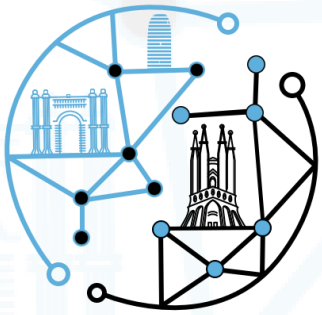


XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle -- Universitat Ramon Llull

Barcelona 2023

LIBRO DE ACTAS

CaixaBank

SCITEL
Sociedad Científica de Ingeniería Telemática

laSalle
UNIVERSIDAD RAMON LLULL

ISBN: 978-84-09-58148-1

Editores:
Guiomar Corral Torruella
Julia Sánchez Rodríguez

XVI Jornadas de Ingeniería Telemática – JITEL 2023

JITEL 2023

Presentación

Al mirar hacia atrás en el tiempo, es fascinante reflexionar sobre el estado de las tecnologías relacionadas con las comunicaciones en redes e Internet hace 25 años. En ese entonces, estábamos en los albores de una era digital incipiente, y la rápida evolución ha transformado radicalmente el panorama de la Telemática. Desde las conexiones dial-up hasta la actualidad del 5G (y pensando ya en el 6G), hemos pasado por la era de Internet, la explosión de las redes sociales, el auge de las tecnologías 4G LTE, y la proliferación de "cosas" conectadas dando lugar al Internet de las cosas (IoT). El despliegue de tecnologías como Wi-Fi, la evolución de protocolos como TCP/IP, y la aparición de servicios en la nube y big data han modificado de manera sustancial la forma en que nos conectamos y compartimos información. En estas jornadas conmemorativas, hemos reflexionado sobre toda esta evolución y compartido el papel esencial de la labor docente e investigadora en este progreso continuo.

Desde hace 25+1 años que se celebró la primera, las Jornadas de Ingeniería Telemática (JITEL) han establecido un foro consolidado para el intercambio de conocimientos entre investigadores. Un espacio de reunión, debate y divulgación que permite a los grupos dedicados a la enseñanza e investigación en redes y servicios telemáticos compartir experiencias y resultados.

Es importante reconocer la importancia de la labor docente en este ámbito. Los avances tecnológicos no solo surgen de la investigación, sino también del compromiso de los educadores que comparten su conocimiento y experiencia con las generaciones futuras. En estas jornadas, celebramos también la contribución significativa de la comunidad académica, que impulsa el desarrollo y la innovación en el campo de la Ingeniería Telemática.

Esta dieciseisava edición de las Jornadas (JITEL 2023), celebrada en Barcelona los días 8, 9 y 10 de noviembre de 2023, ha sido organizada por el área de Ingeniería Telemática de La Salle - Universitat Ramon Llull y la Sociedad Científica de Ingeniería Telemática (SCITEL), con el apoyo y patrocinio de Caixabank.

En este libro de actas se recogen las 67 contribuciones aceptadas en las jornadas cuyas sesiones se organizan en los siguientes ámbitos temáticos: *Análisis de datos de redes (machine learning), cloud, edge and fog computing; Aplicaciones y servicios; Calidad en comunicaciones, redes y sistemas (parámetros y percepción); Docencia en telemática; Inteligencia Artificial para redes; Investigación en tecnología educativa; Gestión y operación de redes y sistemas; Multimedia, salud y sociedad digitales; Redes dedicadas (IoT, m2m, e2e, redes de sensores, redes Ad-Hoc…); Redes de nueva generación (5G, 6G, IoT, slicing…); Seguridad en comunicaciones, redes y sistemas; Virtualización de redes y servicios (SDN/NFV, orquestación de recursos, slicing…); además de una sesión de posters gamificada 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 aquí inspiren nuevas ideas y colaboraciones que moldeen el camino de la Ingeniería Telemática en los años venideros.

Comité de Programa de JITEL 2023

JITEL 2023

Comité de Programa

Presidenta del Comité de Programa:

Corral Torruella, Guiomar (La Salle – Universitat Ramon Llull)

Miembros del Comité de Programa:

Agüero Calvo, Ramón (Universidad de Cantabria)
Bagnulo, Marcelo (Universidad Carlos III de Madrid)
Bote Lorenzo, Miguel Luis (Universidad de Valladolid)
Carmona Murillo, Javier (Universidad de Extremadura)
Carneiro Díaz, Victor (Universidade da Coruña)
Estepa Alonso, Antonio José (Universidad de Sevilla)
Felici Castell, Santiago (Universitat de Valencia)
Fernández Navajas, Julián (Universidad de Zaragoza)
García Fernández, Roberto (Universidad de Oviedo)
Hesselbach Serra, Xavier (Universitat Politècnica de Catalunya)
Higuero Aperribay, Maria Victoria (Euskal Herriko Unibertsitatea)
Lloret Mauri, Jaime (Universitat Politècnica de Valencia)
Manzanares López, Pilar (Universidad Politécnica de Cartagena)
Navarro Ortiz, Jorge (Universidad de Granada)
Ramis Bibiloni, Jaume (Universitat Illes Balears)
Rojas Sánchez, Elisa (Universidad de Alcalá)
Román Castro, Rodrigo (Universidad de Málaga)
Ruiz Martínez, Antonio (Universidad de Murcia)
Suárez Sarmiento, Álvaro (Universidad de Las Palmas de Gran Canaria)
Valero Duboy, Miguel Ángel (Universidad Politécnica de Madrid)

