análisis.

IV. Resultados

Esta sección presenta los dispositivos usados para implementar nuestra red basada en computación autonómica y los resultados obtenidos en relación con el consumo energético de cada nodo. Finalmente, se ha monitorizado el ancho de banda que consumiría un nodo de estas características, mandando los datos a un servidor para su posterior procesamiento.

Para poder probar el sistema propuesto, se han empleado dispositivos embebidos de pequeño tamaño con integrado ESP32 [11]. El NodeMCU ESP32 es un microcontrolador con módulos WiFi y Bluetooth integrados. Se caracteriza por su sencillez a la hora de programarlo y usarlo. Es ligero y tiene

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

mayor memoria y capacidad de cálculo que un Arduino. Esto la convierte en una placa ideal para aprender a programar y desarrollar objetos o servidores conectados. El microcontrolador NodeMCU ESP32 funciona en un rango de tensión de 7-12V gracias a su regulador de tensión integrado, mientras que el microprocesador funciona a una tensión de 3,3V. En funcionamiento normal, el microcontrolador consume hasta 45mA (si no recibe alimentación externa) y puede aceptar una corriente máxima de 40mA en cada uno de sus pines IO. La Tabla I resumen las características principales de este tipo de integrados.

Tabla I. Características de un NodeMCU ESP32

Microcontrolador	
Nombre	ESP32
Marca:	Espressif
Características	
CPU:	ESP-WROOM-32
Tensión de alimentación:	7-12V
Tensión lógica:	3.3V
E/S digitales:	14
Entradas analogicas:	6
Memoria Flash:	4000kB
SRAM:	520kB
EEPROM:	448kB
Fecuencia de reloj:	240 MHz
Wifi:	Sí
Bluetooth:	Sí
Zocalo SD card:	No
UART/SPI/I2C/I2S:	Sí/Sí/Sí/Sí

Tras programar y configurar los diferentes nodos, se emula la presencia de eventos que fuercen el paso de un estado a otro con el objetivo de monitorizar los consumos de energía en cada uno de sus estados y el consumo del mismo nodo, en caso de hacer una ejecución secuencial del código.

A. Tiempos en el que el nodo permanece en cada estado

La Fig. 6 muestra los valores temporales durante el cual el nodo debe permanecer en cada estado para realizar la tarea asociada a cada modo de funcionamiento. Como se observa, los procedimientos de activación y desactivación de la transmisión son despreciables, en comparación a los procedimientos de monitorización.

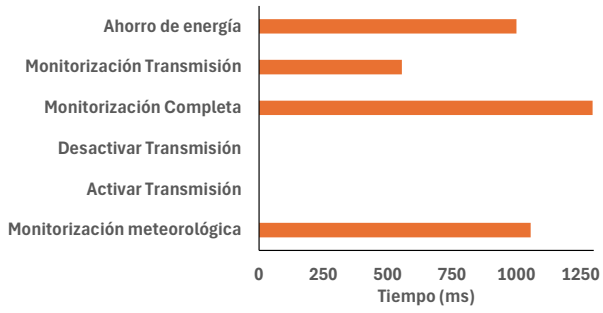


Fig. 6. Tiempo del nodo en cada estado

Siguiendo la distribución de tiempos mostrada en la figura anterior y usando un medidor de energía, se toman las medidas reales de consumo del nodo en cada estado. La Fig. 7 muestra el consumo por hora de los distintos modos de funcionamiento del sistema. En ella, se observa que, al aplicar la computación autonómica, los diferentes modos de funcionamiento tienen consumos muy diferentes por hora. Respecto a una ejecución secuencial del código, el consumo global del sistema es bastante más alto que a los diferentes modos de funcionamiento de la computación autonómica. Si se considera el tiempo y energía consumido en la tarea de envío de datos al servidor, se observa que el consumo es más alto que la captura de los datos por parte de los sensores. Esto demuestra que aplicar una computación autonómica aumenta la eficiencia del sistema respecto a una computación de tipo secuencial.

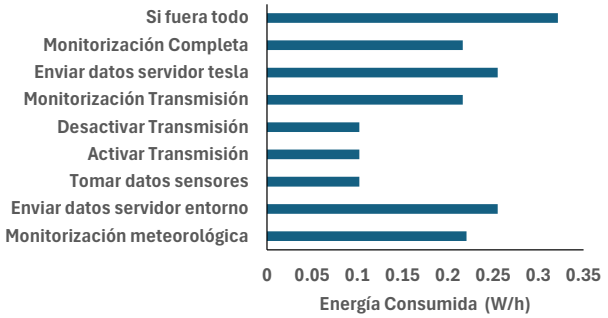


Fig. 7. Consumo de energía del nodo en función de los distintos modos de funcionamiento

Finalmente, en la Fig. 8 se ha calculado la diferencia de consumo por día respecto a si se ejecutara el código secuencialmente. En ella se ve que hay una diferencia de 2,5 W/día en favor de la computación autonómica realizando las tareas de monitorización completa, monitorización meteorológica y de la transmisión con respecto a si se ejecutara todo de manera secuencial. Si únicamente se trabajara con los modos de funcionamiento de activación/desactivación de la transmisión, la reducción en el consumo de energía roza los 5,3 W/día. Considerando las restricciones energéticas en las redes de sensores, se considera que este ahorro de energía es tremendamente importante.

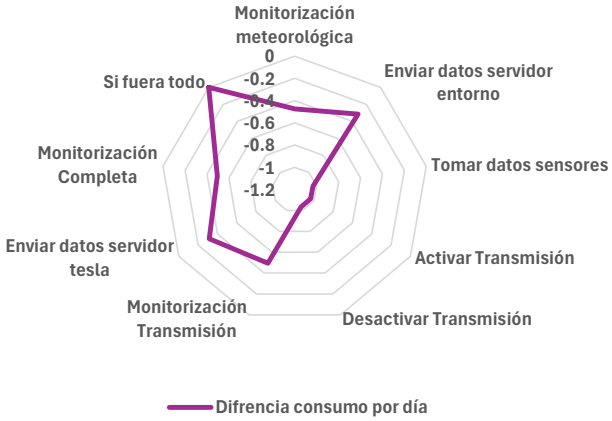


Fig. 8. Diferencia de consume del nodo en cada estado

B. Medidas de tráfico de red generado por un nodo

Finalmente, se realizan medidas reales del tráfico generado por este tipo de nodos. En la Fig. 9 se puede ver el tráfico entrante al servidor en los distintos modos de funcionamiento

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

del nodo transmisor. En el primer pico se trata del modo de monitorización del entorno estando alrededor a los 20 paquetes por segundo. El valle de la función se centra en los modos de desactivar y activar la transmisión estando alrededor de los 3 paquetes por segundo. Por último, la parte final de la captura muestra de los modos de monitorización de la transmisión y la monitorización completa, consumiendo un ancho de banda entorno a los 25 paquetes por segundo. Con ello, se puede observar, que, en volumen de tráfico de la red, dependerá en gran medida del modo en el que estas aumentan o disminuye el tráfico de red.

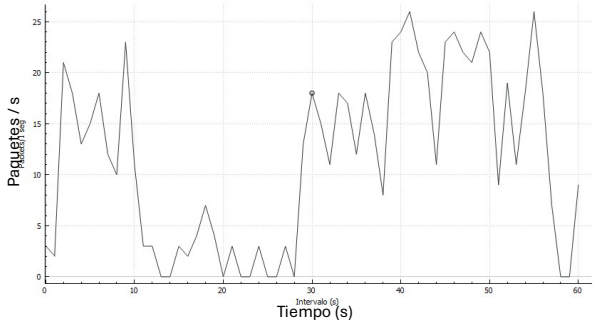


Fig. 9. Tráfico de red entrante al servidor por un nodo transmisor

La Fig. 10 muestra el tráfico entrante al servidor por parte de las peticiones de un cliente web, al solicitar información al servidor. En ella se observa que cuando se inicia una sesión de un cliente web, se genera un pico puntual de 7000 paquetes por segundo al servidor.

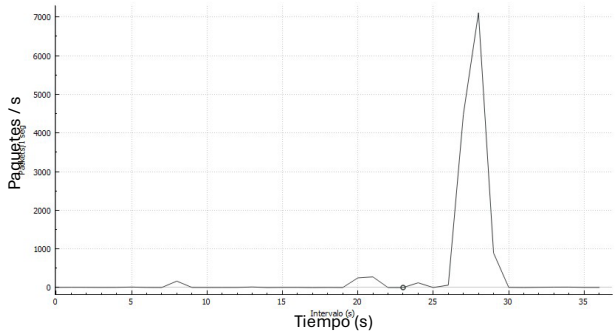


Fig. 10. Tráfico de red entrante al servidor por parte de un cliente web al solicitar información

V. Conclusiones

Con la proliferación de los sistemas IoT y las redes de sensores inalámbricas, nacieron nuevos retos a los que hacer frente. Entre ellos, siempre se ha encontrado la necesidad de reducir el consumo energético de los dispositivos y el de mejorar la eficiencia en la computación de la información por parte de los nodos sensores.

Siguiendo estas ideas, en este artículo, se ha presentado una nueva propuesta de computación autonómica en nodos sensores basadas en el modelo MAPE-K, que nos ha permitido crear sistemas de monitorización multihilo en dispositivos que a priori son incapaces de hacerlo como son los NodeMCU ESP32. Tras probarlo y compararlo con la ejecución de los mismos códigos en modo secuencial, hemos observado que por término medio seríamos capaces de ahorrar unos 2,5 W/día, valor nada despreciable, si se considera que los consumos de este tipo de dispositivos, en sus tareas de transmisión podría llegar a los 100 mA, en algunos casos.

Como trabajos futuros, se trataría de integrar este concepto y modo de trabajo en sistemas de transmisión de potencia inalámbrica de largo alcance para alimentar sensores en áreas poco accesibles [12]. Así mismo, también se pretende trabajar con un mayor número de nodos y de este modo poder comprobar a nivel práctico la escalabilidad real del sistema.

Agradecimientos

This publication is part of the project PID2023-153039OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by Generalitat Valenciana (THINKINAZUL/2021/002).

Referencias

[1] Eger, A. M., Earp, H. S., Friedman, K., Gatt, Y., Hagger, V., Hancock, B., ... & Reeves, S. (2022). The need, opportunities, and challenges for creating a standardized framework for marine restoration monitoring and reporting. Biological Conservation, 266, 109429.

[2] Rehmani, M. H., & Pathan, A. S. K. (Eds.). (2016). Emerging communication technologies based on wireless sensor networks: current research and future applications.

[3] How to use ESP32 Dual Core with Arduino IDE. [Artículo en linea]. Disponible en: https://randomnerdtutorials.com/esp32-dual-core-arduino-ide/#:~:text=The%20ESP32%20comes%20with%202%20Xtensa%2032-bit%20LX6,on%20the%20ESP32%20second%20core%20by%20creating%20tasks [Ultimo acceso: 13 de junio de 2025]

[4] Alam, M. G. R., Hassan, M. M., Uddin, M. Z., Almogren, A., & Fortino, G. (2019). Autonomic computation offloading in mobile edge for IoT applications. Future Generation Computer Systems, 90, 149-157

[5] Parashar, M., Hariri, S. (2005). Autonomic Computing: An Overview. In: Banâtre, JP., Fradet, P., Giavitto, JL., Michel, O. (eds) Unconventional Programming Paradigms. UPP 2004. Lecture Notes in Computer Science, vol 3566. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11527800_20

[6] IBM: An architectural blueprint for autonomic computing. IBM White Paper. [Artículo en línea]. Disponible en: https://www-03.ibm.com/autonomic/pdfs/ACBlueprintWhitePaperV7.pdf (2005). [Ultimo acceso: 13 de junio de 2025]

[7] Rutten, E., Marchand, N., & Simon, D. (2018, January). Feedback control as MAPE-K loop in autonomic computing. In Software Engineering for Self-Adaptive Systems III. Assurances: International Seminar, Dagstuhl Castle, Germany, December 15-19, 2013, Revised Selected and Invited Papers (pp. 349-373). Cham: Springer International Publishing.

[8] Fortino, G., Russo, W., Savaglio, C., Shen, W., & Zhou, M. (2017). Agent-oriented cooperative smart objects: From IoT system design to implementation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48(11), 1939-1956.

[9] Sendra, S., Ivars-Palomares, A., Zaragoza-Esquerdo, M., & Lloret, J. (2023, September). Group-Based Collaborative Environments for Coastal Areas Monitoring. In International Conference on Cooperative Design, Visualization and Engineering (pp. 202-213). Cham: Springer Nature Switzerland.

[10] Huebscher, M. C., & McCann, J. A. (2008). A survey of autonomic computing—degrees, models, and applications. *ACM Computing Surveys (CSUR)*, *40*(3), 1-28.

[11] Caracterisitcas del módulo NodeMCE ESP32. [Artículo en línea]. Disponible en: https://www.circuitos-electricos.com/esp32-especificaciones-y-disenos/ [Ultimo acceso: 13 de junio de 2025]

[12] Zaragoza-Esquerdo, M., Sendra, S., Ahmad, A., & Lloret, J. (2024, September). Collaborative Wireless Energy Charging System in Wireless Sensor Networks. In International Conference on Cooperative Design, Visualization and Engineering (pp. 107-116). Cham: Springer Nature Switzerland.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Impacto de la distancia de visualización en la percepción de calidad de representaciones humanas volumétricas comprimidas en entornos de realidad virtual

Mohamad Hjeij[1,2], Mario Montagud[1,3], David Rincón[2], Sergi Fernández[1,2]

[1]Fundación i2CAT, [2]Universitat Politècnica de Catalunya, [3]Universitat de València

hjeij.mohamad@i2cat.net, mario.montagud@i2cat.net , david.rincon@upc.edu , sergi.fernandez@i2cat.net .

Resumen—La compresión eficiente del video volumétrico es crucial para habilitar servicios escalables y de baja latencia de telepresencia holográfica (holoportación), manteniendo al mismo tiempo la calidad de experiencia (Quality of Experience, QoE) percibida. Aunque investigaciones previas han puesto énfasis en el rendimiento de compresión y el renderizado en tiempo real, se ha prestado poca atención a cómo los artefactos de compresión afectan la percepción del usuario según la distancia de visualización en entornos de Realidad Virtual (VR). Este estudio presenta primero una comparación objetiva entre un método novedoso de compresión de nubes de puntos basado en video (V*-PCC) y un códec geométrico de última generación (G-PCC), demostrando el mejor desempeño del V*-PCC. A partir de ello, se lleva a cabo un estudio con usuarios para evaluar la calidad percibida de hologramas humanos volumétricos dinámicos comprimidos a distintos niveles y mostrados a tres distancias dentro de un entorno VR. Los resultados destacan dos hallazgos clave: (i) la QoE percibida mejora al aumentar la distancia de visualización para un mismo nivel de compresión, y (ii) los artefactos de compresión se vuelven menos percept y más aceptables a medida que aumenta la distancia. Estas conclusiones subrayan la importancia de integrar factores perceptivos en el diseño de sistemas y apuntan hacia estrategias de codificación adaptativas y sensibles a la distancia como una vía prometedora para mejorar la adaptabilidad, escalabilidad y la experiencia del usuario en aplicaciones volumétricas inmersivas.

Palabras clave—Compresión de Video Volumétrico, Realidad Virtual, Distancia de Visualización, Evaluación Perceptual

I. INTRODUCCIÓN

El video volumétrico y las tecnologías holográficas se están consolidando como habilitadores claves de las experiencias inmersivas de próxima generación**.** Sin embargo, aún existen diversos desafíos para que estas tecnologías logren una adopción generalizada, especialmente en escenarios multiusuario desplegados en entornos distribuidos, ubicuos y heterogéneos. Entre los elementos habilitadores fundamentales, la compresión de video volumétrico desempeña un papel central al reducir la demanda de procesamiento y ancho de banda, al tiempo que mantiene la calidad perceptiva.

Si bien la investigación previa se ha centrado mayoritariamente en la eficiencia de compresión y el rendimiento en tiempo real, se ha prestado menos atención al impacto perceptivo de la compresión, en particular cuando los usuarios observan hologramas humanos dinámicos con distintos niveles de calidad y a diferentes distancias dentro de entornos de Realidad Virtual (*Virtual Reality*, VR). Estas condiciones son típicas de los escenarios inmersivos, lo que convierte esta cuestión en un área de investigación altamente relevante y, sin embargo, poco explorada. Abordar esta problemática es esencial para avanzar en el diseño de sistemas y optimizar la experiencia del usuario, con posibles implicaciones para la escalabilidad, interoperabilidad y sostenibilidad de los servicios emergentes de medios volumétricos.

La importancia de la calidad perceptiva en el video volumétrico no ha pasado desapercibida. Estudios recientes han comenzado a explorar la calidad de experiencia (*Quality of Experience*, QoE) en relación con los niveles de compresión bajo diversas condiciones, tales como:

- Formatos y representaciones volumétricas: incluyendo nubes de puntos [1–5], compresión basada en video frente a geometría [6], codificación adaptativa por bloques, y alternativas basadas en mallas [5,7].
- Impacto de las distorsiones geométricas y de textura en la compresión de nubes de puntos [3].
- Dinámicas de cambio de calidad durante la reproducción [1,4,8].
- Modalidades y condiciones de visualización: como visualización en 2D [1,7], experiencia inmersiva VR con 6 *Degrees of Freedom* (6DoF) [2–5], y

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

comparaciones entre experiencias en 3DoF y 6DoF [6].
- Efecto de la distancia entre el observador y el holograma [5,7,8].

Este artículo se basa en dichos estudios previos y examina específicamente el efecto de la distancia de visualización en la calidad percibida, abordando varias limitaciones presentes en la literatura actual:

- Los estudios anteriores suelen utilizar secuencias de alta calidad provenientes de datasets públicos [9], capturadas con configuraciones volumétricas profesionales. Aunque útiles para la reproducibilidad, estas secuencias son breves (~10s), carecen de contenido conversacional y no reflejan el tipo de capturas alcanzables con cámaras RGB-D asequibles como Microsoft Azure o Orbbec Femto Bolt, dispositivos más accesibles para usuarios generales.
- En [5,7,8], los efectos relacionados con la distancia se abordaron parcialmente, pero con limitaciones clave. Por ejemplo, [7] se centró únicamente en visualización 2D; [5] utilizó conversión a mallas e introdujo niveles fijos de calidad sin abordar la compresión como tal; y [8] se desarrolló en Realidad Aumentada (*Augmented Reality*, AR), limitada por el campo de visión (*Field of View*, FoV), lo que impidió la visualización de cuerpos completos a corta distancia.

En cambio, este trabajo aborda estas limitaciones mediante:

1. La creación de secuencias más largas y de alta calidad en nube de puntos, capturadas con tres cámaras Orbbec Femto Bolt (color 1080p, profundidad 720p), representando a un sujeto humano frontal con gestos conversacionales naturales. Estas secuencias se publicarán tras la revisión.
2. La evaluación de la calidad percibida en tres niveles de compresión y tres distancias de visualización fijas en VR, deshabilitando el 6DoF para asegurar condiciones uniformes de visualización. Se utilizaron secuencias más largas y contenido alternante para mitigar la pérdida de atención y mejorar la fiabilidad del QoE.
3. La combinación de metodologías de estímulo único (Single Stimulus Methodology, SSM) [1,4,6] y estímulo doble (Double Stimulus Methodology DSM) [3,5,7] para fortalecer la fiabilidad de la evaluación y explorar la interacción entre ambos protocolos perceptivos, un tema aún debatido en la comunidad [3,6].