Comité de Organización

Presidenta del Comité de Organización:

Corral Torruella, Guiomar (La Salle – Universitat Ramon Llull)

Miembros del Comité de Organización:

Briones Delgado, Alan (La Salle – Universitat Ramon Llull)
Sánchez Rodríguez, Julia (La Salle – Universitat Ramon Llull)
Zaballos Diego, Agustín (La Salle – Universitat Ramon Llull)

Revisores

Agüero Calvo, Ramón (Universidad de Cantabria)
Alario-Hoyos, Carlos (Universidad Carlos III de Madrid)
Alcaraz, Cristina (Universidad de Málaga)
Alorda, Bartomeu (Universitat Illes Balears)
Álvarez Díaz, Manuel (Universidade da Coruña)
Álvarez, Marco (Universidade da Coruña)
Angel Irastorza, Jose (Universidad de Cantabria)
Atutxa, Asier (Euskal Herriko Unibertsitatea)
Bernardos, Carlos (Universidad Carlos III de Madrid)
Briones, Alan (La Salle – Universitat Ramon Llull)
Cabot, Miquel A. (Universitat Illes Balears)
Cacheda, Fidel (Universidade da Coruña)
Caleya, Julia (Universidad de Granada)
Calle-Cancho, Jesús (Universidad de Extremadura)
Calveras, Anna (Universitat Politècnica de Catalunya)
Carmona Murillo, Javier (Universidad de Extremadura)
Carneiro, Victor (Universidade da Coruña)
Chinchilla-Romero, Natalia (Universidad de Granada)
Corcoba Magaña, Víctor (Universidad de Oviedo)
Corral Torruella, Guiomar (La Salle – Universitat Ramon Llull)
Cortés-Polo, David (Universidad de Extremadura)
de la Cruz, Luis (Universitat Politècnica de Catalunya)
de la Oliva, Antonio (Universidad Carlos III de Madrid)
Delgado-Ferro, Félix (Universidad de Granada)
Diez, Luis (Universidad de Cantabria)
Estepa, Antonio (Universidad de Sevilla)
Estepa, Rafa (Universidad de Sevilla)
Felici Castell, Santiago (Universitat de Valencia)
Fernandez, Diego (Universidade da Coruña)
Fernández-Navajas, Julián (Universidad de Zaragoza)
Franco, David (Euskal Herriko Unibertsitatea)
Gamiz, Idoia (Euskal Herriko Unibertsitatea)
García Fernández, Roberto (Universidad de Oviedo)
García Gutiérrez, Alberto Eloy (Universidad de Cantabria)
García Moros, José (Universidad de Zaragoza)
García Pañeda, Xabiel (Universidad de Oviedo)
García Zarza, Pablo (Universidad de Valladolid)

García, Laura (Universitat Politècnica de Valencia)
García, Marta (Universidad de Cantabria)
Garcia-Carrillo, Dan (Universidad de Oviedo)
Garcia-Pineda, Miguel (Universitat de Valencia)
Genovard Oliver, Josep (Universitat Illes Balears)
Gómez Sánchez, Eduardo (Universidad de Valladolid)
González-Ortega, David (Universidad de Valladolid)
Hernandez, Angela (Universidad de Zaragoza)
Hernandez, Jose Alberto (Universidad Carlos III de Madrid)
Hesselbach Serra, Xavier (Universitat Politècnica de Catalunya)
Hinarejos, M. Francisca (Universitat Illes Balears)
Jacob, Eduardo (Euskal Herriko Unibertsitatea)
Jiménez, José (Universitat Politècnica de Valencia)
Lloret, Jaime (Universitat Politècnica de Valencia)
Lopez, Miguel (Universidad de Granada)
López-Vizcaíno, Manuel (Universidade da Coruña)
Macias Lopez, Elsa (Universidad de Las Palmas de Gran Canaria)
Malgosa-Sanahuja, Josemaría (Universidad Politécnica de Cartagena)
Manzanares López, Pilar (Universidad Politécnica de Cartagena)
Marrero, Domingo (Universidad de Las Palmas de Gran Canaria)
Martínez, Ignacio (Universidad de Zaragoza)
Mata, Jorge (Universitat Politècnica de Catalunya)
Mayor, Vicente (Universidad de Sevilla)
Melendi, David (Universidad de Oviedo)
Montagud, Mario (I2CAT Foundation)
Muñoz, Antonio (Universidad de Málaga)
Muñoz, Luis (Universidad de Cantabria)
Muñoz-Gea, Juan Pedro (Universidad Politécnica de Cartagena)
Mut, Macià (Universitat Illes Balears)
Navarro Ortiz, Jorge (Universidad de Granada)
Navarro, Joan (La Salle – Universitat Ramon Llull)
Novoa, Francisco J. (Universidade da Coruña)
Parra, Lorena (Universitat Politècnica de Valencia)
Payeras-Capellà, M. Magdalena (Universitat Illes Balears)
Pericàs Gornals, Rosa (Universitat Illes Balears)
Ramis Bibiloni, Jaume (Universitat Illes Balears)
Ramos-Munoz, Juan (Universidad de Granada)
Ríos, Rubén (Universidad de Málaga)
Rodríguez, Manuel (Universidad de Valladolid)
Rodríguez-Martín, Pablo (Universidad de Granada)

Rodríguez-Pérez, Francisco-Javier (Universidad de Extremadura)
Rojas Sánchez, Elisa (Universidad de Alcalá)
Roman Castro, Rodrigo (Universidad de Málaga)
Ruiz Martínez, Antonio (Universidad de Murcia)
Ruiz-Mas, Jose (Universidad de Zaragoza)
Salazar, Jose (Universidad de Zaragoza)
Sánchez, Julia (La Salle – Universitat Ramon Llull)
Sanchez, Luis (Universidad de Cantabria)
Sanchez-Fernandez, Luis (Universidad Carlos III de Madrid)
Sanchez-Iborra, Ramon (Universidad de Murcia)
Sanchez-Vital, Roger (Universitat Politècnica de Catalunya)
Sanz, Ane (Euskal Herriko Unibertsitatea)
Sasiain, Jorge (Euskal Herriko Unibertsitatea)
Segura Garcia, Jaume (Universitat de Valencia)
Sendra, Sandra (Universitat Politècnica de Valencia)
Simmross-Wattenberg, Federico (Universidad de Valladolid)
Soriano Asensi, Antonio (Universitat de Valencia)
Suarez, Alvaro (Universidad de Las Palmas de Gran Canaria)
Taha, Miran (Universitat Politecnica de Valencia)
Toledo, Nerea (Euskal Herriko Unibertsitatea)
Verdú, María Jesús (Universidad de Valladolid)
Zaballos, Agustín (La Salle – Universitat Ramon Llull)

Índice de trabajos

[SRS1] Seguridad en comunicaciones, redes y sistemas 1

[RNG] Redes de nueva generación (5G, 6G, IoT, slicing…)

[CCRS-IAN] Calidad en comunicaciones, redes y sistemas (parámetros y percepción) / Inteligencia Artificial para redes

[VRS] Virtualización de redes y servicios (SDN/NFV, orquestación de recursos, slicing…)

[SRS2-AS1] Seguridad en comunicaciones, redes y sistemas 2 / Aplicaciones y servicios 1

[AS2] Aplicaciones y servicios 2

[AS3] Aplicaciones y servicios 3

[DT-ITE] Docencia en telemática / Investigación en tecnología educativa

[ADR-GORS] Análisis de datos de redes (machine learning), cloud, edge and fog computing / Gestión y operación de redes y sistemas

[RD] Redes dedicadas (IoT, m2m, e2e, redes de sensores, redes Ad-Hoc…)

[MSSD] Multimedia, salud y sociedad digitales

Trabajos presentados en Póster

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Uso práctico del modelo ATT&CK para la detección de ciberataques

Elvira Castillo-Fernández*, J. Diaz-Verdejo*, R. Estepa Alonso†, A. Estepa Alonso†, Javier Muñoz-Calle†

* Dpt. Teoría de Señal, Telemática y Comunicaciones, CITIC, Univ. de Granada
C/ Periodista Daniel Saucedo Aranda, s/n, 18071 Granada (Spain)
E-mail: elviracastillo@ugr.es, jedv@ugr.es

† Dpt. Ingeniería Telemática, Escuela Superior de Ingenieros, Univ. de Sevilla
C/ Camino de los Descubrimientos s/n, 41092 Sevilla (Spain)
E-mail: {rafaestepa,aestepa,fjmc}@us.es

Resumen

ATT&CK establece un modelo donde se especifican las fases secuenciales de un ciberataque, así como las técnicas que suelen ser usadas en cada paso del ataque. Sería interesante incorporar este modelo en el proceso de detección de los ciberataques ya que facilitaría la correlación de las numerosas alertas generadas por los sistemas de monitorización de red. Sin embargo, la aplicación del modelo en los procesos de correlación de eventos no es inmediata, ya que no está formulado en términos de eventos observables y/o detecciones sino de acciones a realizar. En el presente trabajo exploramos y evaluamos los elementos necesarios para incorporar el modelo ATT&CK en el procesamiento de la información generada por los sistemas de monitorización de la seguridad en la red.

***Palabras Clave*—detección de intrusiones, modelado de ataques, correlación de alertas**

I. INTRODUCCIÓN

Los ciberataques son cada vez más frecuentes y con efectos más relevantes, haciendo que la seguridad de las redes adquiera una importancia crítica. Una buena defensa pasa, necesariamente, por el conocimiento del estado de la red y los sistemas y la rápida detección de los incidentes. Para ello los oficiales de ciberseguridad (CSO) recurren a los denominados sistemas de monitorización de la seguridad (NSM, del inglés *Network Security Monitoring*) [1].