Tras establecer el rendimiento objetivo superior de un nuevo método de compresión de nubes de puntos basado en Video (V*-PCC, *Point Cloud Compression*) frente a un códec de referencia basado en Geometría (G-PCC) [6,10], se realizó un estudio con usuarios en VR para evaluar cómo la distancia de visualización afecta la calidad percibida bajo distintos niveles de compresión. Los resultados revelaron dos hallazgos principales: (i) la calidad percibida mejora al aumentar la distancia a un nivel de compresión fijo, y (ii) los artefactos de compresión se vuelven menos molestos y más aceptables a mayores distancias.

Estos hallazgos respaldan la necesidad de estudios perceptivos más avanzados y granulares para determinar los umbrales de perceptibilidad de los artefactos de compresión en función de la distancia entre el observador y elementos objetivo. Además, refuerzan la necesidad de estrategias de codificación adaptativas y conscientes del contexto de visualización, lo que podría mejorar significativamente la escalabilidad, adaptabilidad y uso de recursos de los sistemas de video volumétrico en aplicaciones inmersivas.

II. Trabajos Relacionados

Como se describió en la Sección I, estudios previos han explorado la calidad perceptiva de secuencias de video volumétrico con sujetos humanos bajo diversas condiciones de prueba y configuraciones.

El trabajo en [1] investigó el impacto del cambio de calidad en la QoE al visualizar secuencias de nubes de puntos de 10 segundos en pantallas 2D mediante transmisión adaptativa HTTP. Los resultados mostraron que la calidad de la textura desempeña un papel crucial en la satisfacción del usuario, y que el contenido con bajo contraste puede mejorar la calidad percibida global. En [2], los autores presentaron una caja de herramientas para diseñar y realizar experimentos de QoE relacionados con la visualización de secuencias de nubes de puntos comprimidas en entornos de realidad virtual. Su marco metodológico permite evaluaciones inmersivas con navegación 6DoF, proporcionando una base experimental para pruebas perceptuales en contextos XR.

El estudio en [7] se centró en el efecto de la compresión y la distancia de visualización sobre la QoE, comparando representaciones holográficas basadas en nubes de puntos y en mallas visualizadas en pantallas 2D. Se identificaron dos hallazgos clave: (i) distancias de visualización más cercanas reducen la QoE para un mismo nivel de calidad, y (ii) la compresión de nubes de puntos es más adecuada para escenarios con baja tasa de bits (*bitrate*), mientras que la compresión basada en mallas ofrece mejores resultados a distancias cortas y bitrates altos.

En [6], se realizó una evaluación subjetiva de calidad bajo condiciones inmersivas de VR, comparando modos de visualización en 3DoF y 6DoF, y evaluando dos estrategias de compresión: un códec basado en video y otro basado en geometría con estructura de *octree*. El estudio destacó cómo la interactividad y la libertad espacial influyen en la percepción de calidad.

El trabajo en [3] analizó por separado los efectos de la distorsión geométrica y de textura en la calidad percibida de secuencias de nubes de puntos comprimidas en entornos VR con 6DoF. Se concluyó que ambos componentes tienen un impacto significativo en la experiencia del usuario. En [4], los investigadores evaluaron los beneficios de la transmisión adaptativa basada en mosaicos (*tile-based*) para nubes de puntos, en comparación con una línea base no adaptativa, en contextos de VR inmersiva. Los resultados indicaron que la transmisión adaptativa mejora la calidad perceptiva sin comprometer la eficiencia de recursos.

El estudio en [5] examinó los efectos de convertir nubes de puntos en representaciones en malla para lograr un renderizado más eficiente, utilizando secuencias de 10

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

segundos en entornos VR con 6DoF. Se observó que tanto el nivel de compresión como la distancia de observación afectan a la calidad percibida, mientras que los cambios de calidad influyen poco en la percepción del usuario.

Finalmente, [8] exploró el impacto de cuatro factores en la QoE al visualizar secuencias dinámicas de nubes de puntos en un entorno de AR: configuraciones de compresión, cambio de calidad, distancia de visualización y características del contenido. El estudio confirmó que las distancias cortas y la alta compresión reducen la calidad percibida, mientras que las distancias mayores mitigan las diferencias notables en calidad.

Estos estudios aportan conocimientos valiosos sobre la percepción del video volumétrico. Sin embargo, como se señaló en la Sección I, muchos se basaron en secuencias cortas, visualización en 2D o configuraciones basadas en AR, y utilizaron datasets capturados profesionalmente que no reflejan las condiciones reales de captura con dispositivos de consumo. Este artículo busca ampliar y profundizar el conocimiento actual abordando dichas limitaciones e introduciendo nuevas condiciones experimentales, incluyendo secuencias RGB-D más largas capturadas con dispositivos asequibles, visualización inmersiva en VR a distancias fijas y una metodología combinada de evaluación utilizando tanto SSM como DSM.

III. METODOLOGÍA

Este estudio comprende dos fases experimentales centradas en la codificación de hologramas humanos volumétricos dinámicos.

La Prueba 1 establece una línea base objetiva mediante la comparación del rendimiento de un códec de referencia basado en geometría (G-PCC) [6,10] frente a un nuevo método propuesto de compresión de nubes de puntos basado en video (V*-PCC). Sin embargo, el enfoque principal recae en la Prueba 2, que investiga la calidad perceptiva de secuencias holográficas comprimidas con V*-PCC a diferentes niveles de calidad y distancias de visualización del observador. La Prueba 1 consiste en una evaluación técnica diseñada para analizar el rendimiento objetivo de los códecs G-PCC y V*-PCC en condiciones controladas y a través de tres configuraciones de calidad predefinidas: Alta (*High Quality*, HQ), Media (*Medium Quality*, MQ) y Baja (*Low Quality*, LQ). El contenido volumétrico fue capturado utilizando una configuración de tres cámaras RGB-D Orbbec Femto Bolt, dispuestas en un arco frontal para lograr una cobertura corporal completa y una reconstrucción de alta fidelidad. El sistema de captura operaba a una tasa de fotogramas de 15 FPS, correspondiente a una ventana de procesamiento de 66 ms por fotograma. Cada secuencia tenía una duración de 90 segundos y fue grabada bajo condiciones idénticas mismo actor, movimientos y configuración de iluminación para garantizar la comparabilidad.

Las configuraciones de compresión variaron entre ambas canalizaciones:
(i) El códec G-PCC aplicó cuantificación geométrica basada en vóxeles, con tamaños de 0.05 (LQ), 0.025 (MQ) y 0.0 (HQ).
(ii) El códec V*-PCC* empleó valores de *Constant Rate Factor* (CRF) y *Quantization Parameter* (QP) de 34 (LQ), 17 (MQ) y 0 (HQ).

Estas configuraciones se seleccionaron para alinear los niveles de compresión de ambos códecs y permitir una comparación justa y reproducible (Fig. 1). Esta prueba inicial sirve como referencia para evaluar el potencial del V*-PCC e informar su aplicación en el estudio de usuarios posterior. Se evaluaron las siguientes métricas bajo cada condición de prueba: (i) el recuento de vóxeles, que refleja la densidad de puntos de las secuencias codificadas; (ii) los fotogramas por segundo codificados (FPS), influenciados por la tasa de captura y que indican el rendimiento del proceso de codificación; y (iii) el tiempo de procesamiento, que representa el tiempo necesario para codificar cada fotograma individual.

Tal y como se puede apreciar en la Fig. 1, la comparación visual entre G-PCC y V*-PCC pone de manifiesto diferencias cualitativas entre ambas canalizaciones. Mientras que G-PCC tiende a reducir la densidad de puntos y a introducir pérdidas de detalle geométrico en las representaciones, V*-PCC mantiene una mayor estabilidad visual y consistencia en la textura, incluso en configuraciones con mayor compresión. Esta figura complementa el análisis objetivo, permitiendo al lector identificar claramente los beneficios prácticos del enfoque basado en vídeo frente al geométrico.

La Prueba 2 se centró en la evaluación subjetiva de la calidad percibida al visualizar hologramas humanos volumétricos dinámicos codificados mediante el códec V*-PCC, considerando distintos niveles de compresión y distancias de visualización. Cada participante evaluó nueve condiciones experimentales distintas, que incluían secuencias holográficas de un actor realizando gestos corporales y faciales naturales en un contexto conversacional. Los hologramas se presentaron a tres distancias predefinidas del participante (actuando como observador en el entorno VR): (i) D1 – Cercana (1 metro), (ii) D2 – Media (2,5 metros) y (iii) D3 – Lejana (5 metros).

Para cada distancia, primero se mostró una secuencia de entrenamiento en alta calidad (HQ), seguida de tres secuencias correspondientes a diferentes niveles de compresión: HQ, calidad media (MQ) y baja (LQ). El orden de presentación de los niveles de calidad fue contrabalanceado. La compresión se ajustó utilizando valores de CRF de 0 (HQ), 17 (MQ) y 34 (LQ), logrando aproximadamente una compresión 2× para MQ y 4× para LQ en comparación con HQ. Estos ajustes se aplicaron únicamente al componente de video, sin alterar la textura geométrica.

Fig. 1. *Comparación visual del resultado en alta calidad: G-PCC (izquierda, 12 Mbps) vs. V-PCC (derecha, 9 Mbps)**

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Los resultados visuales correspondientes a la Prueba 2 se ilustran en la Fig. 2, que pone de manifiesto que los diferentes niveles de compresión generan artefactos visuales claramente distinguibles. En la condición LQ (1 Mbps) se aprecia una pérdida de detalle en la textura y suavizado en las expresiones faciales, mientras que en MQ (4 Mbps) los defectos resultan menos intrusivos y la representación conserva una aceptable coherencia visual. La condición HQ (9 Mbps) ofrece la mayor fidelidad, manteniendo bordes definidos y una apariencia más realista. Esta comparación visual refuerza la motivación del estudio perceptivo, al mostrar cómo la degradación introducida por la compresión varía en función de la calidad y anticipa que su impacto puede depender también de la distancia de visualización. El orden de las distancias fue fijo, comenzando por la más cercana (D1) y avanzando hacia la más lejana (D2, D3). Todas las secuencias de video tenían una duración de 30 segundos.

Los participantes comenzaron la sesión firmando un formulario de consentimiento y completando un cuestionario demográfico y de antecedentes. Las condiciones de visualización se implementaron en un entorno personalizado desarrollado en Unity3D, siguiendo el estándar ITU-R BT.500-13 [11], como se ha utilizado en estudios previos [4,8]. Para la evaluación subjetiva, se empleó una metodología de calificación durante la sesión, validada en literatura reciente [4]. El experimento se llevó a cabo en una sala tranquila y cómoda, libre de distracciones visuales y auditivas (véase Fig. 3). Los participantes utilizaron un visor Meta Quest 3 conectado a un ordenador portátil con GPU dedicada. Permanecieron de pie y estáticos, experimentando el contenido en un modo de visualización 3DoF+ (sin movimiento físico ni interacción). Todas las secuencias se mostraron de forma pasiva en una línea temporal lineal, y la sesión completa tuvo una duración aproximada de 20–25 minutos por participante.

Tras la secuencia de entrenamiento en HQ, los participantes visualizaron los tres niveles de calidad en orden aleatorio y respondieron a dos tipos de evaluación:

- Una metodología de estímulo único (SSM), mediante una escala Likert absoluta de 10 puntos (1–10).
- Una metodología de estímulo doble (DSM), utilizando una escala Likert relativa de 5 puntos, que iba desde “Mucho peor” hasta “Mucho mejor” respecto a la referencia.

Las respuestas se registraron mediante el controlador del visor VR. Esta evaluación dual se diseñó para capturar tanto la satisfacción subjetiva con la calidad absoluta como la percepción de degradación en comparación con una referencia conocida, aspecto especialmente relevante dada la fidelidad aún limitada del video volumétrico actual.

Fig. 2. Impacto de la compresión en la calidad visual y el ancho de banda: HQ (izquierda, 9 Mbps), MQ (centro, 4 Mbps) y LQ (derecha, 1 Mbps).

Al finalizar la visualización de todas las secuencias en los tres niveles de distancia (D1, D2, D3), los participantes completaron un breve formulario de retroalimentación general, esta vez en papel, con el fin de evitar prolongar la exposición al entorno de realidad virtual.

IV. Resultados

A. Resultados de la Prueba 1

Los resultados de la Prueba 1 ponen de manifiesto una diferencia significativa en el rendimiento entre los códecs V*-PCC y G-PCC. En términos de densidad de vóxeles, la canalización V*-PCC al estar basada en imágenes mantuvo un recuento de puntos constante en todos los niveles de calidad (HQ, MQ, LQ), con un promedio de aproximadamente 1,1 millones de puntos por secuencia. Esta estabilidad se atribuye a su diseño: la canalización procesa tanto el color como la geometría en forma de fotogramas de imagen, que son gestionados eficientemente por la GPU. Como cada píxel se trata de forma independiente, la densidad de puntos se mantiene prácticamente inalterada por los ajustes de compresión, siempre que los datos quepan dentro del fotograma de video.

En contraste, el rendimiento de G-PCC depende en gran medida de la densidad de vóxeles y de su distribución espacial. Su compresión se basa en estructuras jerárquicas como octrees, que deben particionar el espacio de forma recursiva y calcular la distribución de vóxeles. Esta complejidad aumenta considerablemente con nubes de puntos más densas, lo que a menudo imposibilita el procesamiento en tiempo real. A medida que se incrementa la compresión, G-PCC muestra una fuerte caída en el recuento de vóxeles: desde un promedio de 203.720 vóxeles en HQ hasta menos del 20% de esa cantidad en los niveles MQ y LQ. Esta pérdida de resolución espacial pone en evidencia una limitación fundamental de G-PCC en escenarios que requieren transmisión adaptativa y fidelidad geométrica constante.

Otras métricas de rendimiento refuerzan la ventaja del enfoque V*-PCC:

- Los fotogramas por segundo (FPS) se mantuvieron estables en todos los niveles de calidad para V*-PCC, oscilando entre 14,7 y 15 FPS, coincidiendo con la tasa original de captura.
- Por el contrario, G-PCC mostró una variabilidad notable, con una caída a 10,18 FPS en HQ, aunque recuperándose hacia los 15 FPS en niveles de calidad más bajos (véase Fig. 3).

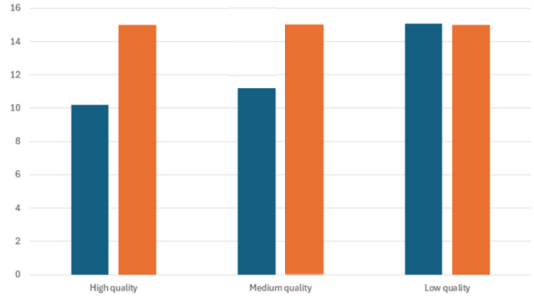


Fig. 3. *Comparación de FPS en la Prueba 1 (Azul: FPS de G-PCC; Naranja: FPS de V*-PCC).*

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

Sin embargo, la diferencia más llamativa se encuentra en el tiempo de procesamiento por fotograma:

- V*-PCC logró tiempos de codificación consistentemente bajos, entre 7,89 ms y 8,61 ms, independientemente del nivel de calidad.
- G-PCC, en cambio, requirió 93,5 ms por fotograma en HQ, y aun en LQ, siguió siendo más del doble de lento que V*-PCC (17,8 ms).

Estos resultados demuestran que V*-PCC ofrece una eficiencia de procesamiento superior, una fidelidad visual estable y una mayor escalabilidad, lo que lo convierte en una opción adecuada para aplicaciones de video volumétrico en tiempo real y adaptativas. En términos prácticos, G-PCC solo puede mantener un rendimiento en tiempo real (≥15 FPS) con nubes de puntos de menos de 30.000 puntos, debido a su elevado coste computacional. En cambio, V*-PCC puede mantener de forma fiable hasta 30 FPS con nubes de puntos densas (como se confirmó en esta prueba con aproximadamente 1 millón de puntos), lo que justifica su elección para la evaluación subjetiva en la Prueba 2.

B. Resultados de la Prueba 2:

Un total de 24 participantes (13 hombres y 11 mujeres), con edades comprendidas entre 21 y 62 años, participaron en el experimento de evaluación subjetiva.

Resultados DSM:

Los participantes evaluaron la calidad percibida de cada secuencia en comparación con una referencia de alta calidad, con resultados que variaron según el nivel de compresión y la distancia de visualización, tal como se anticipaba.

En D1 – Distancia cercana (1 metro) (véase Fig. 4), los participantes demostraron una alta sensibilidad a los artefactos de compresión.

- HQ fue valorado como "Equivalente" por el 72 % de los participantes y como "Ligeramente mejor" por un 8 %, posiblemente debido a gestos expresivos o a la exposición previa a contenido de menor calidad.
- MQ recibió una evaluación mixta: el 40 % lo consideró "Equivalente" y el 52 % lo calificó como "Peor", lo que indica una disminución moderada en la calidad percibida.
- LQ obtuvo los peores resultados, con un 84 % de las valoraciones ubicadas en las categorías de "Ligeramente peor" "Mucho peor".

En D2 – Distancia media (2,5 metros) (véase Fig. 5), las diferencias perceptivas comenzaron a disminuir:

- HQ mantuvo puntuaciones altas, con un 52 % de respuestas "Equivalente" y un 16 % "Mejor".
- MQ le siguió de cerca, con un 48 % de valoraciones como "Equivalente" y un 16 % como "Mejor".
- LQ mostró una ligera mejora, aunque el 76 % de los participantes aún lo calificó como "Peor."

En D3 – Distancia lejana (5 metros) (véase Fig. 6), el impacto de la compresión se volvió menos perceptible:

- HQ, MQ y LQ fueron valorados como "Equivalente" por el 60 %, 52 % y 48 % de los participantes, respectivamente.
- Una proporción moderada de respuestas (hasta un 16 % para HQ y un 12 % para MQ) calificó las secuencias como "Ligeramente mejor" o superior, lo que sugiere una mayor tolerancia perceptiva a medida que aumenta la distancia.

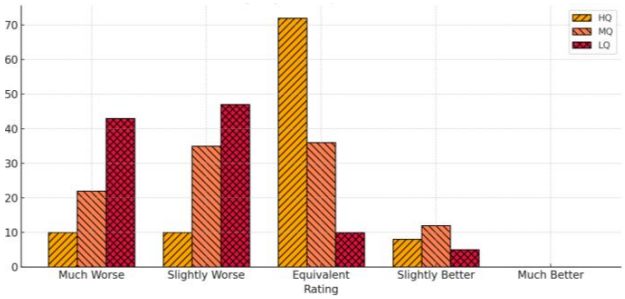


Fig. 4. Distribución de valoraciones DSM (escala Likert) para D1 en la Prueba 2.

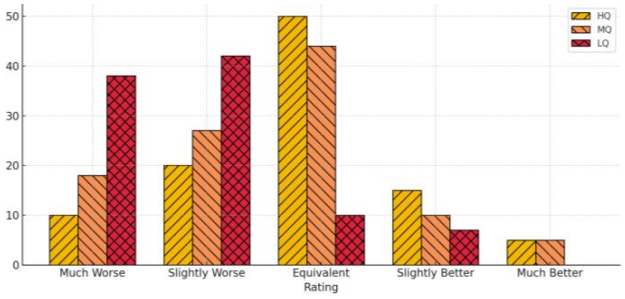


Fig. 5. Distribución de valoraciones DSM (escala Likert) para D2 en la Prueba 2.