Los NSM procesan información de múltiples fuentes y diversa naturaleza, como los flujos de tráfico, o las alertas generadas por los sistemas de detección de intrusiones (IDS, del inglés *Intrusion Detection Systems*) desplegados [2]. De esta forma, en un escenario típico, los NSM deben gestionar una enorme cantidad de eventos, lo que se suele traducir en un elevado número de alarmas que deben ser supervisadas. Lamentablemente, muchas de las alertas suelen corresponder a falsos positivos, a un mismo incidente o no resultan relevantes para el sistema monitorizado. Por ello, se suelen desplegar técnicas de correlación de alertas que intentan relacionar y agrupar todas las alertas asociadas a un mismo ataque. En la literatura hay numerosas propuestas en esta línea [3] [4] [5], la mayoría en base a la información de las propias alertas y a relaciones temporales entre ellas. Sin embargo, estas adolecen de importantes limitaciones, siendo necesario mejorar sus capacidades tanto en tasa de detección de ataques como en relación al número de alertas y falsos positivos (FP) generados.

Por otra parte, la mayoría de los ciberataques responden a una secuencia de acciones, que suelen dar lugar a múltiples alertas. En este sentido, el modelo ATT&CK de Mitre [6] establece un marco de trabajo basado en ataques reales para la descripción de los ataques. Por ello, sería interesante la incorporación de éste modelo en el proceso de detección y, particularmente, en la correlación de alertas. Sin embargo, el punto de vista considerado en el modelo ATT&CK está centrado en la ejecución del ataque [7], a partir del establecimiento de las diferentes fases (*tácticas*) y los métodos específicos posibles para su ejecución (*técnicas*). Consecuentemente, la utilización del modelo ATT&CK durante el proceso de detección de ataques requiere de un cambio de perspectiva, ya que es necesario identificar las técnicas a partir de los eventos observados. En primera instancia, esta identificación requiere del mapeo de dichas técnicas a las diferentes alertas generadas por los IDS.

El objetivo del presente trabajo es explorar los métodos y procedimientos necesarios para la aplicación del modelo ATT&CK. Nos planteamos cuatro cuestiones de interés:

- Detectabilidad: ¿cuál es la capacidad de los IDS para generar alertas para las diferentes técnicas?
- Identificación: ¿es posible identificar la técnica usada por un atacante a partir de la/s alerta/s observadas?

- Información de contexto: ¿resulta útil la incorporación de información contextual en el proceso?
- Incompletitud: ¿cómo se puede aplicar el modelo si no se detecta alguna fase/táctica?

El resto de este artículo se estructura como sigue. La sección II presenta un resumen de los trabajos relacionados y la motivación de la propuesta. La sección III describe el esquema de asociación entre eventos detectados por el sistema de monitorización y las técnicas de ataque del modelo. La sección IV plantea dos casos de estudio basados en nuestro esquema de asociación. Finalmente, la sección V presenta las conclusiones preliminares y las líneas de avance de este trabajo en curso.

II. TRABAJOS PREVIOS Y MOTIVACIÓN

Existen múltiples iniciativas en la línea de mejorar las capacidades de los NSM mediante la aplicación de técnicas de IA a fin de agregar la información relevante y descartar los FP. Así, hay numerosos trabajos orientados a la reducción de falsos positivos y a la agregación de alertas, p.e. [5]. No obstante, hasta la fecha, la mayoría de los trabajos presentan resultados poco significativos y centrados exclusivamente en la información contenida en las propias alertas [4], por lo que parece interesante explorar nuevas técnicas que puedan aplicarse de forma paralela o secuencial a la utilización de técnicas de IA.

Por otra parte, los incidentes de seguridad se corresponden habitualmente con una secuencia de acciones. En este sentido, se han establecido varios modelos de fases entre los que podemos destacar el modelo ATT&CK, que establece una secuencia de posibles fases en forma de tácticas Sin embargo, en la mayoría de los trabajos en la bibliografía se obvia la incorporación de estas fases o cualquier información adicional diferente a las propias alertas. Así, en [4] se hace un estudio detallado de los métodos de detección especialmente diseñados para hacer frente a los ataques multietapa. En ese trabajo se seleccionan 181 publicaciones como relevantes de las que solo 18 plantean el empleo de fuentes de información diferentes a las alertas y, de esos, tan solo 5 combinan otras huellas dejadas por los atacantes con alertas generadas por el IDS. En trabajos más recientes se considera explícitamente la inclusión del modelo en el proceso de correlación, típicamente en forma de grafos de ataque, p.e. [8], [9], [10].

En este contexto, nuestro objetivo es incorporar modelos de fases de ataque en el proceso de correlación de alertas/eventos a partir de la identificación de las técnicas/tácticas. Para ello, asumimos que la consecución de algunas de esas fases deja evidencias a modo de alertas en el IDS. En tal caso, es posible establecer una agrupación de alertas individuales y eventos asociados a las mismas (p.e. IPs, flujos, protocolos, etc.) en una única *hiperalerta* que contendrá toda la información conocida asociada al ataque.

Por otra parte, cada una de las fases puede implicar eventos, legítimos o no, que no generen alertas, pero que dejarán diversas evidencias en el NSM. La inclusión de estos eventos en el proceso de correlación puede aportar mejoras en el mismo y posibilitar la identificación de fases que no generen alertas. En este contexto, en [10] se usa el modelo para identificar posibles fases no detectadas y buscar eventos potencialmente relacionados en las trazas.

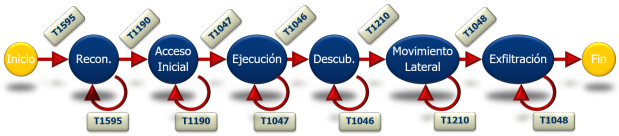


Figura 1. Modelo (autómata de estados finitos) para la secuencia de técnicas para uno de los ataques evaluados.

Nuestra propuesta se orienta, de forma similar a la propuesta en [10], a la incorporación del modelado en el proceso de correlación considerando eventos adicionales a las alertas. En particular, se plantea el uso del modelo ATT&CK como núcleo del sistema. A partir de este modelo, es posible definir autómatas asociados a los diferentes ataques (Fig. 1) en el que los estados son las diferentes tácticas involucradas y las transiciones se activan ante la observación de las técnicas utilizadas durante la ejecución del ataque. A diferencia de otros trabajos similares, p.e [8], que recurren a grafos de alertas/ataques en base a acciones previamente establecidas (p.e. un intento de acceso), nuestro foco se centra en identificar la técnica concreta que está siendo usada en cada fase (táctica). Por tanto, para la aplicación del modelo (autómata) se hace necesario mapear las técnicas a eventos observables. Como ya se ha mencionado, la aproximación inmediata sería establecer una correspondencia entre alertas y técnicas a través de los identificadores de las firmas (SID, del inglés *Signature IDentification*).

Se plantean, sin embargo, dos cuestiones a analizar. Por una parte, la correspondencia entre técnica y SID no será biunívoca, es decir, una técnica puede disparar múltiples alertas y, por otra parte, dada una alerta (SID) es posible que no pueda determinarse la técnica que la ha disparado de entre las posibles. Esto sugiere la necesidad de estudiar empíricamente las propiedades de este mapeo y la posible incorporación de información adicional discriminativa (p.e. tipos de flujos). Por otra parte, algunas técnicas podrían no generar alertas, bien por no existir actividades ilegítimas o firmas asociadas, bien por no existir actividad observable desde el punto de vista de tráfico generado en la red.

En este sentido, en trabajos previos (p.e. [11] y [12]), hemos constatado la existencia de un número significativo de técnicas que no generan alertas, por lo que es necesario abordar métodos adicionales para su detección en base a otras fuentes de información y/o para aplicar el modelo en ausencia de ellas. En consecuencia, en este trabajo nos planteamos también la incorporación de información contextual en el mapeo de técnicas a eventos detectados, tal como se propone en [10].

III. CARACTERIZACIÓN DE TÉCNICAS Y EVENTOS

En este trabajo utilizaremos un sistema de detección basado en *eventos* desarrollado por los autores en el que consideramos información de diversos sensores de red.

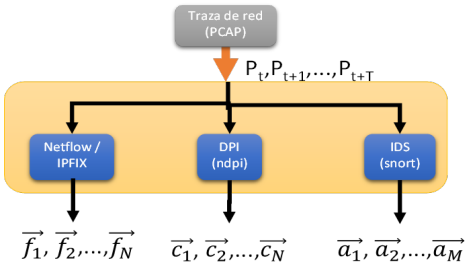


Figura 2. Eventos de entrada al sistema.

Una descripción detallada del mismo se puede encontrar en [11]. Inicialmente se consideran tres sensores que toman como entrada la secuencia de paquetes P_t , P_{t+1} , ... , P_{t+T} y generan a la salida, respectivamente, los siguientes observables (Fig. 2):

- Sensor IDS: Se usa un NIDS (*snort*) para generar una secuencia de alertas de intrusión, $\overrightarrow{a_1}$, $\overrightarrow{a_2}$, $\cdots$, $\overrightarrow{a_M}$. Estas contienen, al menos, marcas temporales, IPs implicadas e identificador de la firma activada (SID).
- Sensor DPI (*Deep Packet Inspection*): Se usa un clasificador de tráfico (*ndpi*) para generar una secuencia con los flujos existentes y sus correspondientes clasificaciones (protocolos), $\overrightarrow{c_1}$,$\overrightarrow{c_2}$, $\cdots$, $\overrightarrow{c_N}$. Incluyen, al menos, marca temporal, tupla y protocolo de capa de aplicación transportado.
- Sensor Netflow: Se usa IPFIX o equivalente para generar una secuencia de flujos y características de los mismos $\overrightarrow{f_1}$, $\overrightarrow{f_2}$, $\cdots$, $\overrightarrow{f_N}$. En este caso se incluye, al menos, marca temporal, tupla, duración y número de paquetes intercambiados.