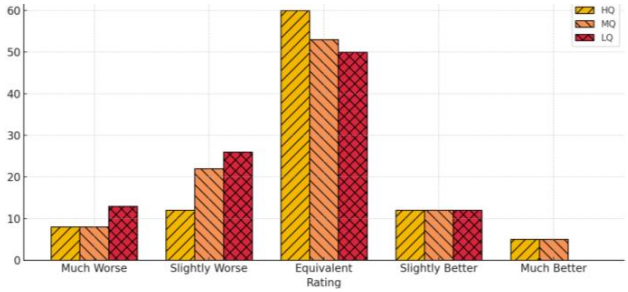


Fig. 6. Distribución de valoraciones DSM (escala Likert) para D3 en la Prueba 2.

En general, los resultados del DSM indican una tendencia clara: (i) las secuencias HQ obtuvieron sistemáticamente las valoraciones más altas (ii) Las secuencias MQ siguieron un patrón similar, aunque ligeramente inferior (iii) las secuencias LQ mostraron una mejora notable en la calidad percibida a medida que aumentaba la distancia de visualización. Estos hallazgos coinciden con estudios previos [5, 7, 8] y confirman que la distancia de visualización desempeña un papel significativo en la tolerancia perceptiva a los artefactos de compresión en video volumétrico.

Para evaluar los efectos de la distancia de visualización y el nivel de compresión sobre la experiencia percibida, se realizó un análisis de varianza de dos factores (ANOVA bidireccional) sobre las valoraciones recogidas. El análisis reveló efectos principales estadísticamente significativos tanto para la distancia ($F(2, 207) = 4.17$, $p = 0.017$) como para el nivel de calidad ($F(2, 207) = 41.93$, $p < 0.001$), lo que confirma que ambos factores influyeron en las valoraciones de calidad realizadas por los participantes. Además, se observó una interacción significativa entre distancia y calidad ($F(4, 207) = 4.73$, $p < 0.01$), lo que sugiere que el impacto de la compresión está modulado por la proximidad del observador. Estos resultados se visualizan en la Fig. 7, que presenta las valoraciones medias para cada condición con intervalos de

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

confianza del 95 %. Las secuencias de alta calidad (HQ) recibieron sistemáticamente valoraciones más altas que las de calidad media (MQ) y baja (LQ) en todas las distancias. Sin embargo, MQ mostró una caída pronunciada en la distancia más cercana (D1, media = 2.38), con mejoras en D2 (2.58) y D3 (2.83). LQ siguió una trayectoria ascendente similar, aunque se mantuvo consistentemente más baja. En general, la brecha perceptiva entre niveles de calidad fue más evidente a corta distancia y se redujo a medida que aumentaba la distancia, especialmente entre MQ y LQ.

Resultados SSM:

Los resultados obtenidos mediante SSM aportan varios hallazgos relevantes (véase Fig. 8). En general, las valoraciones absolutas en escala de 1 a 10 para los tres niveles de calidad reflejaron las mismas tendencias observadas en el análisis DSM. Las secuencias de alta calidad (HQ) obtuvieron puntuaciones elevadas de forma constante en todas las distancias, aunque sin alcanzar el umbral de "Excelente". Las secuencias de calidad media (MQ) mostraron mejoras en sus valoraciones a medida que aumentaba la distancia (D2 y D3), mientras que las secuencias de baja calidad (LQ) evidenciaron una diferenciación más marcada entre distancias, con puntuaciones significativamente más bajas en D1 y una mejora progresiva en D2 y D3. La brecha perceptiva entre niveles de calidad fue más evidente en la distancia más cercana (D1), donde las diferencias entre HQ, MQ y LQ fueron claramente distinguibles. En D2, LQ continuó siendo significativamente inferior a HQ y MQ, aunque estas dos últimas comenzaron a converger. Finalmente, en D3, las valoraciones para los tres niveles de calidad especialmente entre MQ y LQ tendieron a alinearse, lo que sugiere una menor sensibilidad perceptiva a mayores distancias de visualización.

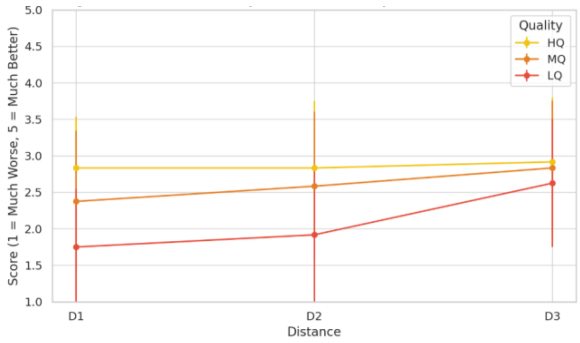


Fig. 7. Valoraciones medias de los usuarios según distancia y nivel de calidad en la Prueba 2.

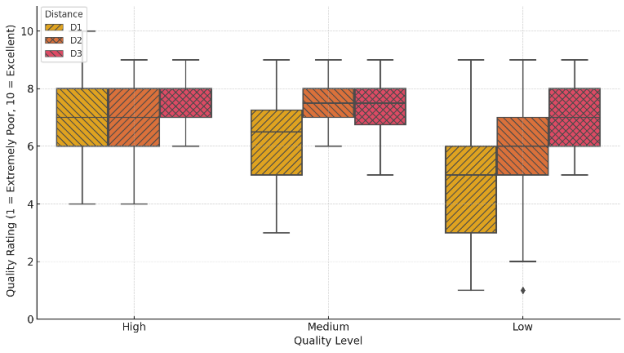


Fig. 8. Diagrama de cajas de las valoraciones de calidad (SSM) según distancia y nivel de calidad en la Prueba 2.

Para validar estadísticamente estas tendencias, se realizó un análisis ANOVA bidireccional sobre las valoraciones del SSM. El análisis reveló efectos principales significativos tanto para la calidad de compresión ($F(2, 207) = 193.47$, $p < 0.001$) como para la distancia de visualización ($F(2, 207) = 35.74$, $p < 0.001$), además de una interacción significativa entre ambos factores ($F(4, 207) = 21.44$, $p < 0.001$). Estos resultados confirman que la calidad percibida está influenciada conjuntamente por la distancia y el nivel de compresión.

Para profundizar en estos efectos, se aplicó la prueba post-hoc de Tukey (HSD). En D1, todas las comparaciones por pares entre HQ, MQ y LQ fueron estadísticamente significativas ($p < 0.001$). En D2, se mantuvieron diferencias significativas entre HQ–LQ y MQ–LQ ($p < 0.05$), pero no se detectaron diferencias significativas entre HQ y MQ. Finalmente, en D3, no se encontraron diferencias significativas entre los tres niveles de calidad ($p > 0.1$), lo que refuerza la idea de que el impacto de la compresión disminuye a medida que aumenta la distancia de visualización.

Calidad Percibida Global e Impacto:

Tras la experiencia en VR, los participantes respondieron a tres afirmaciones posteriores al test, diseñadas para evaluar la calidad perceptiva y la aplicabilidad general de las secuencias de alta calidad (de referencia).

Q1 planteaba: *"En general, para el nivel máximo de calidad (secuencias de referencia), la diferencia de calidad entre las tres distancias ha sido perceptible."* Como se muestra en la Fig. 9, una gran mayoría (79 %) estuvo total o parcialmente de acuerdo, lo que indica que incluso con la calidad más alta, la distancia de visualización influyó en la experiencia perceptiva de los usuarios.

Q2 indicaba: *"Con respecto al nivel máximo de calidad (secuencias de referencia), la calidad de video lograda es satisfactoria para ofrecer una experiencia inmersiva."* Las respuestas fueron mayoritariamente favorables, con un 66 % de los participantes expresando acuerdo. Estos resultados sugieren que el nivel de codificación en alta calidad fue, en general, percibido como suficiente para sustentar experiencias inmersivas en VR.

Q3 abordaba el posible impacto en aplicaciones reales: *"Con respecto al nivel máximo de calidad (secuencias de referencia), la calidad de video lograda ya puede tener un impacto en el consumo de medios inmersivos o en el sector de la comunicación"*. Aquí, el 52 % de los participantes estuvo de acuerdo, mientras que el 36 % respondió de forma neutral. Esta evaluación más reservada puede reflejar un optimismo prudente: los participantes reconocieron el potencial de la calidad alcanzada, pero posiblemente se vieron influenciados por factores adicionales como el rendimiento del hardware, el formato del contenido o las limitaciones de interacción que podrían afectar su implementación en entornos reales.

Finalmente, para complementar los resultados principales, se preguntó a los participantes si las diferencias de calidad entre las tres secuencias eran claramente distinguibles en cada distancia. Como muestra la Fig. 10, las respuestas siguieron un patrón dependiente de la distancia: en D1 (distancia cercana), la mayoría percibió claramente las diferencias; en D2, las respuestas fueron más variadas; y en D3 (distancia lejana), la mayoría no notó diferencias, lo que refuerza la disminución de la sensibilidad perceptiva con el aumento de la distancia.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

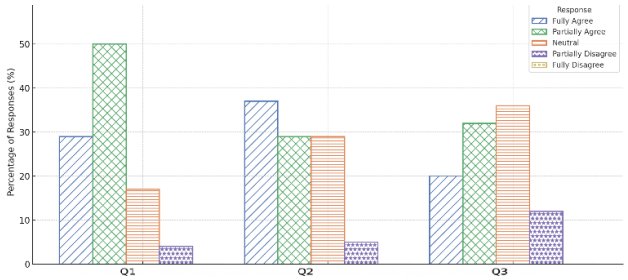


Fig. 9. Opiniones de los participantes en la Prueba 2.

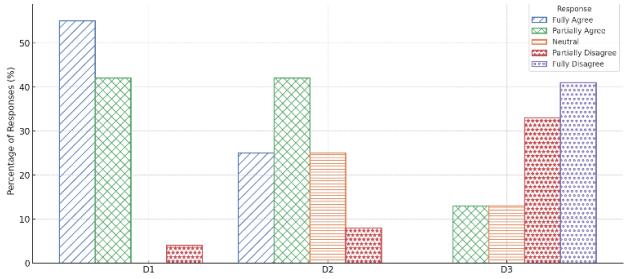


Fig. 10. Diferencia de calidad percibida entre secuencias mostradas a cada distancia en la Prueba 2.

Además de las valoraciones perceptuales, resulta relevante considerar el compromiso entre calidad visual y la tasa binaria (bitrate). En los experimentos presentados, las condiciones HQ, MQ y LQ correspondieron aproximadamente a 9 Mbps, 4 Mbps y 1 Mbps, respectivamente. Estos valores ponen de manifiesto que es posible reducir el ancho de banda hasta 4× sin comprometer severamente la QoE a distancias medias o lejanas, lo que ofrece una guía práctica para la selección de niveles de compresión en función de las restricciones de red. De este modo, los resultados no solo aportan evidencias perceptuales, sino también parámetros de referencia útiles para la optimización sistemas multimedia distribuidos.

V. Limitaciones

Aunque el estudio ofrece aportes valiosos sobre el impacto perceptivo de la compresión y la distancia de visualización en experiencias inmersivas de video volumétrico, es importante reconocer varias limitaciones. En primer lugar, el experimento se realizó en condiciones controladas de VR, utilizando un modo de visualización fijo en 3DoF y una única configuración de dispositivo (Meta Quest 3 con GPU de alto rendimiento). Esto limita la generalización de los resultados a otros contextos inmersivos, como la navegación en 6DoF, la Realidad Aumentada (AR) o entornos con hardware de menor capacidad. Además, solo se evaluaron hologramas humanos orientados frontalmente realizando gestos conversacionales, lo cual no representa la diversidad de tipos de contenido y complejidad de movimiento presentes en casos reales (por ejemplo, interacciones grupales, entornos dinámicos o escenas centradas en objetos).

En segundo lugar, la evaluación de compresión se centró exclusivamente en un códec novedoso basado en video (V*-PCC) y en una alternativa de referencia basada en geometría (G-PCC). Si bien estas opciones son representativas de las estrategias actuales de codificación volumétrica, no se incluyeron técnicas emergentes como la compresión neuronal o el *Gaussian Splatting*, que podrían arrojar resultados distintos tanto en eficiencia como en impacto perceptivo. Asimismo, solo se evaluaron tres niveles discretos de compresión y distancia; pruebas más granulares podrían ayudar a definir con mayor precisión los umbrales de aceptabilidad perceptiva. Por último, aunque la combinación de metodologías SSM y DSM mejora la solidez del enfoque experimental, la retroalimentación de los participantes se basó en una exposición de corta duración y puede no reflejar completamente fenómenos como la adaptación a largo plazo, la fatiga o el nivel de implicación en experiencias inmersivas prolongadas.

VI. Conclusiones

Este estudio evaluó el rendimiento y el impacto perceptivo de un nuevo método de compresión de nubes de puntos basado en video (V*-PCC) para la codificación y renderizado de hologramas humanos volumétricos dinámicos. La investigación siguió un enfoque en dos fases: (i) una comparación objetiva con el códec G-PCC basado en geometría, y (ii) un estudio perceptivo centrado en el usuario que analizó la influencia de los niveles de compresión y las distancias de visualización en un entorno inmersivo de realidad virtual. Los resultados objetivos confirmaron la eficiencia técnica y estabilidad de V*-PCC, validando su idoneidad para aplicaciones en tiempo real y su uso en evaluaciones subjetivas de la QoE. La evaluación perceptiva demostró además que los artefactos de compresión son más perceptibles a distancias cortas, menos intrusivos a distancias medias y en gran medida imperceptibles a mayores distancias, especialmente en el caso de secuencias de baja calidad (LQ), que fueron valoradas de forma similar al contenido de mayor calidad cuando se visualizaron desde lejos.

Estos hallazgos preliminares ofrecen una guía valiosa para el diseño de sistemas adaptativos de transmisión volumétrica y abren múltiples líneas para investigaciones futuras. El trabajo futuro se centrará en estudios de QoE más detallados que permitan identificar umbrales perceptivos tanto para los niveles de compresión como para las distancias de visualización. En paralelo, se integrarán otros códecs volumétricos, incluido G-PCC, en evaluaciones perceptivas ampliadas para comparar el rendimiento bajo diferentes condiciones. Finalmente, los conocimientos obtenidos servirán de base para el desarrollo de estrategias de codificación y flujos de transmisión sensibles a la distancia, con el objetivo de optimizar tanto el uso de recursos como la experiencia del usuario en escenarios de comunicación inmersiva en tiempo real.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por: el Programa Horizonte Europa de la Unión Europea, bajo el acuerdo nº 101135637 (proyecto PRESENCE) y nº 101070250 (proyecto XRECO); y por la Agencia Estatal de Investigación (AEI), en el marco del Proyecto Estratégico Orientado a la Transición Ecológica y a la Transición Digital (TED) 2021, bajo los acuerdos TED2021-131690B-C32 y TED2021-131690A-C33 (proyecto REVOLUTION) y en el marco de Proyectos Generación de Conocimiento 2022, bajo el acuerdo PID2022-140749OB-I00 (proyecto EVOLVE), financiado por MICIU/AEI/10.13039/501100011033 y FEDER, UE. El trabajo de Mohamad Hjeij ha sido financiado por la Generalitat de Catalunya y el Fondo Social Europeo (becas para la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

contratación de personal investigador en formación FI 2022). El trabajo de Mario Montagud ha sido financiado por MCIN/AEI/10.13039/501100011033, bajo la subvención RYC2020-030679-I.

REFERENCIAS

[1] J. van der Hooft, M. T. Vega, C. Timmerer, A. C. Begen, F. De Turck, and R. Schatz, "Objective and subjective QoE evaluation for adaptive point cloud streaming," in *Proc. 12th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.

[2] E. Alexiou, N. Yang, and T. Ebrahimi, "PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality," in *Proc. 12th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2020, pp. 1–6.

[3] X. Wu, Y. Zhang, C. Fan, J. Hou, and S. Kwong, "Subjective quality database and objective study of compressed point clouds with 6DoF head-mounted display," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 12, pp. 4630–4644, Dec. 2021.

[4] S. Subramanyam, I. Viola, J. Jansen, E. Alexiou, A. Hanjalic, and P. Cesar, "Subjective QoE evaluation of user-centered adaptive streaming of dynamic point clouds," in *Proc. 14th Int. Conf. Quality Multimedia Exper. (QoMEX)*, Sep. 2022, pp. 1–6.

[5] S. Van Damme, I. Mahdi, H. K. Ravuri, J. van der Hooft, F. De Turck, and M. T. Vega, "Immersive and interactive subjective quality assessment of dynamic volumetric meshes," in *Proc. 15th Int. Conf. Quality Multimedia Exper. (QoMEX)*, Jun. 2023, pp. 141–146.

[6] S. Subramanyam, J. Li, I. Viola, and P. Cesar, "Comparing the quality of highly realistic digital humans in 3DoF and 6DoF: A volumetric video case study," in *Proc. IEEE Conf. Virtual Reality 3D User Interface (VR)*, Mar. 2020, pp. 127–136.

[7] K. Cao, Y. Xu, and P. Cosman, "Visual quality of compressed mesh and point cloud sequences," *IEEE Access*, vol. 8, pp. 171203–171217, 2020.

[8] M. Nguyen, S. Vats, S. Van Damme, J. van der Hooft, M. T. Vega, T. Wauters, and H. Hellwagner, "Characterization of the quality of experience and immersion of point cloud videos in augmented reality through a subjective study," *IEEE Access*, vol. 11, pp. 128898–128910, 2023.

[9] E. d'Eon, J. Fillion, D. Tardif, and F. Gosselin, "8i voxelized full bodies A voxelized point cloud dataset," ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document WG11M40059/WG1M74006, 2017.

[10] R. Mekuria, K. Blom, and P. Cesar, "Design, implementation, and evaluation of a point cloud codec for tele-immersive video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 4, pp. 828–842, Apr. 2017.

[11] *Methodologies for the Subjective Assessment of the Quality of Television Images*, ITU-R Recommendation BT.500-15, International Telecommunication Union, Geneva, Switzerland, May 2023.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Aproximación hacia arquitecturas modulares asistidas por computación en el borde para optimización de comunicaciones holográficas

Genís Castillo, Sergi Fernández, David Rincón, Mario Montagud.
genis.castillo@i2cat.net, sergi.fernández@i2cat.net, david.rincon@upc.edu, mario.montagud@i2cat.net

Resumen

Este artículo presenta una aproximación inicial hacia arquitecturas modulares asistidas por computación en el borde (Edge Computing) orientadas a una optimización de la prestación de servicios de comunicaciones holográficas multiusuario. Se propone una modularización de componentes y módulos en el lado servidor que, por una parte, permite desacoplar el plano de control del plano de datos de usuario y, por otra parte, facilita una gestión dinámica de conexiones en función de requerimientos específicos y a los recursos disponibles. La arquitectura propuesta, junto con sus mecanismos de orquestación asociados, se adapta y reconfigura mediante la especificación e integración de APIs de red compatibles con iniciativas de estandarización actuales. Los beneficios de esta aproximación se validan preliminarmente a través de pruebas experimentales que evidencian, por un lado, las ventajas de utilizar servidores de borde para la gestión de comunicaciones, y por otro, el impacto positivo del balanceo de carga entre componentes servidor en cuanto a la escalabilidad y estabilidad del servicio.