Cada observable se etiqueta con un identificador único. Se considera que todo incidente debe generar al menos una alerta. A partir de los eventos observados se generan *hiperalertas*, h_i, que resultan de la agregación de los mismos de acuerdo a los procedimientos de correlación activos. Asimismo, la correlación de hiperalertas genera nuevas hiperalertas de niveles superiores, h_i^l, pudiendo establecerse diversos niveles de agregación, l. Un hiperalerta deben contener toda la información de los elementos que la integran.

En este contexto, la finalidad del modelo es guiar el proceso de agregación de hiperalertas a través de los diferentes estados y transiciones del autómata. De esta forma, el uso efectivo del modelo ATT&CK implica el establecimiento de una aplicación, $\mathcal{M}$, que asigne una correspondencia entre el conjunto de técnicas, $\mathcal{T}$, y las hiperalertas, $\mathcal{H}$,

$$\mathcal{M} : \mathcal{T} \rightarrow \mathcal{H} \tag{1}$$

Las propiedades de la aplicación dependerán de las de las hiperalertas, que a su vez dependerán tanto de los procedimientos de correlación activos como de la información a considerar. Dado que nuestro objetivo es evaluar estas propiedades, consideraremos en este trabajo una función de correlación ideal que incorpore la información en base a la técnica correspondiente. Así, obtendremos un conjunto de hiperalertas, $\mathcal{H}$, cada una asociada a una instanciación, k, de una técnica, t, diferente:

$$\mathcal{H} = \{h_{kt}/t \in \mathcal{T}\} \tag{2}$$

Las cuestiones más relevantes a analizar son: determinar si es posible establecer elementos comunes en las hiperalertas asociadas a una misma técnica; si existen solapamientos entre esos elementos comunes para diferentes técnicas; y si estos permiten inferir de forma unívoca la técnica.

IV. CASO DE ESTUDIO

Para obtener una primera evaluación de las características y propiedades tanto de $\mathcal{H}$ como de $\mathcal{M}$, se ejecutan y analizan diversas técnicas individualmente. Esto permite obviar los procedimientos de correlación, ya que todos los eventos identificados podrán ser incorporados en la hiperalerta correspondiente.

En la elección de técnicas a implementar se ha considerado la generación de actividad en la red susceptible de ser monitorizada y detectada [12], descartándose las restantes.

A efectos de aplicar el modelo, se usan combinaciones de técnicas de acuerdo a los diferentes pasos y fases de ataques multietapa seleccionados. Por razones de brevedad, nos centraremos en el análisis de un ataque de exfiltración de datos con 6 fases (Fig. 2) en el que se ejecutan 6 técnicas diferentes, cada una asociada a una de las fases (Cuadro I).

A. *Uso de alertas IDS*

La solución más simple para identificar las técnicas es la utilización únicamente de las alertas activadas. En este caso, la hiperalerta únicamente incorpora como información relevante los identificadores de las alertas (SIDs).

$$h_{kt} = \langle sid_1, sid_2, \cdots, sid_m \rangle \tag{3}$$

En el Cuadro I se muestra el número de alertas diferentes (algunas se activan múltiples veces) y los SID activados en cada técnica para el ataque multietapa considerado. Se utiliza Snort con reglas Talos en una configuración de máxima detección (todas las reglas activadas).

Se pueden observar tres cuestiones relevantes. En primer lugar, hay técnicas que no activan alertas. Esto era algo esperado. De hecho, en [12] hemos identificado que sólo 18 de las 32 técnicas evaluadas han generado alertas. Lógicamente, los resultados empeoran con la configuración por defecto para las reglas Talos. Por tanto, aparece un problema de detectabilidad. En segundo lugar, el número de SIDs es muy variable en función de la técnica. Como se puede observar en el Cuadro I, hay técnicas que disparan hata 46 SIDs, mientras que otras únicamente disparan 1. En tercer lugar, si se evalúan las reglas y los motivos de detección, nos encontramos con que en muchos casos se ha realizado una detección genérica. A modo de ejemplo, tanto la técnica T1047 como la T1210 activan la regla 12, que corresponde a *stream5: TCP Small Segment Threshold Exceeded*; o la técnica T1190 activa una detección por *SQL xp_cmdshell attempt*. Es decir, se activan alertas por el uso de fragmentos muy pequeños o el intento de acceso al

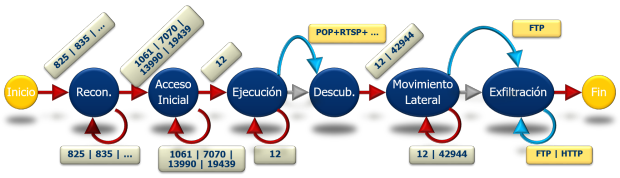


Figura 3. Autómata en base a SIDs y flujos para el ataque de ejemplo. En gris, transiciones por SID, en amarillo transiciones por tipo de flujo.

recurso *xp_cmdshell*, respectivamente. Por tanto, no parece que sea posible identificar la técnica empleada únicamente por las SIDs activadas.

Por otra parte, los resultados obtenidos para otras técnicas muestran solapamientos, incluso completos, en las SIDs activadas. Por tanto, no es posible identificar la técnica empleada únicamente a partir de las mismas.