***Palabras Clave*—Arquitecturas de Comunicaciones, Comunicaciones Holográficas, Calidad de Servicio, Computación en el Borde, Escalabilidad, Streaming Adaptativo**

I. INTRODUCCIÓN

En los últimos años están emergiendo plataformas innovadoras de comunicaciones holográficas que permiten superar limitaciones esenciales de los servicios de videoconferencia 2D tradicionales en cuanto a niveles de calidad de interacción, inmersión y co-presencia [1] [2].

Sin embargo, a pesar de su indudable potencial, las escasas soluciones existentes presentan asimismo limitaciones en cuanto a su capacidad de adaptabilidad y escalabilidad, debido a sus estrictos requisitos en cuanto a latencia, ancho de banda y capacidad de procesado, especialmente cuando están basadas en formatos de vídeo volumétrico capturado en tiempo real [3] [4].

Este artículo presenta una aproximación inicial hacia arquitecturas modulares asistidas por computación en el borde (Edge Computing) orientadas a una optimización de la prestación de servicios de comunicaciones holográficas multiusuario, en cuanto a su escalabilidad, consumo de recursos y niveles de Quality of Service (QoS) ofrecidos. En particular, se propone una modularización de componentes en el lado servidor que, por una parte, permite desacoplar el plano de control del plano de datos de usuario y, por otra parte, facilita una gestión dinámica de conexiones en función de requerimientos específicos y a los recursos disponibles. La arquitectura propuesta, junto con sus mecanismos de orquestación asociados, se adapta y reconfigura mediante la especificación e integración de APIs de red compatibles con iniciativas de estandarización actuales.

Los beneficios de esta aproximación se validan preliminarmente a través de pruebas experimentales en despliegues a través de entornos de red distribuidos reales y evidencian dos ventajas fundamentales: (i) la capacidad de seleccionar servidores en el borde de la red, cercanos a los clientes, permite reducir significativement el retardo de las comunicaciones (y el uso de recursos global); y (ii) la capacidad de balancear la carga de tráfico entre componentes servidor distribuidos contribuye a una mayor escalabilidad y estabilidad del servicio.

II. ESTADO DEL ARTE

A. *Estudios de revisión*

Los artículos en [3] y [4] proporcionan una revisión profunda sobre avances y retos pendientes en el campo del vídeo volumétrico. Por un lado, estos estudios reflejan los altos requerimientos de ancho de banda, procesamiento, y retardo que suelen presentar los servicios de eXtended Reality (XR) interactivos e inmersivos asociados a estos formatos emergentes. Por otro lado, subrayan la relevancia de adoptar funciones de procesamiento asistido por red (p.ej., transcodificación, conversión entre formatos, renderizado remoto) para optimizar el rendimiento, interoperabilidad y minimizar los costes asociados.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

B. Plataformas de comunicaciones holográficas

Muchas plataformas de Social Virtual Reality (VR) se han propuesto con tal de superar limitaciones intrínsicas de los sistemas de videoconferencia 2D. No obstante, la mayoría de estas plataformas se basa en avatares, lo que conlleva limitaciones importantes en cuanto a la preservación de la identidad, y calidad de interacción / comunicación. En este contexto, estudios recientes (e.g. [1]) han demostrado que: (i) las plataformas de Social VR basadas en avatares ofrecen una mejor calidad de interacción y niveles de (co-)presencia mayores que las plataformas de videoconferencia 2D; y (ii) estos niveles mejoran aún más cuando se utilizan representaciones realistas de los usuarios mediante video volumétrico (esto es, hologramas), ofreciendo una experiencia de interacción cercana a las reuniones presenciales.
Estos hallazgos son prometedores, especialmente considerando que la resolución y calidad de los hologramas en las plataformas Social VR utilizadas en dichos estudios (e.g., [1], [2], [5]) todavía presentan un amplio margen de mejora.

Se prosigue ahora con las principales plataformas Social VR/XR que admiten representaciones holográficas 3D. En [5] se presenta una plataforma basada en en el uso sensores RGB-D, tales como Kinect, y entornos virtuales estáticos basados en vídeo 360º. Los flujos Video + Depth (RGB-D) se codifican mediante codecs de vídeo 2D (como H.264 o VP8) y se transmiten vía Web Real Time Communications (WebRTC). El sistema soporta una resolución de 1080×960 píxeles, que requiere un ancho de banda superior a 5Mbps por cada flujo, y permite sesiones con hasta 3 usuarios simultáneos.
En [2] se presenta una plataforma de comunicaciones holográficas multi-usuario en tiempo real, utilizando también sensores RGB-D, como Kinect. En esta plataforma los usuarios pueden ser capturados por uno o varios sensores, y sus representaciones se transmiten como flujos de Nube de Puntos mediante Socket.io (TCP websockets) a través de una Selective Forwarding Unit (SFU) en la nube. Con una sola cámara, cada flujo requiere entre 5 y 7 Mbps (15 fps y ∼50K puntos por fotograma), observándose cuellos de botella a partir de 4 usuarios por sesión. Esta plataforma, denominada HoloMIT, constituye el punto de partida de este trabajo.

C. Optimización de servicios de comunicaciones holográficas

Aunque las representaciones realistas de usuarios mediante vídeo volumétrico mejoran la experiencia frente al uso de avatares [1], también plantean retos de escalabilidad e interoperabilidad por sus altos requerimientos de procesamiento y ancho de banda [2]. Por ello, la comunidad científica ha centrado esfuerzos en optimizar los servicios de comunicaciones holográficas y XR en general. En este contexto, destacan trabajos recientes orientados a mejorar la escalabilidad y adaptabilidad de estos servicios inmersivos en tiempo real. En primer lugar, se han propuesto estrategias avanzadas de procesamiento y distribución basadas en campo de visión, en el lado cliente, con tal de mejorar la adaptabalidad y estabilidad de servicios de streaming de nubes de puntos. Por ejemplo, [6] y [7] introdujeron técnicas de transmisión adaptativa basadas en *tiling*, dividiendo los flujos volumétricos en secciones cúbicas codificadas a distintas tasas de bits, y entregando las calidades más apropiadas de estos cubos segun la posición y campo de visión del usuario objetivo, así como al ancho de banda disponible, en cada momento. También se han desarrollado optimizaciones gestionadas desde el lado servidor para la transmisión de contenidos XR y servicios de comunicaciones holográfica. En [8] se presentó una versión mejorada de la plataforma presentada en [5], incorporando una Multipoint Control Unit (MCU) en la nube que genera un mosaico fusionado a partir de flujos RGB-D individuales desde cada cliente, usando códecs de vídeo 2D. Evaluada con usuarios simulados, la plataforma soportó hasta 8 participantes (con una resolución de 540×800 píxeles por flujo), disminuyendo la carga de tráfico en los clientes en comparación al uso de una SFU, aunque sin considerar métricas como latencia y fluidez en el estudio presentado. Sin embargao, esta solución no soporta representaciones volumétricas completas, ni aprovecha plenamente las características intrínsicas de los entornos 3D, como son las posiciones y campos de visión variantes. De forma similar, [9] introdujo una MCU para hologramas pre-grabados y entornos 3D, incorporando fusión de escenas, transcodificación dinámica y envío personalizado para cada cliente en función de su posición y campo de visión instantáneo. En [10], se integró una versión evolucionada de esta MCU en la plataforma HoloMIT [2], ejecutando en tiempo real dichas estrategias para generar flujos personalizados según distancias y puntos de vista para cada cliente. Esto permitió prácticamente duplicar el número de usuarios por sesión (desde 7 hasta 13), reducendo el uso de recursos computacionales (CPU, GPU, RAM) y ancho de banda, así como manteniendo la latencia dentro de límites aceptables.

En un contexto relacionado, [11] propuso técnicas para mejorar la escalabilidad y estabilidad de servicios de holoconferencia multi-usuario basados en web y vídeo 2D, mediante: (i) un esquema de QoS que prioriza flujos según la posición de los usuarios; (ii) filtrado por *frustum culling*; y (iii) un gestor de recursos que distribuye conexiones entre servidores según la proximidad en el entorno virtual. Además, [12] propuso una arquitectura para servicios colaborativos multi-usuario en XR, basada en renderizado remoto y sincronización Cloud-Edge, logrando mejoras significativas en cuanto a escalabilidad (hasta un 50 % más de usuarios) y rendimiento (frecuencia de renderizado casi 8 veces superior) frente a soluciones previas. De forma complementaria, [13] mostró que el renderizado remoto para habilitar clientes web ligeros en servicios de comunicaciones holográficas puede añadir unos 200ms de latencia, considerada aceptable para experiencias interactivas en tiempo real.

Finalmente, [14] ofrece una revisión detallada sobre habilitadores tecnológicos para servicios XR en la nube,

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

incluyendo avances en estandarización como European Telecommunications Standards Institute (ETSI) Multi-access Edge Computing (MEC), nuevas funcionalidades en 3rd Generation Partnership Project (3GPP) para mejorar la QoS en XR, y nuevos formatos y protocolos Moving Picture Experts Group (MPEG).

D. *Contribuciones de este trabajo*

Este artículo parte de la plataforma HoloMIT [2], y se centra en su extensión y evolución con tal de soportar arquitecturas de despliegue modulares y asistidas por computación en el borde (Edge) que permitan un instanciamiento dinámico de SFUs para sesiones compartidas, mejorando así el rendimiento, la escalabilidad y la sostenibilidad de los servicios de comunicación holográfica. Aunque no se introduce un nuevo componente tecnológico específico para procesamiento XR en la nube, las contribuciones pueden considerarse un habilitante de escalabilidad *per se*, y son compatibles con futuras extensiones que integren funcionalidades de procesado en la nube / borde como, por ejemplo una MCU [10] o un Renderizador Remoto [13], facilitando su incorporación en fases posteriores de investigación.

III. PLATAFORMA EXTENDIDA

A. *Plataforma de Partida: HoloMIT*

A fecha de inicio de este trabajo, el componente Orquestador (Orchestrator) de HoloMIT se basaba en una implementación monolítica, integrando en un único paquete software las funcionalidades de gestión de usuarios, sesiones, conexiones y reenvío de datos multimedia mediante SFU [2]. Aunque funcional, esta implementación claramente no era la más adecuada ni eficiente, ya que no permitía una escalabilidad efectiva. Además, combinaba funciones del plano de control y del plano de datos en un mismo módulo.

B. *Orquestrador Modular*

Este trabajo introduce tres mejoras clave para el componente Orquestador, mediante la especificación y adopción de una arquitectura modular y desacoplada (véase Figura 1). Primero, se ha desacoplado el plano de control del plano de datos de usuario, permitiendo instancias únicas y reutilizables de servicios generales (como gestión de usuarios y sesiones) en la nube, así como instancias múltiples bajo demanda de servicios / módulos específicos (como las SFUs). Segundo, todos los nuevos servicios modulares han sido virtualizados, lo que permite su despliegue dinámico en servidores Edge o en la nube, en base a recursos y demandas existentes, por cada sesión. Tercero, se han incorporado nuevos servicios para integrarse con plataformas de orquestación en el Edge y otros elementos / APIs de red para exponer recursos y funcionalidades de interés, como detección de congestión o estrategias de selección de servidores Edge en los que instancar funcionalidades de procesado multimedia específicas.

A continuación, se presentan los diferentes módulos que forman parte de los servicios que lo componen:

- **User Manager**: Responsable del registro, gestión y provisión de datos de los clientes, escenarios y otros componentes en la nube, utilizando MongoDB [15].
- **Session Manager**: Encargado de gestionar el ciclo de vida de las sesiones multi-usuario (creación, unión, abandono y destrucción), tanto para cada cliente como para cada escenario virtual seleccionado, almacenando la información correspondiente en MongoDB.
- **Clock Manager**: Responsable de garantizar sincronización entre todas las entidades involucradas en la sesión multimedia. En caso de no disponer de acceso a un servidor Network Time Protocol (NTP), permite proporcionar sincronización mediante un protocolo ad-hoc de intercambio y alineamiento de marcas de temporización, sin necesidad de disponer de conexión a Internet.
- **Index Manager**: Nuevo manager que decide y orquesta los servidores en el Edge o en la nube donde se desplegarán managers del plano de usuario (como SFUs, MCUs), y gestiona su ciclo de vida en coordinación con una plataforma de orquestación de recursos Edge. También actúa como interfaz de acceso para explotar funcionalidades / APIs de red específicas (por ejemplo, detección de congestión, identificación de ubicación, descubrimiento y selección de nodos Edge) [16].

C. *Arquitectura Multi-SFU*

En este trabajo, la SFU ha sido evolucionada e integrada como un conjunto de módulos que actúan como managers o servicios:

- **Media Manager**: Gestiona el intercambio de flujos audiovisuales entre los clientes (es decir, actúa como una SFU para audio y vídeo). Puede interactuar con otras SFUs si es necesario e identificar tuplas *[dirección Internet Protocol (IP), puerto, protocolo]* para cada conexión activa, permitiendo así que otros elementos y APIs de red apliquen ciertas optimizaciones (por ejemplo, Access Traffic Steering, Switching and Splitting (ATSSS), slicing, etc.).
- **Events Manager**: Encargado de reenviar eventos o metadatos relevantes entre los clientes, como sus posiciones en el espacio virtual o información de control relacionada con la sesión multimedia (p.ej., interacciones con el entorno, actualizaciones de estado, etc.).

D. *Cliente web sin interfaz gráfica (headless)*

La disponibilidad de clientes / reproductors completos en plataformas nativas y dispositivos con disponibilidad de GPU es esencial para ofrecer servicios de comunicación holográfica multi-usuario de manera eficiente, debido a los exigentes requisitos de ancho de banda y procesamiento [2], [10]. En la plataforma HoloMIT, el player nativo se

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

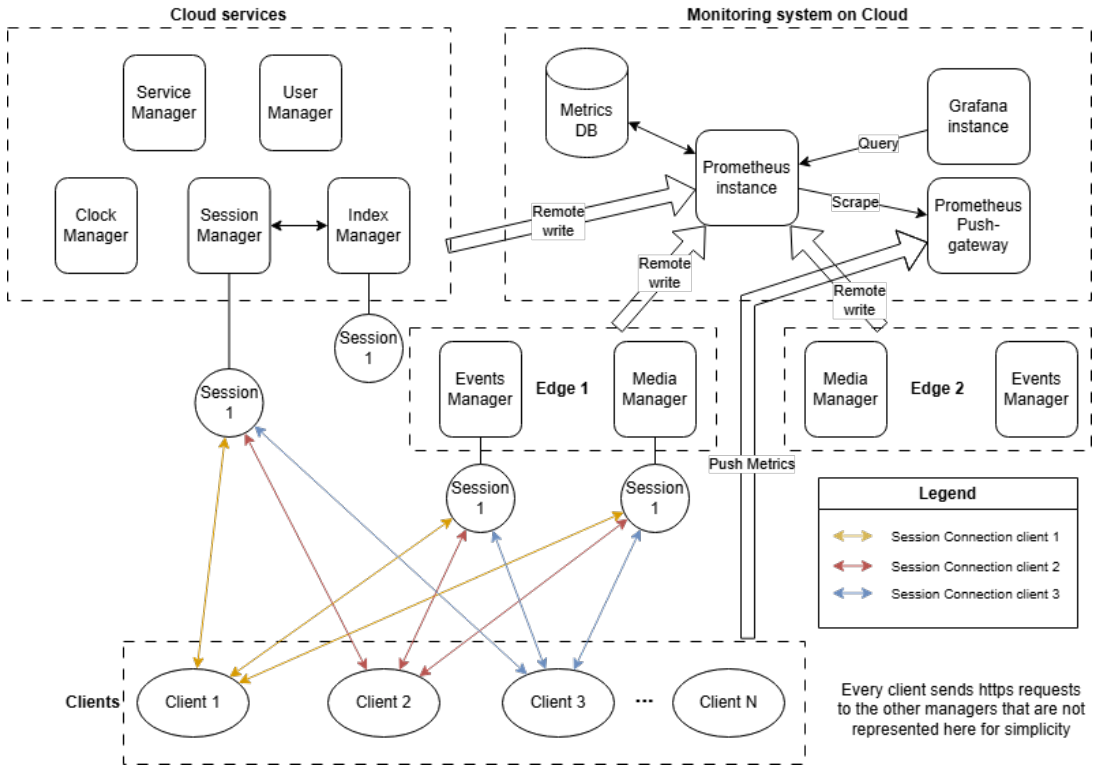


Figura 1: Nueva versión del orquestrador modular con un subsistema de métricas integrado

ha implementado en Unity y está diseñado para sistemas operativos Windows, debido a dependencias específicas de ciertas librerías y módulos software. Sin embargo, utilizar reproductores nativos completos no resulta eficiente en términos de esfuerzo y coste para realizar pruebas extensivas de escalabilidad y rendimiento. Por este motivo, en este trabajo se ha desarrollado un nuevo reproductor ligero sin interfaz gráfica (*headless*) basado en tecnologías web. El objetivo es disponer de un reproductor que requiera menor consumo de recursos (espacio, GPU y memoria), y seea más fácil y ágil de desplegar para pruebas experimentales. Además, dado que este trabajo se centra principalmente en pruebas objetivas de rendimiento y uso de recursos, no es necesario que los hologramas se capturen en origen ni se visualicen en destino, lo que permite omitir funciones de procesamiento intensivo como la captura, codificación, decodificación y el renderizado, en línea con las características de un cliente *headless*.

Concretamente, el cliente web *headless* se ha implementado mediante Node.js [17], emulando el comportamiento y las funcionalidades de transmisión y recepción del reproductor nativo, interactuando con los managers necesarios del Orchestrator y comunicándose con el resto de clientes de la sesión a través de los managers de comunicación seleccionados (p.ej. Media o Events Manager).

E. Subsistema de métricas

Se ha desarrollado y adoptado un subsistema de métricas (véase Figura 1) basado en Prometheus [18] y Grafana [19], con el fin de medir, registrar y monitorizar métricas de rendimiento y uso de recursos de distintos componentes de HoloMIT, durante sesiones de prueba.

Se emplean distintos exportadores de Prometheus para recopilar métricas de los módulos de interés. Por ejemplo, Node Exporter [20] proporciona métricas como uso de CPU, memoria y ancho de banda, mientras que Cadvisor [21] ofrece métricas relacionadas con contenedores Docker, como el ancho de banda utilizado por componentes como la SFU. Además, se ha integrado Prometheus Push Gateway [22] para exportar métricas desde el cliente nativo Unity y el cliente web *headless*. Adicionalmente se han adoptado dos exportadores adicionales para obtener métricas de los clientes nativos Unity: (i) Windows Exporter [23], que resulta más adecuado para sistemas Windows que Node Exporter; y (ii) dcgm-exporter [24], útil para recopilar métricas relacionadas con la GPU.

Dado que Prometheus funciona como un sistema de métricas por extracción (*pull*), se ha desplegado una instancia de Prometheus como agente en cada máquina. Esto permite realizar escrituras remotas (esto es, *remote writing*) hacia una instancia general de Prometheus y visualizar todas las métricas en un único lugar, sin necesidad de conocer la IP de cada máquina, lo cual representa una ventaja clara para la ejecución de pruebas.

A continuación, se enumeran las principales métricas que se soportan:

- **SFU**: latencia de procesado por fotograma (ms), desviación de sincronización con el Clock Manager (ms), uso de CPU (%), uso de memoria (MB), uso

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

de ancho de banda (Mbps).

- **Cliente web *headless***: latencia desde transmisión hasta recepción (ms), desviación de sincronización con el Clock Manager (ms), uso de CPU (%), uso de memoria (MB), uso de ancho de banda (Mbps).
- **Cliente nativos Unity**: tiempo desde captura hasta renderizado (ms), tiempo de codificación (ms), fps de codificación, tiempo de decodificación (ms), fps de decodificación, desviación de sincronización con el Clock Manager (ms), uso de CPU (%), uso de GPU (%), uso de memoria (MB), uso de ancho de banda (Mbps).

IV. RESULTADOS

A continuación se presentan resultados para 2 tipos de tests realizados: (i) comparación entre el uso de una SFU cercana vs una SFU más lejana; (ii) comparación de límites de escalabilidad cuando se usa una sola SFU vs cuando se utilizan múltiples SFUs para una sesión dada.

A. *Test 1. SFU cercana vs SFU lejana*

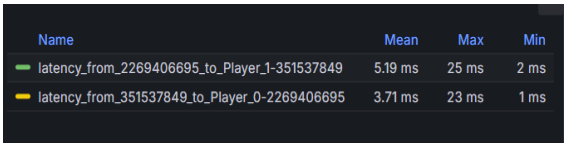
A1. Descripción y Objetivos: Se llevan a cabo sesiones de 2 usuarios cuando se usa una SFU cercana (p.ej. en la misma red local) vs una SFU más lejana (esto es, en una red remota).

El objetivo es evaluar la existencia de diferencias en términos técnicos y beneficios potenciales (p.ej., latencia, estabilidad, consumo de recursos) cuando se posibilita el uso de SFUs en servidores Edge o en la nube. Para ello, se recrean y evaluan dos escenarios análogos con 2 usuarios, uno con una SFU cercana y otro con una SFU remota. En ambas condiciones de prueba se miden y comparan métricas relevantes como latencia, fps y consumo de ancho de banda, durante sesiones de aproximadamente 5 minutos.

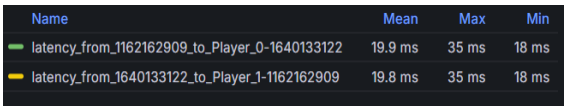
A2. Configuración y Despliegue: Para ambas condiciones de test se configura una sesión de HoloMIT con 2 clientes remotos (web sin interfaz gráfica) y una SFU (Media / Events Manager). Para la condición de test de SFU cercana tanto los clientes como la SFU se despliegan en la misma infraestructura de Azure en el Norte de Europa (Irlanda). Para la condición de test de SFU lejana, los clientes se trasladan a la región Oeste de Europa.

A3. Resultados:

A3a. ***Test 1.1 - SFU cercana****:* Como era de esperar, los niveles de latencia obtenidos fueron bajos, aunque con ligeras fluctuaciones (véase Figura 2a) posiblementes por desviaciones en el proceso de sincronización de relojes. En particular, la latencia media fue de aproximadamente 4 ms, y la latencia de procesamiento introducida por la SFU fue de 1.2 ms, un valor muy bajo pero relevante para esta prueba. Este retardo de procesamiento se midió mediante Wireshark [25], aunque debe tenerse en cuenta que las mediciones asociadas pueden verse afectadas por el reloj interno de la máquina, interrupciones de la tarjeta de red o multitarea del sistema operativo. Aun así, ofrecen una buena estimación del procesamiento de la SFU.



(a) Latencia en el Test 1.1 (SFU cercana)



(b) Latencia en el Test 1.2 (SFU lejana)

Figura 2: Resultados de latencia experimentada cuando se usa SFU cercana vs lejana

A3b. ***Test 1.2 - SFU lejana****:* La Figura 2b muestra los valores de latencia experimentados por los clientes al conectarse a una SFU lejana (clientes en Europa Occidental y SFU en Europa del Norte). Como era previsible, la latencia aumentó con respecto al Test 1.1, situándose en torno a los 18 ms, con ligeras fluctuaciones atribuibles también a la sincronización de relojes.

Adicionalmente, también se comprobó experimentalmente que la diferencia de retardos entre la selección de diferentes servidores Azure desde una localización dada para un cliente local externo puede alcanzar niveles en torno a 200ms. Ello refleja la relevancia de poder selección el servidor de interés para ofrecer servicios de comunicaciones en tiempo real, sensibles al retardo.

B. *Test 2. Escalabilidad con una sola SFU*

B1. Descripción y Objetivos: Se persigue determinar el límite de escalabilidad cuando se adopta una única SFU por sesión, dada una configuración especifífica de despliegue (véase Figura 3), mediante el aumento progresivo del número de usuarios por sesión. Se empieza con sesiones de 2 usuarios, y se va incrementando el número de usuarios de dos en dos en cada iteración del test. Una vez alcanzado el límite superior, se puede reducir el número en una unidad para identificar el número máximo de usuarios que aún permite un rendimiento estable.

B2. Configuración y Despliegue: Dado que las especificaciones y recursos del servidor que ejecuta la SFU pueden desempeñar un papel clave, se han llevado a cabo condiciones de prueba utilizando dos máquinas virtuales de Azure con diferentes configuraciones.

Los clientes sin interfaz gráfica se ejecutaron en una máquina virtual de Azure con especificaciones Standard_DS1_v2, desplegada en Europa Occidental. Los subcomponentes del Orquestador y los módulos/managers de la SFU también se desplegaron en una máquina virtual de Azure, en este caso ubicada en Europa del Norte. En una de las condiciones de prueba se utilizó una máquina virtual Standard_DS1_v2, mientras que en otra se empleó una Standard_DS2_v2 para ejecutar la SFU, con tal de evaluar el impacto de la diferencia der recursos y capacidades del

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

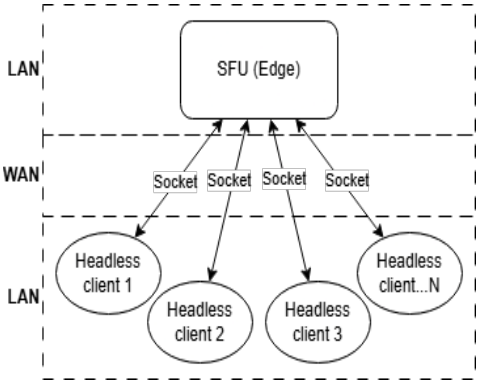


Figura 3: Configuración de test para el límite de escalabilidad con SFU única

servidor en la nube en términos de escalabilidad de la sesión.

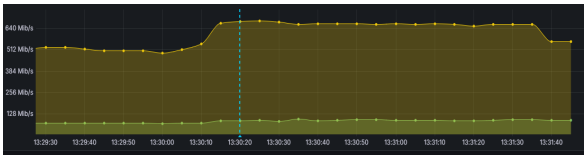
En cada una de las condiciones de test, los clientes sin interfaz gráfica se han añadido progresivamente a una sesión compartida. Al alcanzarse el límite de escalabilidad (es decir, cuando se observan fallos), se reduce el número de clientes en una unidad. Si la sesión resulta estable, ese es el número de clientes que determina el límite de escalabilidad en régimen estable; en caso contrario, se considera el número de clientes antes de los fallos. En cada condición de test, se miden y comparan métricas clave como los retardos, el consumo de ancho de banda y el uso de recursos computacionales, durante sesiones de aproximadamente 5 minutos.

Además, se contempla otra condición de test adicional, que consiste en el uso de sockets volátiles (sin retransmisión) para determinar si bajo esta configuración de streaming mejoran las prestaciones; en caso de ser así, se adoptará esta configuración para cada una de las sesiones de test.

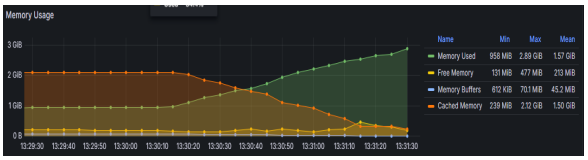
B3. Resultados:

B3a. ***Test 2.1****:* En esta condición de test, la diferencia entre el ancho de banda recibido y transmitido por la SFU, así como la cantidad total de datos intercambiados, comenzó a ser significativa al llegar a 8 clientes por sesión, lo que podría representar una barrera en ciertos escenarios de despliegues. En esta situación, mientras que la SFU recibe $N_{Clients} = 8$ flujos (aproximadamente 75,5 Mbps), debe transmitir $N_{Clients} * (N_{Clients} - 1) = 56$ flujos (más de 500 Mbps) a todos los clientes remotos. A pesar de ello, la sesión se ejecutó sin problemas.

Al llegar a 10 clientes, ya se empezaron a percibieron problemas de estabilidad y escalabilidad. En este caso, la SFU debería recibir $N_{Clients} = 10$ flujos (alrededor de 100 Mbps) y transmitir $N_{Clients} * (N_{Clients} - 1) = 90$ flujos (aproximadamente 1 Gbps) a todos los participantes. Sin embargo, como se observa en la Figura 4a, el ancho de banda transmitido no superó los ∼682 Mbps cuando se unieron los clientes 9 y 10, ya que el límite de la máquina virtual adoptada para desplegar la SFU era de 750 Mbps, y ya se estaban recibiendo ∼105 Mbps en el canal de subida. En este caso, la SFU experimentó sobrecargas de memoria debido a su incapacidad para gestionar el volumen total de datos (véase Figura 4b), lo que provocó el apagado automático del contenedor Docker para preservar los recursos de la máquina virtual. Este fue el punto de fallo para el Test 2.1.



(a) Fallo de Test 2.1 con 10 usuarios – valores de ancho de banda



(b) Fallo de Test 2.1 con 10 usuarios – valores de memoria

Figura 4: Resultados del Test 2.1

Tras analizar los resultados de este primer Test 2.1, se observó que el desbordamiento de memoria (Figura 4b) fue potencialmente provocado por la retransmisión de segmentos TCP (socket.io). Así pues, se consideró el uso de sockets volátiles con el fin de evaluar posibles mejoras. En este condición de test, aunque no se logró aumentar el número de clientes con dicha configuración, no se produjo desbordamiento de memoria, aunque sí se registró una pérdida significativa de paquetes. Ello también puede considerarse un límite operativo, aunque resulta mucho más estable y conveniente que la inhabilitación del contenedor por exceso de uso de memoria.

B3b. ***Test 2.2****:* En este caso, el uso de una máquina virtual con especificaciones más altas (Standard_DS2_v2) para el despliegue de la la SFU resultó en el soporte de un mayor número de usuarios por sesión, llegando hasta 12 usuarios sin pérdidas ni estabilidades significativas. Esto de alguna manera confirma que el límite en Test 2.1 fue más debido a la tarjeta de red que no a la CPU o la memoria. Sin el uso de sockets volátiles, ocurrió una situación similar con 13 usuarios a la ocurrido con 10 usuarios en el Test 2.1. Esto valida que las fórmulas analíticas dan una idea aproximada adecuada del límite de escalabilidad según el ancho de banda máximo disponible, además de que el límite siempre se encuentra en el ancho de banda transmitido.

C. Test 3. SFU única vs SFU múltiples

C1. Descripción y Objetivos: Tras identificar los límites de escalabilidad con una única SFU, se preparan escenarios con múltiples SFUs por sesión para evaluar si se obtienen mejoras en cuanto a escalabilidad y, en tal caso, en qué medida. También se considera el despliegue de las SFUs en distintas regiones continentales. En cada una de las condiciones experimentales, al igual que para los tests anteriores, los clientes sin interfaz gráfica se

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

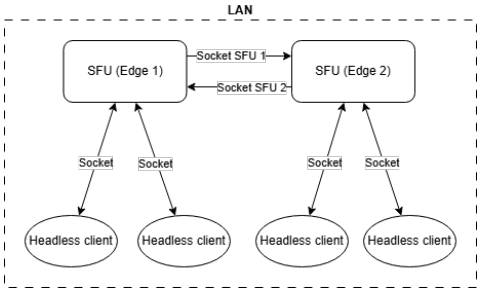
incorporan progresivamente en grupos de dos a una sesión compartida, distribuyéndolos de manera equilibrada entre las distintas SFUs. Para cada condición de prueba, se miden y analizan métricas relevantes como la latencia, el consumo de ancho de banda y el uso de recursos computacionales, durante sesiones de aproximadamente 5 minutos.

C2. Configuración y Despliegue: Se consideran despliegues de la SFU en distintas regiones geográficas. Como se muestra en la Figura 5a, se lleva a cabo una primera prueba (Test 3.1) interconectando dos SFUs que atienden a los clientes de una misma sesión multi-usuario. Posteriormente, se realiza una prueba similar (Test 3.2), sin interconectar las SFUs, pero modificando el comportamiento de los clientes para que puedan recibir datos desde más de una SFU en el canal de bajada, como se ilustra en la Figura 5b. En este segundo caso, los clientes siguen enviando sus datos a una única SFU en el canal de subida, pero reciben datos de otros clientes desde múltiples SFUs. Finalmente, se realiza una tercera prueba (Test 3.3) desplegando tres SFUs en distintas regiones de Azure. Esta prueba tiene un doble objetivo: (i) evaluar si el uso de tres SFUs mejora la escalabilidad en comparación con escenarios con dos SFUs; y (ii) comprobar si el efecto multiplicativo en el canal de bajada puede mitigarse al desplegar las SFUs en diferentes regiones y, por tanto, en redes distintas. En este caso, las conexiones entre SFUs seguirán el patrón del Test 3.2, ya que tanto empíricamente como analíticamente se comprobó que se reduce la sobrecarga de red en comparación con el Test 3.1.

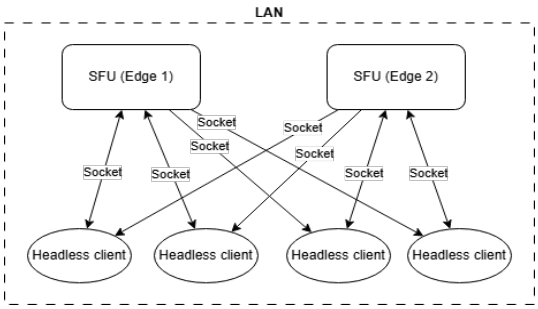
En cuanto al despliegue específico, los subcomponentes del Orquestador y los módulos o managers de SFU fueron desplegados en una máquina virtual Azure Standard_DS1_v2. En los Test 3.1 y 3.2, las SFUs se desplegaron en la región de Europa del Norte, mientras que en el Test 3.3 se distribuyeron entre las regiones de Europa del Norte, Europa Occidental e Italia del Norte, con el objetivo de evaluar sesiones distribuidas geográficamente. Los clientes sin interfaz gráfica también se desplegaron en Azure (en Europa Occidental para los Test 3.1 y 3.2, y distribuidos entre regiones en el Test 3.3).

C3. Resultados:

C3a. ***Test 3.1 - 2 SFUs interconectadas****:* En una condición de prueba con 4 clientes (2 por SFU), los valores analíticos esperados de ancho de banda son 76.8 Mbps para la transmisión y 38.4 Mbps para la recepción. Estos valores teóricos presentan una alta correlación con los valores obtenidos experimentalmente, que fueron de 76.2 Mbps y 38.65 Mbps para el ancho de banda transmitido y recibido, respectivamente. En una condición de prueba con 12 clientes (6 por SFU) se está cerca del límite superior. Mientras que el ancho de banda recibido coincide con el valor analítico (alrededor de 115 Mbps), el ancho de banda transmitido real (aproximadamente 600 Mbps) fue ligeramente inferior al valor teórico (691.2 Mbps). Esto indica que se alcanzó el límite de escalabilidad. No obstante, se realizó una prueba adicional con 13 usuarios



(a) SFUs interconectadas



(b) SFUs no interconectadas

Figura 5: Configuración de Test 3

para determinar con mayor precisión el límite operativo de la SFU. Los resultados se muestran en la Figura 6, en la que puede observarse que una de las SFUs alcanza su límite de capacidad, lo cual se evidencia en la línea casi recta del gráfico correspondiente al ancho de banda transmitido por la SFU en cuestión, indicando que no es posible enviar más datos y que se produce una pérdida de paquetes.

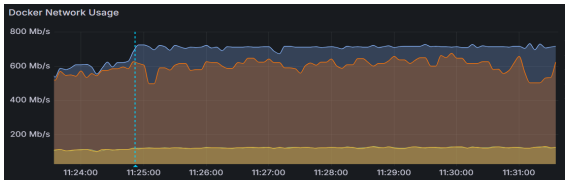


Figura 6: Ancho de banda del Test 3.1 con 13 clientes

C3b. ***Test 3.2 - Dos SFU no interconectadas por sesión****:* En esta test, los clientes se añadieron progresivamente a la sesión y se asignaron de manera balanceada a una de las dos SFUs para la transmisión de datos, aunque recibían datos de ambas SFUs. En una condición de prueba con 4 clientes (2 por SFU), los valores experimentales de ancho de banda transmitido y recibido mostraron una alta correlación con los valores analíticos (19.2 Mbps y 57.6 Mbps, respectivamente), lo que evidencia un rendimiento satisfactorio. Con 12 clientes (6 por SFU), la prueba se mantuvo estable, con valores de ancho de banda recibido muy próximos a los analíticos (57 Mbps). No obstante, ya se observaron desviaciones en el ancho de banda transmitido respecto al valor teórico (633 Mbps). Se añadió

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

un cliente adicional para evaluar el rendimiento en dicha condición de test. Con 13 clientes, las desviaciones en el ancho de banda transmitido fueron más notables que con 12 clientes, aunque la evolución del ancho de banda se mantuvo relativamente estable. Con 14 clientes, el ancho de banda no aumentó en comparación con las condiciones de test anteriores, lo que indica que se alcanzaron los límites de capacidad de la SFU. Por tanto, se puede considerar que el límite de escalabilidad y estabilidad para este Test 3.2 se sitúa en 12 clientes.

C3c. ***Test 3.3 - Tres SFUs no interconectadas por sesión en distintas regiones****:* Como primera condición de test, se ejecutó una sesión con 12 usuarios (4 conectados a cada SFU). En la Figura 7 se observa que se obtuvieron tres grupos principales de valores de latencia, uno por región, como era de esperarse. No obstante, estas diferencias de retardo se mantienen dentro de los márgenes aceptables para comunicaciones en tiempo real exitosas [10]. Posteriormente, se realizó una condición de test con 15 clientes (5+5+5), la cual también mostró un rendimiento estable y consistente de las SFUs en términos de transmisión y recepción de ancho de banda. Los valores de latencia fueron similares a los observados con 12 usuarios (véase Figura 7). Finalmente, se ejecutó una condición de test con 18 clientes (6+6+6), la cual evidenció limitaciones de rendimiento tanto en la gestión del ancho de banda por parte de las SFUs como en los niveles de latencia experimentados por los clientes.

Así pues, se confirma que el número de clientes conectados puede incrementarse mediante el despliegue de más SFUs en distintas regiones geográficas, y que los niveles de latencia asimétricos experimentados siguen siendo aceptables en este tipo de escenarios.

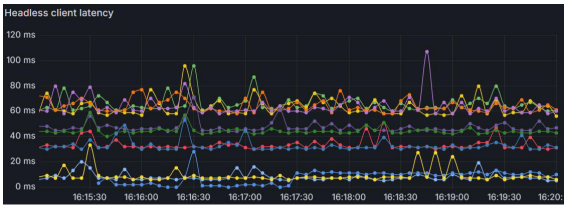


Figura 7: Resultados para el Test 3.3 con 12 clientes

V. CONCLUSIONES

En este trabajo se ha evolucionado una plataforma de comunicaciones holográficas multiusuario pionera con nuevas contribuciones que contribuyen a mejorar su rendimiento en cuanto a escalabilidad, estabilidad, retardos y estabilidad, mediante la adopción adaptativa de arquitecturas modulares asistidas por computación en el borde y APIs de red. Las pruebas de validación realizadas aportan evidencias claras sobre los beneficios aportados, tanto por la posibilidad de desplegar y seleccionar módulos de comunicaciones en los servidores de interés (ej. mimimizar retardos, garantizar recursos necesarios...), como por la posibilidad de instanciar nuevos módulos de comunicaciones en diferentes servidores para conseguir un mejor balanceo de carga y mayor escalabilidad por cada sesión activa .