B. *Uso adicional de información de flujos*

Una posible solución consiste en añadir información adicional. En este caso, hemos considerado los tipos de flujos identificados,

$$h_{kt} = \langle sid_1, \cdots, sid_m, tipo_1, \cdots, tipo_n \rangle \qquad (4)$$

Los resultados para el ejemplo considerado se muestran en el Cuadro II. Como se puede observar, de nuevo aparecen diferencias significativas en el número de elementos para cada técnica. Resulta relevante que la fase de descubrimiento (T1046) genera un muy elevado número de flujos, a la vez que pasa inadvertido por el IDS. Por el contrario, la de exfiltración (T1048), que tampoco dispara alertas, únicamente genera 50 flujos de tipo HTTP y FTP_CONTROL.

La exploración realizada muestra que el mapeo entre técnicas y eventos observables no es un problema trivial y que el uso exclusivo de las alertas resulta insuficiente. La inclusión de información sobre flujos puede ser de utilidad para la aplicación del modelo (Fig. 3), ya que se podrían activar transiciones en base a los tipos de flujos observados. Así, en el ejemplo considerado, la observación de un elevado número de flujos de diferente naturaleza procedentes de la IP considerada atacante una vez alcanzado el estado de ejecución se podría considerar suficiente para activar la transición al estado de descubrimiento. Análogamente, a partir del movimiento lateral, la existencia de flujos FTP_CONTROL haría transicionar al estado de exfiltración de datos, finalizándose la ejecución del ataque multietapa.

V. CONCLUSIONES Y TRABAJO FUTURO

Este trabajo realiza una exploración preliminar del mapeo entre técnicas de ataque y los eventos observables en el tráfico de red. Nuestros primeros resultados sugieren que los IDS no ofrecen suficiente detectabilidad y que la identificación de la técnica es difícil. Esto aconseja el uso futuro de técnicas de ML para la asociación entre hiperalertas y técnicas donde se considere información de contexto. Sería necesario recopilar y analizar muchas otras técnicas para poder generalizar los resultados preliminares obtenidos.

Cuadro I
SIDs ACTIVADOS PARA CADA TÉCNICA

Técnica	Ndif	SIDs
T1595	46	825, 835, 839, 845, 849, 853, $\cdots$, 43290
T1190	4	1061, 7070, 13990, 19439
T1047	1	12
T1046	0	-
T1210	2	12, 42944
T1048	0	-

Cuadro II
FLUJOS Y TIPOS DETECTADOS PARA CADA TÉCNICA

Técnica	N	Tipo de flujos
T1595	13093	DNS, ... (+50 tipos)
T1190	156	HTTP
T1047	1736	HTTP, MDNS, IGMP, IGMPv6
T1046	131366	POP, RTSP, ... (+150 tipos)
T1210	4	ICMP
T1048	50	HTTP, FTP_CONTROL

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por FEDER / Junta de Andalucía - Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto A-TIC-224-UGR20 y por el proyecto de I+D+i PID2020-115199RB-I00 financiado por MICIN/AEI/10.13039/501100011033.

REFERENCIAS

[1] I. Ghafir; V. Prenosil; J. Svoboda; M. Hammoudeh; "A Survey on Network Security Monitoring Systems", Proc. 2016 IEEE 4th Int. Conf. on Future Internet of Things and Cloud Workshops (FiCloudW), 77-82, 2016.

[2] P. García-Teodoro; J. Díaz-Verdejo; G. Maciá-Fernández; E. Vázquez; "Anomaly-based network intrusion detection: techniques, systems and challenges", Computers & Security, 28:18-28, 2009.

[3] S. Salah; G. Maciá-Fernández; J. E. Díaz-Verdejo; "A model-based survey of alert correlation techniques", Computer Networks, 57(5):1289-1317, 2013.

[4] J. Navarro; A. Deruyver; P. Parrend; "A systematic survey on multi-step attack detection", Computers & Security, 76:

[5] Spathoulas, G.; Katsikas. S.; "Enhancing IDS performance through comprehensive alert post-processing", Computers & Security, 37:176-196, 2013.

[6] Strom, Blake E., et al., "MITRE ATT&CK: Design and Philosophy", The MITRE Corporation, Technical Report No. MP180360R1, 2020.

[7] Strom, Blake E., et al., "Finding cyber threats with ATT&CK-based analytics", The MITRE Corporation, Technical Report No. MTR170202, 2020.

[8] Sen, O, et al., "On using Contextual Correlation to detect Multi-stage Cyber Attacks in Smart Grids", Sustainable Energy, Grids and Networks, 32, 100821, 2022.

[9] Milakerdi, S., et al., "HOLMES: Real-time APT Detection through Correlation of Susipicious Information Flows", in Proc. 2019 IEEE Symp. on Security & Privacy, 1137-1152, 2019.

[10] Wang, Y, Guo, Y., Fang, C., "An end-to-end method for advanced persistent threats reconstruction in large-scale networks based on alert and log correlation", Journal of Information Security and Apps., 71, 103373, 2022.

[11] E. Castillo-Fernández, et al. "Multistep Cyberattacks Detection using a Flexible Multilevel System for Alerts and Events Correlation, Proc. European Interdisciplinary Cybersecurity Conference (EICC 2023), pp. 6, 2023.