VI. TRABAJOS FUTUROS

En primer lugar, se pretende extender las contribuciones de este trabajo mediante la integración de nuevos mecanismos de distribución multimedia basado en posiciones relativas y campos de visión de cada usuario. En segundo lugar, se pretende incorporar servicios de transcodificación virtualizados, instanciables bajo demanda para cada flujo de vídeo volumétrico entrante y/o cada cliente objetivo. Estos dos nuevos habilitadores tecnológicos permitirán incrementar la adaptabilidad y escalabilidad de los servicios de comunicaciones holográficas, de una manera más ligera, modular y escalable que cuando se adaptan estrategias de transcodificación en la nube, mediante el uso de MCUs monolíticas desplegadas para cada sesión. Finalmente, se tratará de explotar el potencial de APIs de red, siendo especificadas en el seno de iniciativas de estandarización como CAMARA (https://camaraproject.org/).

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por: el Programa Horizonte Europa de la Unión Europea, bajo el acuerdo nº 101135637 (proyecto PRESENCE), y por el Programa Smart Networks and Services Joint Undertaking (SNS JU), bajo el acuerdo nº 101096838 (proyecto 6G-XR) y nº 101139257 (proyecto SUNRISE-6G); por la Agencia Estatal de Investigación (AEI), en el marco de Proyectos Generación de Conocimiento 2021 (PID2021-126551OB-C21) y 2022 (PID2022-140749OB-I00 y PID2022-137329OB-C41, financiados por MICIU/AEI/10.13039/501100011033 y FEDER, UE); y por el Ministerio para la Transformación Digital y la de Función Pública, en el marco del programa UNICO 5G I+D 2021, bajo el proyecto con Ref. TSI-063000-2021-4 (6G-OPENVERSO-HOLO). El trabajo de Mario Montagud ha sido financiado por MCIN/AEI/10.13039/501100011033, bajo la subvención RYC2020-030679-I.

REFERENCIAS

[1] M. Montagud, J. Li, G. Cernigliaro, A. El Ali, S. Fernández, and P. Cesar, "Towards socialvr: evaluating a novel technology for watching videos together," *Virtual Real.*, vol. 26, no. 4, p. 1593–1613, Dec. 2022. [Online]. Available: https://doi.org/10.1007/s10055-022-00651-5

[2] S. F. Langa, M. Montagud, G. Cernigliaro, and D. R. Rivera, "Multiparty holomeetings: Toward a new era of low-cost volumetric holographic meetings in virtual reality," *IEEE Access*, vol. 10, pp. 81 856–81 876, 2022.

[3] Y. Jin, K. Hu, J. Liu, F. Wang, and X. Liu, "From capture to display: A survey on volumetric video," 2023. [Online]. Available: https://arxiv.org/abs/2309.05658

[4] J. van der Hooft, H. Amirpour, M. T. Vega, Y. Sanchez, R. Schatz, T. Schierl, and C. Timmerer, "A tutorial on immersive video delivery: From omnidirectional video to holography," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 1336–1375, 2023.

[5] S. N. B. Gunkel, H. M. Stokking, N. Prins, Martin Jand van der Stap, F. B. T. Haar, and O. A. Niamut, "Virtual reality conferencing," in *Proceedings of the 9th ACM Multimedia Systems Conference*. New York, NY, USA: ACM, Jun. 2018.

[6] J. Park, P. A. Chou, and J.-N. Hwang, "Rate-utility optimized streaming of volumetric media for augmented reality," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 9, no. 1, pp. 149–162, Mar. 2019.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

[7] S. Subramanyam, I. Viola, A. Hanjalic, and P. Cesar, "User centered adaptive streaming of dynamic point clouds with low complexity tiling," in *Proceedings of the 28th ACM International Conference on Multimedia*. New York, NY, USA: ACM, Oct. 2020.

[8] S. N. B. Gunkel, R. Hindriks, K. M. E. Assal, H. M. Stokking, S. Dijkstra-Soudarissanane, F. t. Haar, and O. Niamut, "Vrcomm: an end-to-end web system for real-time photorealistic social vr communication," in *Proceedings of the 12th Multimedia Systems Conference*, ser. MMSys '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 65–79. [Online]. Available: https://doi.org/10.1145/3458305.3459595

[9] G. Cernigliaro, M. Martos, M. Montagud, A. Ansari, and S. Fernandez, "Pc-mcu: point cloud multipoint control unit for multi-user holoconferencing systems," in *Proceedings of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, ser. NOSSDAV '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 47–53. [Online]. Available: https://doi.org/10.1145/3386290.3396936

[10] S. Fernandez, M. Montagud, D. Rincón, J. Moragues, and G. Cernigliaro, "Addressing scalability for real-time multiuser holoportation: Introducing and assessing a multipoint control unit (mcu) for volumetric video," in *Proceedings of the 31st ACM International Conference on Multimedia*, ser. MM '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 9243–9251. [Online]. Available: https://doi.org/10.1145/3581783.3613777

[11] M. Dasari, E. Lu, M. W. Farb, N. Pereira, I. Liang, and A. Rowe, "Scaling VR video conferencing," in *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE, Mar. 2023.

[12] Y. Huang, H. Wang, X. Qiao, X. Su, Y. Li, S. Dustdar, and P. Zhang, "SCAXR: Empowering scalable Multi-User interaction for heterogeneous XR devices," *IEEE Netw.*, vol. 38, no. 4, pp. 250–258, Jul. 2024.

[13] I. Yeregui, D. Mejías, G. Pacho, R. Viola, J. Astorga, and M. Montagud, "Edge rendering architecture for multiuser XR experiences and E2E performance assessment," in *2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, Jun. 2024, pp. 1–7.

[14] T. Theodoropoulos, A. Makris, A. Boudi, T. Taleb, U. Herzog, L. Rosa, L. Cordeiro, K. Tserpes, E. Spatafora, A. Romussi, E. Zschau, M. Kamarianakis, A. Protopsaltis, G. Papagiannakis, and P. Dazzi, "Cloud-based XR services: A survey on relevant challenges and enabling technologies," *J-NaNA*, vol. 2, no. 1, pp. 1–22, 2022.

[15] MongoDB, "Mongodb," last Access in: August 2024. [Online]. Available: https://www.mongodb.com/

[16] e. a. M. Montagud, "Awarexr: A naas architecture to enhance xr services over beyond 5g networks," , IEEE Network, To Appear in 2025.

[17] nodejs.org, "Node.js," last Access in: August 2024. [Online]. Available: https://nodejs.org/en

[18] Prometheus, "Prometheus," last Access in: August 2024. [Online]. Available: https://prometheus.io/

[19] Grafana, "Grafana," last Access in: August 2024. [Online]. Available: https://grafana.com/

[20] Prometheus, "Node exporter," last Access in: August 2024. [Online]. Available: https://github.com/prometheus/node_exporter

[21] Google, "Cadvisor," last Access in: August 2024. [Online]. Available: https://github.com/google/cadvisor

[22] Prometheus, "Prometheus push gateway," last Access in: August 2024. [Online]. Available: https://github.com/prometheus/pushgateway

[23] P. Community, "Windows exporter," last Access in: August 2024. [Online]. Available: https://github.com/prometheus-community/windows_exporter

[24] Nvidia, "Dcgm-exporter," last Access in: August 2024. [Online]. Available: https://github.com/NVIDIA/dcgmexporter

[25] Wireshark, "Wireshark," last Access in: September 2024. [Online]. Available: https://www.wireshark.org/download.html

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

XVII Jornadas de Ingeniería Telemática
JITEL 2025
Universidad de Extremadura

Actas de las XVII Jornadas de Ingeniería Telemática (JITEL 2025), Cáceres (España), 12-14 de noviembre de 2025.

ISBN

Optimización del Despliegue de SDNs Mediante Algoritmos de Colonias de Hormigas

Luis M. Chaves, Antonio M. Mora, Juan F. Valenzuela-Valdés
Departamento de Teoría de la Señal, Telemática y Comunicaciones
Universidad de Granada
ETSIIT, CITIC-UGR, Granada, España
lmchaves@correo.ugr.es, amorag@ugr.es, juanvalenzuela@ugr.es

Las Redes definidas por software (SDN, por *Software Defined Networks*) se han estandarizado en los últimos años. Dentro de éstas, su despliegue efectivo sigue siendo un problema complejo para los administradores, cuando se espera conseguir un balanceo de carga adecuado, atendiendo a criterios de minimización de retardo, mientras se maximiza el aprovechamiento de la capacidad de los enlaces.

Este trabajo propone un algoritmo metaheurístico de Optimización basada en Colonias de Hormigas (ACO, por *Ant Colony Optimization*), aplicado para diseñar el mejor despliegue de una SDN sobre el controlador Ryu de OpenFlow, considerando restricciones de retardo, carga de los enlaces y pérdida de paquetes, al tiempo que se pretende minimizar la ruta para el tráfico entre dos nodos de la red. El algoritmo propuesto ha demostrado un gran rendimiento, siendo capaz de encontrar soluciones óptimas (según dichas restricciones) en escenarios dinámicos. Además, se ha desarrollado una herramienta visual que ayuda al investigador (y al administrador de la red) a monitorizar la evolución de las soluciones de dicho algoritmo, pudiendo fijar algunos parámetros del estado de la red en tiempo real, para ver cómo se comporta dicho algoritmo.

***Palabras Clave*- Redes Definidas por Software, SDN, Algoritmo basado en Colonias de Hormigas, ACO, Balanceo de carga, Load Balancing.**

I. Introducción

El crecimiento exponencial de las redes modernas, debido, entre otros factores, al gran número de usuarios móviles, el big data, la computación en la nube y el Internet de las cosas, ha desembocado en una gestión cada vez más compleja. Para abordar este desafío, surgió hace unos años el concepto de Redes Definidas por Software (SDN) [1], que propone una separación del plano de control del plano de datos, permitiendo una gestión de red centralizada.

Las SDNs, posibilitan aplicar técnicas como la monitorización continua del estado de la red, la reconfiguración dinámica de rutas y servicios en tiempo real.

Aún así, el verdadero reto no reside en recopilar información del plano de datos, sino en tomar decisiones óptimas, como enrutar el tráfico de manera de que minimice la congestión, el equilibrio de las cargas en los enlaces, la reducción de latencias y la pérdida de paquetes.

Junto a esto, el citado crecimiento de las redes exige, a su vez, alta calidad de servicio y una gestión eficiente del tráfico, siendo la ***optimización del balanceo de carga*** indispensable, especialmente en centros de datos [2], ya que los recursos de red son limitados. Dicho balanceo consiste en distribuir el tráfico de la red entre múltiples rutas o servidores con el objetivo de evitar la congestión, maximizar la utilización de los recursos disponibles y mejorar el rendimiento general [3]. En este contexto, los controladores SDN, gracias a su visión global de la red, pueden facilitar la realización de un balanceo de carga óptimo.

Este problema es de naturaleza NP-completa [4], lo que significa que no puede resolverse de forma exacta en un tiempo razonable para redes de gran tamaño.

Los algoritmos tradicionales de balanceo de carga, como Dijkstra [5] y Round Robin (RR) [6], fueron concebidos para entornos estáticos y son insuficientes para la naturaleza dinámica y compleja de las SDN modernas [7]. Dijkstra, por ejemplo, es un método estático que se basa en heurísticas predefinidas y métricas fijas como el número de saltos o el coste de los enlaces, ignorando factores de la red que pueden cambiar en tiempo, como la carga de los nodos. Esto puede llevar a soluciones subóptimas y a que se produzca congestión, ya que el tráfico se envía continuamente por la ruta "más corta" sin considerar otras opciones. Además, Dijkstra es incapaz de optimizar múltiples métricas simultáneamente, como la pérdida de paquetes, el retardo o el ancho de banda disponible en los enlaces, que son requisitos considerados para garantizar una Calidad de Servicio (QoS) mínima. Por su parte, Round Robin distribuye las peticiones cíclicamente asumiendo capacidades idénticas en todos los servidores o enlaces, lo que no siempre es cierto y puede llevar a un balanceo de carga desigual, así como sobrecarga en los conmutadores con menos recursos.

Estas limitaciones de los algoritmos estáticos provocan una cascada de efectos negativos, incluyendo un aumento de la latencia, la pérdida de paquetes o la reducción del rendimiento general de la red.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

De modo que se hace recomendable la aplicación de algún método de optimización. Entre dichos métodos, las técnicas metaheurísticas bioinspiradas, como la Optimización basada en Colonias de Hormigas o *Ant Colony Optimization* (ACO) [8], emergen como soluciones prometedoras, ya que pueden encontrar soluciones competentes en un tiempo aceptable (limitado) e incluso en escenarios cambiantes dinámicamente. Como su nombre indica, ACO se basa en el comportamiento de las hormigas para encontrar rutas óptimas en la naturaleza. Éstas utilizan una sustancia llamada *feromona* como mecanismo de comunicación y retroalimentación (es lo que se conoce como *estigmergia* [9]). De modo que la aplicación de esta metaheurística, permitiría encontrar rutas eficientes dentro de la red para el envío del tráfico, pudiendo considerar, a su vez, métricas como la pérdida de paquetes, la carga de los nodos y el retardo del envío de datos.

La idea que perseguimos consiste en aprovechar la sinergia entre el control centralizado de SDN y la inteligencia adaptativa y distribuida de algoritmo de hormigas. SDN proporciona una visión global de la red y la capacidad de ajustar dinámicamente las reglas de reenvío, mientras que ACO ofrece la inteligencia para realizar estos ajustes de forma adaptativa, basándose en las condiciones de la red en tiempo real, lo que permite explorar espacios de soluciones complejos para encontrar rutas óptimas y asignar recursos

Existen trabajos anteriores que aplican metaheurísticas a la optimización de SDNs [10], entre ellas, una que plantea el uso de ACO [11], denominada LLBACO; si bien, el modelo de ACO propuesto es el enfoque más simple. Este estudio pretende mejorar dicho trabajo, diseñando y probando una versión de ACO más avanzada y que normalmente proporciona mejores resultados en problemas de optimización: un *Ant Colony System* (ACS) [8]. Dicho algoritmo se integrará dentro de un controlador muy extendido en el ámbito de las SDNs: *Ryu* (https://ryu-sdn.org/) junto con *OpenFlow* (https://opennetworking.org/).

El objetivo principal será desarrollar e integrar un algoritmo ACO que optimice el balanceo de carga en redes SDN (desplegadas con el software *Mininet* (https://mininet.org/)), proponiendo rutas alternativas para el tráfico en función del estado de la red en cada instante.

En segundo lugar, y como complemento de investigación, se ha desarrollado una herramienta web interactiva que permite hacer una monitorización del funcionamiento del ACS sobre el despliegue de la SDN en tiempo real. Esta aplicación también permitirá cambiar factores de la configuración de la red, que harán que la instancia a resolver cambie dinámicamente, como por ejemplo reducir o aumentar el retardo, la tasa de paquetes perdidos o el ancho de banda en un enlace o generar tráfico sintético para aumentar la carga o saturación de los nodos.

Finalmente, se ha llevado a cabo una fase experimental completa en un entorno SDN realista, comparando configuraciones del algoritmo propuesto, así como el enfoque planteado en [11] (LLBACO) y un algoritmo tradicional (Dijkstra).

II. Conceptos Preliminares

Esta sección presenta los conceptos que se utilizarán en el trabajo, como las SDN, así como el Algoritmo de Colonias de Hormigas

A. Redes definidas por Software (SDN)

Las Redes Definidas por Software (SDN) representan un nuevo enfoque en la gestión de redes, ofreciendo una arquitectura más flexible, dinámica y manejable. Todas estas características las hacen muy adecuadas para la naturaleza dinámica de las aplicaciones de red de hoy en día. La gran innovación es que desacopla el plano de control del plano de datos de un dispositivo de red [12].

En este modelo, el plano de control es trasladado de los dispositivos de la red a un controlador lógico centralizado que actúa como sistema operativo de la red. Proporcionando una visión global de la red, que, a diferencia de las redes tradicionales, puede realizar cambios globales de una manera centralizada. Uno de los protocolos más representativos en este contexto es *Openflow* (https://opennetworking.org/), el cual permite una comunicación estandarizada entre el controlador y los dispositivos del plano de datos.

La arquitectura tradicional de SDN está compuesta de tres capas principales [13] (ver Figura 1):

- *Capa de datos (Data plane)*: Es la capa inferior, y está formada por dispositivos de red físicos o virtuales como switches, routers y puntos de acceso. Estos dispositivos no toman decisiones, sino que simplemente reenvían paquetes según las reglas instaladas por el controlador. Se comunican con la capa de control a través de interfaces como el protocolo *OpenFlow*, que permite acceder al plano de reenvío de forma directa y segura (con TLS).
- *Capa de control (Control plane)*: Es la capa intermedia compuesta por controladores SDN basados en software. Estos controladores centralizan la lógica de control de red, monitorean el comportamiento del tráfico y gestionan la infraestructura mediante la interfaz con el plano de datos. También exponen interfaces hacia las aplicaciones de red, conocidas como A-CPI (Application to Controller Plane Interface), que permite a las aplicaciones obtener una visión abstracta del estado de la red y modificar su comportamiento. Un controlador típico puede incluir componentes funcionales como virtualizadores, coordinadores, balanceadores de carga, sistemas de detección de intrusos o monitorización de red.
- *Capa de aplicación (Application plane)*: Es la capa superior, donde se desarrollan e implementan las aplicaciones que utilizan los servicios ofrecidos por el controlador SDN. Estas aplicaciones pueden incluir funciones de ingeniería de tráfico, balanceo de carga, enrutamiento, seguridad o virtualización. Se comunican con el plano de control mediante la interfaz A-CPI, accediendo a una visión global y abstracta de la red para la toma de decisiones.

OpenFlow es un protocolo de comunicación de red fundamental en la arquitectura de redes definidas por software. Se utiliza como un protocolo que facilita la comunicación entre el controlador SDN, que actúa como el "cerebro" de la red, y los dispositivos de reenvío de tráfico, como conmutadores y routers. De hecho, *OpenFlow* se considera el API (*Application Programming Interface*) *southbound* estándar de la industria para la implementación de entornos SDN.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

En este trabajo también se han utilizado dos herramientas adicionales:

- *Mininet*: (https://opennetworking.org/mininet/) un emulador de red el cual crea una red virtual,con hosts, switches, controladores y enlaces. Perfecto para desarrollar y experimentar con SDNs. Los hosts de Mininet se ejecutan en Linux y los switches utilizan Openflow. Una de las grandes ventajas de Mininet es que permite crear prototipos de redes complejas (virtuales) en un único equipo o ejecutar aplicaciones reales en los hosts emulados e interactuar con ellos, acelerando el ciclo de desarrollo y prueba.
- *Ryu*: (https://ryu-sdn.org/) se trata de un controlador de SDN, el cual es un framework de código abierto escrito en Python. El framework facilita el desarrollo proporcionando funcionalidades básicas para controlar la capa de datos y las funciones comunes en las aplicaciones SDN.

  Las interfaces *southbound* (ver Figura 1) son los mecanismos que utiliza Ryu para comunicarse con los dispositivos de red físicos, como los switches SDN (el controlador decide "arriba" y se comunica con los switches "abajo"). Por ejemplo: Ryu usa Openflow para enviar instrucciones a los switches del tipo: "si un paquete viene del switch A y va para el switch B, envíalo por el puerto 3" y también para recibir información de ellos: "ha llegado un paquete nuevo que no sé cómo gestionar".

  Ryu puede incorporar en su lógica interna algoritmos, heurísticas o funciones desarrolladas, como será el caso del presente trabajo.

  Las interfaces *northbound* (ver Figura 1) exponen el controlador Ryu hacia arriba, para que aplicaciones externas puedan interactuar con él y darle instrucciones. Es muy común utilizar la arquitectura REST para ello, para poder interactuar con el controlador a través de peticiones HTTP.

B. *Algoritmos basados en Colonias de Hormigas (ACO)*

Ant Colony Optimization (ACO) es una técnica metaheurística ampliamente utilizada para encontrar soluciones a complejos problemas de optimización. Este algoritmo surgió por la observación del comportamiento de las colonias de hormigas reales [14] al buscar las mejores rutas entre su nido y las fuentes de alimento.

Las hormigas son insectos sociales que viven en colonias, aunque tienen habilidades individuales limitadas, logran comportamientos complejos a nivel colectivo. Por ejemplo, son capaces de "construir" las rutas más cortas entre el hormiguero y la comida por medio de una sustancia química que llamada *feromona*. Cuando una hormiga encuentra comida, en su recorrido de vuelta al nido deposita feromona en el medio, que todas pueden oler. Este rastro permite a las hormigas ir definiendo un camino a su hormiguero desde la comida. Si una hormiga encuentra un "atajo", las demás irán siguiendo el nuevo rastro y reforzándolo con mayor frecuencia, de modo que, al poco tiempo, éste tendrá una mayor concentración de feromona y el anterior (más largo) irá "desapareciendo", ya que la sustancia se irá evaporando.

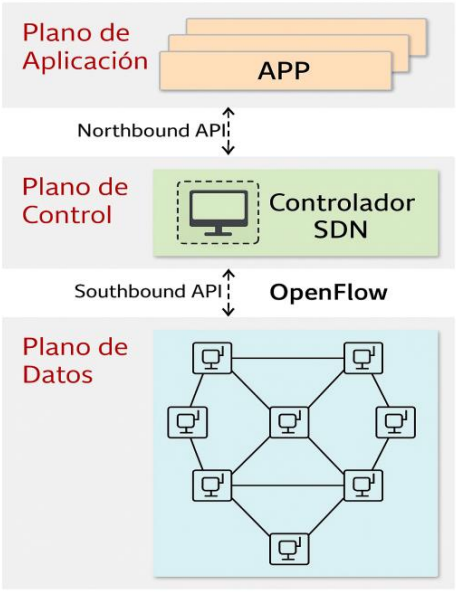


Fig. 1. Arquitectura general de una SDN

El algoritmo se basa en modelar este comportamiento con agentes artificiales (hormigas) que se mueven en un grafo, buscando caminos que minimicen un coste definido. Su movimiento está definido por una regla (*Regla de Transición de Estados*) que determina en cada nodo, cuál sería el siguiente al que pasar en base a la concentración de feromona que haya en el arco que los une y también en base a valores heurísticos que determinan lo deseable que es cada arco, según el problema que se esté resolviendo. Dicha regla combina ambos valores y asigna una probabilidad a cada nodo, que será elegido de manera aleatoria en base a dicha probabilidad (es un algoritmo pseudo-estocástico).

Tanto la feromona como la heurística se modelan con matrices de valores asociados a los arcos, de modo que la feromona se irá modificando según la calidad de la solución que encuentren las hormigas que pasen por dicho arco. Es decir, los arcos de las buenas soluciones se reforzarán positivamente ya que serán más deseables, mientras que los demás arcos verán reducida su feromona por una evaporación simulada.

ACO se ha aplicado a innumerables problemas de optimización, siempre que se puedan modelar como un grafo con pesos en sus arcos. Además, es el algoritmo más utilizado para resolver problemas complejos de *pathfinding* (búsqueda de camino óptimo), en grafos de dimensiones muy grandes o cuando haya que considerar restricciones u objetivos de optimización que no puedan manejar los métodos tradicionales.

Existen muchas variantes del algoritmo, pero los dos modelos fundamentales son: Ant System y Ant Colony System [8].

III. Definición del Problema

Formalmente, el problema se define así: dado un grafo dirigido $G=(V,E)$, donde V representa el conjunto de nodos de la red (switches SDN) y E representa el conjunto de enlaces entre nodos, cada enlace e tiene asociadas las siguientes métricas:

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

- *c(e)*: carga del enlace. Cantidad de ancho de banda utilizado.
- *d(e)*: retardo estimado en milisegundos. Tiempo que tarda un paquete en viajar de un origen a un destino en la red.
- *l(e)*: porcentaje de pérdida de paquetes. Porcentaje de paquetes que no llegan a su destino.

El objetivo es determinar el camino óptimo $P=\{a \rightarrow b\}$ desde un nodo origen $a \in V$ a un nodo destino $b \in V$, que minimice un coste basado en las métricas anteriores.

En consecuencia, el algoritmo a diseñar debe ser capaz de:

- Adaptarse a cambios dinámicos en el estado de la red en tiempo real.
- Tomar decisiones de encaminamiento que distribuyan eficientemente la carga de tráfico.
- Minimizar simultáneamente el retardo, la carga y la pérdida de paquetes.
- Ser implementable y ejecutable sobre un controlador SDN real (como *Ryu*).

La relevancia de este problema radica en que la eficiencia del balanceo de carga no solo afecta al rendimiento de la red, sino que impacta directamente en el uso racional de los recursos y en la Calidad de Experiencia del Usuario (QoE). Su impacto es aún más significativo en contextos como centros de datos, redes móviles o infraestructuras en la nube [4].

Este trabajo busca ofrecer una solución realista, flexible y eficaz, evaluada en un entorno de simulación práctico, lo que permite una comparación directa con enfoques tradicionales y pondrá de manifiesto las ventajas de las técnicas bioinspiradas en entornos SDN [10].

IV. Estado del Arte

El balanceo de carga en Redes Definidas por Software (SDN) es un campo de investigación activo debido a la necesidad de gestionar eficientemente los recursos en redes complejas y dinámicas. Los algoritmos bioinspirados, particularmente Ant Colony Optimization (ACO) y sus variantes, se han consolidado como una línea de solución muy prometedora.

Por ejemplo, el trabajo de Lin y Zhang [15] propone un enfoque jerárquico para el balanceo de carga en SDN, combinando una arquitectura de controladores distribuidos con optimización basada en ACO. La arquitectura incluye un controlador central que monitoriza la red y coordina sub-controladores, cada uno gestionando una subred homogénea. La innovación principal es la clasificación de tareas por demanda de CPU antes de la asignación a subredes especializadas, donde ACO selecciona el nodo específico para cada tarea basándose en el estado actual de la carga de los nodos.

En nuestra propuesta también se aplica ACO para hacer balanceo de carga, pero este estudio ofrece una solución integrada y dinámica que actúa en tiempo real sobre la topología de la red sin necesidad de clasificaciones previas ni la consideración de subredes homogéneas. Además, nuestro objetivo es la elección de rutas óptimas considerando métricas sobre la red (retardo, carga y pérdida de paquetes), obtenidas por monitorización activa, lo que permite una adaptación más flexible en redes cambiantes.

Otro estudio relacionado sería el presentado en [16]. En dicho trabajo, los autores proponen una solución para el balanceo de carga dinámico en SDN, combinando la selección dinámica de rutas y servidores mediante ACS (Ant Colony System). El objetivo es doble: optimizar simultáneamente la selección del servidor con menos carga y la elección de la mejor ruta del cliente al servidor, aprovechando la visión global que ofrece SDN. En este caso, se utiliza el controlador *OpenDaylight* para monitorizar en tiempo real las métricas de la red y las estadísticas de carga de CPU de los servidores. Con esta información, selecciona el servidor menos cargado y ACS encuentra la ruta óptima considerando el estado dinámico de la red y las feromonas depositadas.

En la comparativa con la presente propuesta, ambos trabajos se centran en la optimización del balanceo de carga en SDN, utilizando la variante ACS para la optimización de rutas y considerando la carga como una métrica fundamental para la elección de los nodos a atravesar. No obstante, el algoritmo propuesto en el artículo está diseñado específicamente para escenarios cliente-servidor y la monitorización directa de la CPU de los servidores, lo que limita su aplicabilidad a otras configuraciones más generales. En contraste, esta propuesta aplica ACS para un balanceo de carga más genérico de múltiples nodos y rutas, haciéndolo más adaptable. Además, en nuestro caso, aparte de la carga de los nodos, también se tienen en cuenta más factores sobre el estado de la red (pérdida de paquetes y retardo), lo que supone un valor añadido.

El artículo [17] describe un enfoque híbrido para el balanceo de carga en redes SDN, combinando ACO y Algoritmos Genéticos (GA), denominado G-ACO. Su objetivo es superar las limitaciones de cada estrategia por separado, aprovechando la exploración adaptativa de ACO y la diversidad genética de los algoritmos evolutivos. El algoritmo G-ACO se desarrolla en dos fases: una fase de inicialización (con un ACO clásico) donde las hormigas generan rutas y depositan feromonas considerando la carga y el retardo. Después se realiza una fase de optimización (aplicando GA) donde las rutas generadas por ACO se codifican como cromosomas, aplicando selección (priorizando rutas de menor coste), cruce (combinando rutas exitosas) y mutación (introduciendo variantes aleatorias) para evitar óptimos locales.

Nuestra propuesta resuelve el mismo problema y usa ACO como algoritmo central, sin embargo, la combinación de ACO y GA incrementa considerablemente la complejidad computacional espacial y temporal, dificultando su implantación en controladores SDN en tiempo real o entornos de red limitados. Además, G-ACO fue evaluado solo en simulaciones, mientras que la solución del presente estudio se ha implementado sobre *Mininet* y el controlador *Ryu*, acercándose más a escenarios reales.

Por último, el trabajo en el que se basa nuestra propuesta [11], presenta el algoritmo LLBACO (*Load and Link-based Ant Colony Optimization*), una estrategia de optimización bioinspirada para el balanceo de carga en redes definidas por software, que implementa un ACO simple, basado en el enfoque del *Ant System* (AS), que considera las métricas carga, retardo y tasa de pérdida de paquetes como indicadores del estado de la red (nodos y enlaces). Como se ha mencionado anteriormente, la diferencia principal de nuestro enfoque es la

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

implementación de una variante más eficiente de ACO, el *Ant Colony System* (ACS).

Además de lo comentado en cada caso, en este trabajo se presenta una herramienta que permite monitorizar el desempeño del algoritmo ACO propuesto sobre cada escenario y hacer cambios en tiempo real sobre aspectos de los mismos, lo que también supone una contribución novedosa al estado del arte.

V. Algoritmo Propuesto

El algoritmo desarrollado en este estudio se basa en la propuesta de LLBACO [11], el cual se ha mejorado integrando el enfoque ACS, que introduce mecanismos adicionales como la actualización local de feromona y decisiones más controladas mediante una estrategia de selección en parte *greedy*. Estas modificaciones buscan mejorar tanto la convergencia como la estabilidad del algoritmo en condiciones dinámicas de red.

Como paso previo a las mejoras con ACS, se implementó el algoritmo LLBACO tal como fue descrito originalmente, cuyas fases clave son:

1) **Selección del nodo siguiente** (en la construcción de cada solución) mediante la regla de transición de estados basada en una ruleta de probabilidad, la cual se denomina *Pij*. Es decir, la probabilidad de que una hormiga en el nodo *'i'* seleccione el nodo *'j'* como siguiente viene dada por la expresión de la Ecuación 1.

$$P_{ij} = \frac{\tau_{ij}^{\alpha} \cdot \eta_{ij}^{\beta} \cdot \mu_{ij}^{\gamma}}{\sum_{u \in \text{allowedSet}} \tau_{iu}^{\alpha} \cdot \eta_{iu}^{\beta} \cdot \mu_{iu}^{\gamma}} \quad \text{if } j \in \text{allowedSet} \quad (1)$$

Donde los pesos α, β y γ representan la influencia relativa de cada factor, *allowedSet* son aquellos nodos que las hormigas pueden escoger.

τ_{ij} representa la intensidad de la feromona en el enlace. $\eta_{ij} = 1/Load_{ij}$ representa la influencia de la carga del enlace. $\mu_{ij} = 1/Cost_{ij}$ representa la influencia del coste del enlace (ver Ecuación 4). Ante situaciones con carga y coste mínimos (=0), se considerarían los valores máximos posibles para η y μ.

2) **Cálculo de Carga y Coste de un camino**. La carga de un camino solución se define como la carga máxima de cualquiera de los enlaces que componen ese camino (ver Ecuación 2):

$$\text{Load}(path) = \max_{(i,j) \in path} \text{Load}_{ij} \quad (2)$$

El coste de un camino se define como la suma de los costes de todos los enlaces que componen ese camino (ver Ecuación 3):

$$\text{Cost}(path) = \sum_{(i,j) \in path} \text{Cost}_{ij} \quad (3)$$

Calculando cada coste en un enlace *(i,j)* como una combinación del retardo del enlace (*Delay*) y la Tasa de pérdida de paquetes de dicho enlace (*PacketLoss*), como muestra la Ecuación 4:

$$\text{Cost}_{ij} = \text{Delay}_{ij} + \text{PacketLoss}_{ij} \quad (4)$$

Estos valores se normalizan (min-max en [0-1]) para que su influencia sea similar, aunque se podría aplicar un factor de ponderación a cada uno, si se desea.

3) **Actualización de feromona**. Se trata de un proceso combinado de evaporación y depósito de feromona. En LLBACO el método de actualización de feromona es global. Esto significa que, a diferencia de otros métodos donde cada hormiga deposita la feromona al completar su viaje (llegar al nodo destino), en este caso solo los enlaces del mejor camino encontrado hasta el momento por todas las hormigas reciben un aporte de feromona. Esto ocurre después de que todas las hormigas hayan terminado su búsqueda en una iteración. La actualización de feromona se define a través de la Ecuación 5:

$$\tau_{ij} \leftarrow (1 - \rho) \cdot \tau_{ij} + \Delta\tau_{ij} \quad (5)$$

Siendo:

$$\Delta\tau_{ij} = \begin{cases} \frac{Q}{L_{\text{best}}}, & \text{si } (i,j) \in \text{mejor camino} \\ 0, & \text{en otro caso} \end{cases} \quad (6)$$

Donde $\Delta\tau_{ij}$ es la cantidad de feromona depositada por una hormiga en el enlace entre el nodo *i* y *j*.

$\rho \in (0, 1)$ representa la velocidad de evaporación, Q es la cantidad total de feromona y L el valor calculado como coste del mejor camino.

La feromona es el medio por el cual las hormigas se comunican entre ellas, afectando directamente a la eficiencia de cada hormiga para buscar los caminos en el grafo del problema.

4) **Elección del mejor camino**. El algoritmo pretende buscar los caminos entre nodos de la red que minimicen los criterios elegidos de carga, retardo y pérdida de paquetes, consiguiendo un balanceo de la carga en los enlaces de la red, con vista a asegurar y mantener la calidad de servicio (QoS).

En una red compleja, un camino con poca carga es posible que tenga un alto retardo y pérdida de paquetes. De modo que, para encontrar el mejor camino, el algoritmo establece un umbral dinámico que excluye esos caminos, en los que la carga excede dicho umbral. Una vez filtrados, el camino con coste mínimo es elegido de entre los restantes.

El **umbral** consiste en la carga promedio de todos los caminos encontrados en una iteración dada y se calcula de forma dinámica con la expresión que muestra la Ecuación 7.

$$L_k = \frac{1}{|pathlist_k|} \sum_{path \in pathlist_k} \text{Load}(path) \quad (7)$$

Donde *pathlist* es el conjunto de caminos encontrados por las hormigas en la k-ésima iteración.

Una vez que hemos calculado el umbral (*Lk*), actualizamos la lista de caminos, eliminando de la lista

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

original aquellos caminos cuya carga (*Load(path)*) sea mayor que dicho umbral.

Finalmente, de la lista de caminos filtrados, se selecciona el mejor camino global para esta iteración.

El mejor camino es definido como el camino que tiene el mínimo coste entre todos los caminos que han sobrevivido al filtro de carga. La expresión que muestra esto se ve en la Ecuación 8:

$$\text{best_path} = \arg \min_{path \in filteredPaths} \text{Cost}(path) \quad (8)$$

Sobre este funcionamiento, las mejoras que implementa el enfoque Ant Colony System serían:

- **Parámetro de exploración/explotación (q_0)**:
  En cada paso, una hormiga decide si selecciona el siguiente nodo al que moverse en la construcción de su camino solución de forma determinista (enfoque *greedy*) o probabilística ("ruleta de probabilidad"), controlado por el parámetro q_0. Para ello generamos un número aleatorio q entre 0 y 1:
  - Si ($q < q_0$), se aplica la estrategia de explotación (greedy), seleccionando el nodo que maximiza la expresión:

$$\text{argmax}_{j \in allowedSet} \left(\tau_{ij}^{\alpha} \cdot \eta_{ij}^{\beta} \cdot \mu_{ij}^{\gamma} \right) \quad (9)$$

  - En caso contrario será aplicada una estrategia de exploración basada en la probabilidad, donde el siguiente nodo es seleccionado utilizando el método de la ruleta de probabilidades (calculadas con la expresión de la Ecuación 1), permitiendo encontrar nuevos caminos.
- **Actualización Local de Feromonas (ϕ)**:
  Cada vez que una hormiga recorre un enlace durante su exploración, va aplicando una evaporación y pequeño aporte de feromona sobre ese enlace, aplicando una actualización local.

  Esto permite reducir la atracción por caminos que ya han sido explorados, fomentando una mayor diversidad en la búsqueda. La expresión de la actualización se muestra en la Ecuación 10:

$$\tau_{ij} \leftarrow (1 - \phi) \cdot \tau_{ij} + \phi \cdot \tau_0 \quad (10)$$

  Donde ϕ es la tasa de evaporación local ($\phi \in (0, 1)$) y τ_0 es el valor inicial de la feromona.

  Esta actualización se aplica en tiempo real a medida que la hormiga va recorriendo nodos, a diferencia de la actualización global del AS, que se realiza una vez finalizada la iteración y solamente sobre los arcos del mejor camino.

ACS ha demostrado en múltiples problemas y estudios ser una evolución muy eficiente del AS clásico, ya que permite hacer una búsqueda adaptativa sobre el espacio de soluciones, fomentando la exploración (búsqueda de nuevas soluciones) en unas fases y la explotación (mejora o refinamiento de soluciones ya encontradas) en otras. Esto es crucialmente útil en entornos dinámicos, como es nuestro caso de las redes SDN, donde las condiciones cambian continuamente, siendo necesaria la adaptación constante del enrutamiento.

Este algoritmo se ha incluido dentro de la lógica del controlador Ryu para la SDN simulada.

VI. Herramienta de Visualización

Otra de las contribuciones de este trabajo es una utilidad/aplicación web desarrollada con el propósito de facilitar la interacción del usuario con la SDN para la ejecución y monitorización del ACO que se ejecuta sobre la misma. Por tanto, ésta permite la visualización en tiempo real de su topología y las métricas asociadas a cada enlace, así como la evolución del estado de la red a medida que el algoritmo de optimización propone nuevas soluciones.

De este modo, el usuario puede observar gráficamente la evolución progresiva de la carga, los costes y los caminos seleccionados por el algoritmo en distintos instantes (y estados) de la red.

Su implementación se ha realizado con *Flask* (https://flask.palletsprojects.com/en/stable/api/), servidor y backend de la aplicación; *Socket.IO* (https://socket.io/), comunicación entre navegador web y servidor Flask; *HTTP/SWGI* (interfaz entre controlador Ryu y Flask) y *D3.js* (https://d3js.org/), visualización de la topología.

La visualización de la red se actualiza automáticamente cada 5 segundos gracias a un mecanismo de *snapshots* periódicos. Estas snapshots son capturas del estado actual de la red, generadas por el monitor de Ryu, y contienen información relevante como la carga, el retardo y la pérdida de paquetes en cada enlace. Posteriormente, estos datos se envían al servidor Flask, que los transmite al navegador del usuario a través de Socket.IO, actualizando así la representación gráfica en tiempo real.

La Figura 2 muestra la ventana de interacción de la herramienta para el usuario.

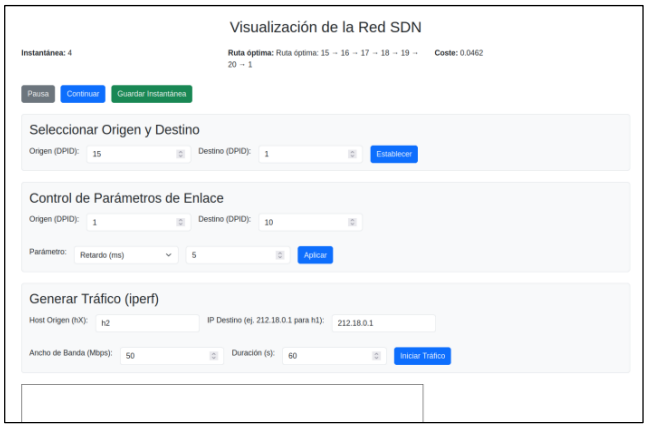


Fig. 2. Herramienta de visualización de la SDN y el ACO, así como la gestión del estado de la red en cada instante.

Como se puede ver, es posible moverse entre instantáneas (estados de la red), lo cual es muy interesante para estudiar cómo se comportan los algoritmos en cada instante. También, se permite elegir los nodos origen y destino para que el ACO realice la búsqueda. De la misma forma, también es posible cambiar de manera manual los valores de las métricas en cada enlace: retardo (en milisegundos), pérdida de paquetes (porcentaje entre 0 y 100), carga (ancho de banda disponible). En relación a la carga, se incluye también una utilidad para

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

generar tráfico sintético entre dos nodos cualesquiera de la red, definiendo el ancho de banda a ocupar y la duración de dicho flujo de tráfico. Esto se hará con *iperf* (https://iperf.fr/).

Como se puede ver en la Figura 3, La herramienta mostrará en todo momento la topología de la SDN, además del camino encontrado por el algoritmo que estemos monitorizando marcado sobre la misma, así como información de cada enlace elegido.

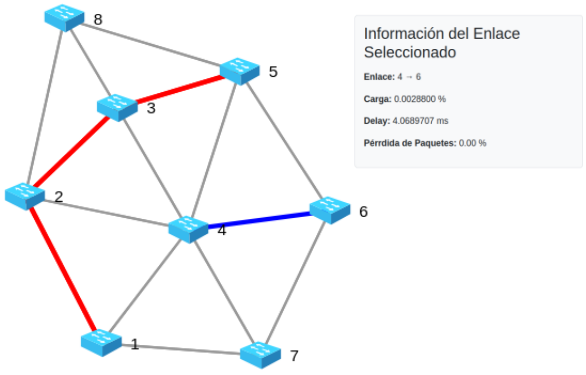


Fig. 3. Herramienta de visualización de la SDN y el funcionamiento del algoritmo ACO. Mejor solución encontrada hasta el momento para un camino (en rojo) y datos de un enlace seleccionado (el azul).

Tanto la herramienta, como el código de los algoritmos (en Python) y del controlador se pueden encontrar en el repositorio: https://github.com/lmchaves/tfg.

VII. Experimentos y Resultados

Para validar el funcionamiento y el rendimiento del algoritmo implementado (***LLBACS***) y compararlo con el *LLBACO* original, así como otros métodos clásicos (Dijkstra), se han realizado una serie de experimentos. En ellos, se analizan diferentes configuraciones, modificando los parámetros del algoritmo para evaluar su impacto sobre el coste de las rutas, así como la carga de la red y el retardo.

Para evaluar los algoritmos, se han definido distintas configuraciones o variaciones de parámetros, que se enumerarán a continuación. Cada una de ellas se ha ejecutado 30 veces (debido al componente estocástico del ACO y ACS) sobre la instancia del problema mostrada en la Figura 4, en la que se tienen 20 nodos sin conectividad total, que modela una red con cierta complejidad.

Hay que aclarar que este trabajo se centra en el estudio de la influencia de los parámetros de los algoritmos, más que en la prueba en instancias de diversa topología para garantizar que se dan situaciones diferentes y, quizá, más realistas. Del mismo modo, la instancia se ha configurado con unos valores "estándar" de: ancho de banda = 100 Mbps, retardo = 10 ms y pérdida de paquetes = 5% en todos los enlaces.

La Tabla I muestra la configuración base de los algoritmos LLBACO y LLBACS (es el que considera los últimos 4 parámetros). En la que se usará un tamaño de colonia estándar para 20 nodos (20 hormigas), se le dará el mismo peso a feromona (α) y las dos funciones heurísticas (β y γ), una tasa de evaporación global (ρ) media, un impacto igualitario de la exploración y explotación en el ACS (q_0) y una tasa de evaporación local (ϕ) en el ACS media.

Tabla I.
Tabla de configuración básica de los algoritmos ACO y ACS

Parámetro	Valor
Iteraciones	200
Tamaño de la colonia	20
α	1.0
β	1.0
γ	1.0
ρ	0.5
Q	1.0
q_0	0.5
ϕ	0.5

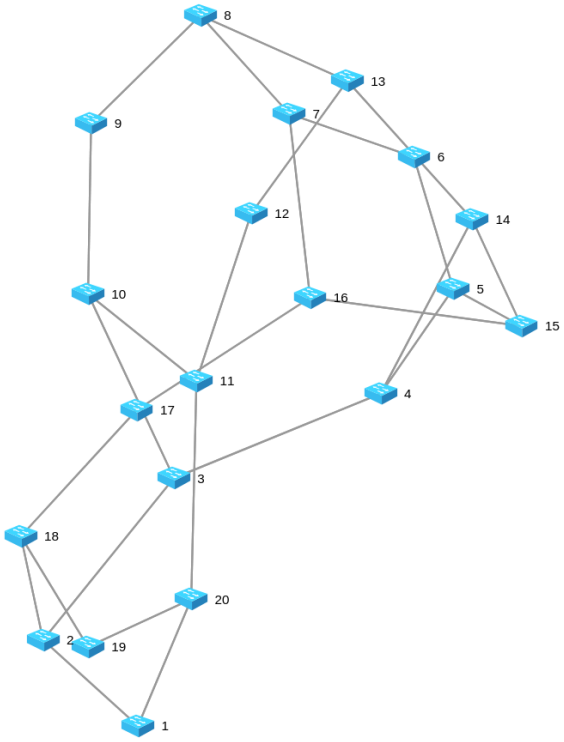


Fig. 4. Topología de la red considerada para los experimentos.

A partir de esta configuración básica, se han diseñado experimentos con variaciones en los parámetros de ACS, a fin de valorar su impacto e influencia en los resultados obtenidos, tanto en los caminos encontrados, como en el estado posterior de la red (buscando el deseado balanceo). Para ello, se han considerado valores extremos en dichos parámetros. Dichas configuraciones se pueden ver en la Tabla II.

Tabla II.
Tabla de configuraciones a probar sobre ACS

#	Configuración	Parámetro cambiado
1	Configuración base	—
2	Menos iteraciones	Iteraciones = 20
3	Más iteraciones	Iteraciones = 400
4	Colonia más pequeña	Tamaño colonia = 5
5	Colonia más grande	Tamaño colonia = 100
6	Alta influencia de feromona	$\alpha = 3.0$
7	Alta influencia de coste	$\beta = 3.0$
8	Alta influencia de la carga	$\gamma = 3.0$
9	Evaporación global mayor	$\rho = 0.95$
10	Evaporación global menor	$\rho = 0.05$
11	Menor exploración greedy	$q_0 = 0.9$
12	Mayor exploración greedy	$q_0 = 0.1$
13	Evaporación local mayor	$\phi = 0.95$
14	Evaporación local menor	$\phi = 0.05$

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

La Figura 5 muestra los resultados referidos a los costes promedio de los mejores caminos obtenidos por el algoritmo LLBACS con cada configuración.

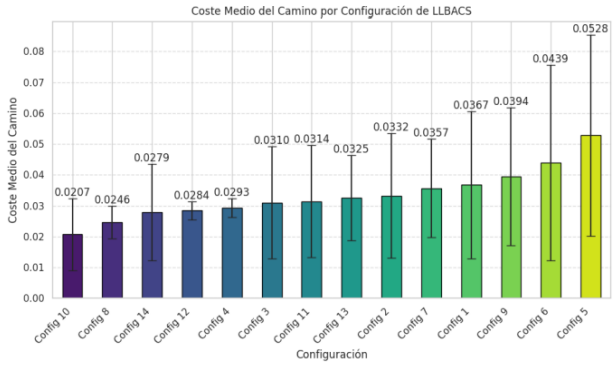


Fig. 5. Costes promedio de los 30 mejores caminos encontrados para cada configuración para el algoritmo ACS. La línea sobre cada barra indica la desviación típica.

Como se puede ver, la configuración que consigue los mejores resultados (menor coste) es la 10, en la que se reduce la evaporación de la feromona global al mínimo, lo que implica que la feromona aportada se mantiene mucho más tiempo, por lo que los enlaces de los mejores caminos persisten más tiempo y se consigue una explotación de las soluciones. La configuración 14 va en la misma línea, tratándose de un parámetro que sólo se utiliza en ACS (no en AS), al igual que la configuración 12, que fomenta la búsqueda *greedy* en ACS. En la comparativa entre variaciones del mismo parámetro, vemos que el peso de la carga de los enlaces en el cálculo de las probabilidades afecta para obtener mejores resultados que los pesos de la feromona o del propio coste.

En cuanto al tiempo de computación, éste varía desde los 0,09 seg de la configuración 2 (la de menos iteraciones), a los 0,88 seg de la configuración base y los 4,36 seg de la configuración 5, que es la que más hormigas utiliza. Los experimentos se han ejecutado en un PC con Procesador Intel Core i7-1195G7 2.90GHz, 16GB de RAM y SSD de 1TB.

En la comparativa entre las soluciones obtenidas por los diferentes algoritmos (ver Tabla III), podemos observar que LLBACS consigue mejorar en las tres métricas los resultados obtenidos por LLBACO y Dijkstra.

Tabla III
RESULTADOS DE LOS DIFERENTES ALGORITMOS SOBRE LA TOPOLOGÍA DE LA FIGURA 4. RESULTADOS MEDIOS DE 30 EJECUCIONES.

Métrica	LLBACO	LLBACS	Dijkstra
Coste Total	12345.67	11340.89	13200.54
Carga Total	9876.54	9340.21	10100.77
Retardo Total	456.78	420.56	510.12

Si comparamos las soluciones obtenidas en 30 ejecuciones sobre un instante concreto del estado de la red (*snapshot* 5), podemos ver en la Figura 6, que LLBACS minimiza en la mayoría de las ejecuciones dicho coste.

La Tabla IV muestra una comparativa entre los algoritmos ACO con resultados numéricos, obtenidos en 30 ejecuciones sobre un instante concreto del estado de la red (*snapshot* 5).

Como se puede ver en la tabla, LLBACS obtiene mejores resultados medios en todas las métricas.

Los tiempos promedio para obtener las soluciones en esa *snapshot* fueron 2,6 seg LLBACO y 3,5 seg LLBACS.

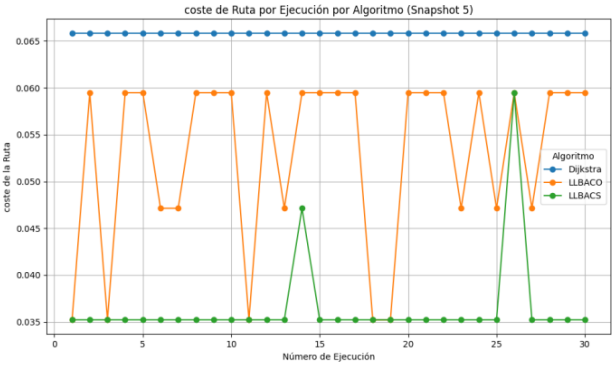


Fig. 6. Mejores costes de los caminos encontrados por cada algoritmo en una *snapshot* (instante particular) de la red en 30 ejecuciones.

Tabla IV
COMPARATIVA DE RESULTADOS MEDIOS PARA TRES MÉTRICAS. 30 EJECUCIONES DE LOS ALGORITMOS SOBRE LA SNAPSHOT 5 DE LA RED.

Métrica	LLBACO	LLBACS
Coste	0.052	0.037
Carga	0.000225	0.000130
Retardo	0.115	0.058

VIII. CONCLUSIONES

En este trabajo hemos presentado dos propuestas de algoritmo basado en colonias de hormigas (ACO), adaptados para buscar caminos óptimos en el despliegue de una Red Definida por Software (SDN), manteniendo un balanceo de carga en los enlaces de la misma.

El primero de los algoritmos ha sido una adaptación del que presentaron los autores en [11], el cual era una implementación de la variante llamada Ant System (AS), el ACO más simple. El segundo algoritmo que hemos desarrollado es la variante Ant Colony System (ACS), que incorpora mecanismos para obtener mejores soluciones con un mejor compromiso entre exploración y explotación.

Ambos algoritmos han sido implementados dentro de la "lógica" del controlador de la SDN, a fin de facilitar la gestión de la red física en base a los criterios que indiquen las soluciones obtenidas por los ACO.

Se han realizado varios experimentos para evaluar el impacto de diferentes parametrizaciones, así como realizar una comparativa entre la calidad de las soluciones obtenidas por cada uno de ellos frente a un método clásico (Dijkstra). Los resultados de la variante ACS han sido bastante mejores que los del AS, si bien su tiempo de cómputo es ligeramente superior.

En próximos trabajos ampliaremos las comparativas, considerando escenarios más variados y realistas. También, se evaluará el método propuesto sobre topologías con mayor dinamismo, ya que se han hecho sobre instancias de la red en las que el único tráfico que había era el de los propios algoritmos en su funcionamiento. Dado que los ACO tienen entre sus virtudes adaptarse a problemas dinámicos, esperamos obtener buenos resultados con ellos. Dentro de esa propuesta, se contemplará también la influencia de distintos flujos de tráfico simultáneos y se valorará la adaptación de los ACO en tiempo real. También se trabajará en definir funciones de coste más adecuadas, tal y como la carga media de los

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

enlaces en lugar de la máxima, que podría estar condicionada por un cuello de botella puntual.

AGRADECIMIENTOS

Este trabajo ha sido financiado, en parte, por el Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación (MICIU/AEI/10.13039/501100011033) a través del proyecto con código PID2023-151462OB-I00 y los proyectos PID2023-147409NB-C21 y PID2020-115570GB-C22. También se ha financiado mediante los proyectos TED2021-131699B-I00 y TED2021-129938B-I00 del MICIU y la Unión Europea NextGenerationEU/PRTR. Además del proyecto C-ING-179-UGR23 concedido por la "Consejería de Universidades, Investigación e Innovación" (Junta de Andalucía, Programa FEDER 2021-2027).

REFERENCIAS

[1] Kamal Benzekki, Abdeslam El Fergougui, Abdelbaki Elbelrhiti Elalaoui. "Software-defined networking (SDN): a survey". Security and communication networks, 2016, vol. 9, no 18, p. 5803-5833.

[2] Amir Sina Milani and Nima Jafari Navimipour. "Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends". Journal of Network and Computer Applications, 71:86–98, 2016.

[3] Pashupati Baniya, Parma Nand, and Bharat Bhushan. Dynamic load balancing schemes for software-defined networking (SDN). In Proceedings of Second International Conference on Intelligent System. Springer, 2024.

[4] Qin Du, Xin Cui, Haoyao Tang, and Xiangxiao Chen. "Review of load balancing mechanisms in SDN-based data centers". https://www.scirp.org/journal/paperinformation?paperid=130469, 2024.

[5] J.R. Jiang, H.W. Huang, J.H. Liao, S.Y. Chen, S. Y. "Extending Dijkstra's shortest path algorithm for software defined networking". In *The 16th Asia-Pacific Network Operations and Management Symposium* (pp. 1-4). IEEE. 2014

[6] Hidayat, T., Azzery, Y., & Mahardiko, R. (2019). "Load balancing network by using round Robin algorithm: a systematic literature review". Jurnal Online Informatika, 4(2), 85-89.

[7] Bruno A. A. Nunes, Marc Mendonca, Xuan-Nam Nguyen, Katia Obraczka, and Thierry Turletti. "A survey of software-defined networking: Past, present, and future of programmable networks". IEEE Communications Surveys & Tutorials, 16(1), 2014.

[8] Ant Colony Optimization by Marco Dorigo and Thomas Stützle, MIT Press, 2004. ISBN 0-262-04219-3

[9] Lewis, Ted G. (June 12, 2012). "Cognitive stigmergy: A study of emergence in small-group social networks". Cognitive Systems Research. 21: 7–21.

[10] Ali Akbar Neghab, Nima Jafari Navimipour, Mehdi Hosseinzadeh, Ali Rezaee. "Nature-inspired meta-heuristic algorithms for solving the load balancing problem in the software-defined network". International Journal of Communication Systems, 31(18), 2018.

[11] Chunzhi Wang, Gang Zhang, Hui Xu, and Hongwei Chen. "An ACO-based link load-balancing algorithm in SDN". In 2016 7th International Conference on Cloud Computing and Big Data (CCBD). IEEE, 2016.

[12] Abdelaziz A., Fong A. T., Gani A., et al. "Distributed controller clustering in software defined networks". PloS One, 12(4): e0174715, 2017.

[13] Murat Karakus and Arjan Durresi. "A survey: Control plane scalability issues and approaches in software-defined networking (SDN)". Computer Networks, 112:279–293, 2017.

[14] Christian Blum. "Ant colony optimization: Introduction and recent trends". Physics of Life Reviews, 2(4):353–373, 2005.

[15] WuCai Lin and LiChen Zhang. "The load balancing research of SDN based on ant colony algorithm with job classification". 2016 International Conference on Smart Grid and Electrical Automation (ICSGEA), pages 1–4, 2016.

[16] Sushma Sathyanarayana and Melody Moh. "Joint route-server load balancing in software defined networks using ant colony optimization". 2017 International Conference on Computing, Networking and Communications (ICNC), pages 338–342, 2017.

[17] Hai Xue, Kyung Tae Kim, and Hee Yong Youn. "Dynamic load balancing of software-defined networking based on genetic-ant colony optimization". Sensors, 19(2):311, 2019.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

UNIVERSIDAD DE EXTREMADURA