[12] Andrés García Coronado (autor), Rafael María Estepa Alonso (tutor); "Detección en fase temprana de ciberataques con rastro en la red", Proyecto Fin de Carrera, Ing. Telecomunicación, Univ. Sevilla, 2022.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Evaluación experimental de las capacidades de detección de ciberataques basados en técnicas del modelo ATT&CK mediante Snort

Javier Muñoz-Calle[1], Javier Fructuoso[1], Rafael Estepa[1], Antonio Estepa[1] y Jesús Díaz-Verdejo[2]
[1] Dpt. Ingeniería Telemática, Escuela Superior de Ingenieros, Univ. de Sevilla
C/ Camino de los Descubrimientos s/n, 41092 Sevilla (Spain)
fmjc@us.es, javierfructuoso8@gmail.com, rafaestepa@us.es, aestepa@us.es
[2] Dpt. Teoría de Señal, Telemática y Comunicaciones, CITIC, Univ. de Granada
C/ Periodista Daniel Saucedo Aranda, s/n, 18071 Granada (Spain)
jedv@ugr.es

Los sistemas de detección de intrusiones (IDS) permiten detectar actividades maliciosas y generar alertas a supervisar por los operadores, constituyendo el núcleo de los sistemas de monitorización de ciberseguridad. Tradicionalmente, se ha asumido que los IDS basados en firmas (SIDS) ofrecen una capacidad de detección apropiada con una baja tasa de falsos positivos, aunque estudios recientes ofrecen datos que apuntan en sentido contrario. En este trabajo experimental se explora la capacidad de detección del SIDS Snort, con los paquetes de reglas gratuitos *Talos* y *ETOpen*, sobre las técnicas de ataque definidas por MITRE ATT&CK. Para ello se implementan y analizan técnicas pertenecientes a las distintas tácticas (fases) de la matriz ATT&CK. Este trabajo se encuentra aún en desarrollo, por lo que en esta contribución temprana se muestran los primeros resultados relativos a la importancia del conjunto de reglas empleado en la capacidad de detección de las diversas fases de los ataques. Los resultados indican que la capacidad de detección varía entre el 72% y el 28% en función de las reglas utilizadas, encontrando siempre una tasa de falsos positivos inferior al 5,5%.

***Palabras Clave*- detección de intrusiones, ciberseguridad, modelo ATT&CK**

I. INTRODUCCIÓN

Los sistemas de detección de intrusos (*Intrusion Detection Systems*, IDS, por sus siglas en inglés) [1] se consideran un elemento fundamental en la protección y monitorización de los sistemas y redes, constituyendo uno de los elementos clave del despliegue de la seguridad en capas. Por ello, resulta relevante que los IDS presenten un alto rendimiento, lo que se traduce en gran capacidad de detección y una tasa de falsos positivos muy contenida.

Los IDS se pueden agrupar en dos categorías principales en función de la forma en la que se realiza la detección [1]: basados en firmas (SIDS, *Signature based IDS*) y basados en anomalías (AIDS, *Anomaly-based IDS*). Los primeros utilizan firmas o patrones extraídos de ataques conocidos, mientras que en los segundos se establecen modelos de actividad que determinan la naturaleza de los eventos. Habitualmente se argumenta que los SIDS presentan altas tasas de detección para ataques conocidos, lo que los hace adecuados para su despliegue en sistemas en explotación. De hecho, la mayoría de los IDS en uso en escenarios reales son de este tipo. Existen múltiples estudios que abordan el uso de SIDS para la detección de ataques [2] [3], pero no cuantifican la capacidad de detección de forma experimental o, caso de hacerlo, emplean *datasets* obsoletos y escasamente representativos de la realidad. Los pocos estudios que sí realizan esta cuantificación se centran en tráfico específico de alguna aplicación concreta. Así, para tráfico web, resultados recientes [4][5] muestran que la capacidad de detección puede variar entre el 7% y el 78% en función de las reglas utilizadas. Aquellos conjuntos de reglas que presentan mayor detección presentan una tasa muy elevada de falsos positivos. Todo esto plantea la necesidad de corroborar de forma adecuada el rendimiento de los SIDS en distintas tipologías de ataque, lo que motiva el presente trabajo.

Dado que Snort [6] es uno de los SIDS gratuitos más ampliamente utilizados y que está presente en la gran mayoría de las instalaciones, utilizaremos dicho software con diversos conjuntos de reglas públicas a fin de evaluar la dependencia de los resultados con el conjunto de reglas utilizado.

Para validar la capacidad de detección de los SIDS resulta necesario disponer de *datasets* de ataques que sean

representativos y que se encuentren actualizados. Sin embargo, nos encontramos con que existen pocos *datasets* públicos disponibles para hacer experimentación sobre SIDS en general (e.g., KDD'99 [7], CAIDA [8], UNSW-NB15 [9], o HIKARI-2021 [10]). En dichos *datasets*, los ataques aparecen en número reducido, son antiguos y mayoritariamente de sistemas simulados. Esta carencia de cercanía a la realidad y la falta de representatividad los hace poco aplicables a entornos de producción [11] y, consecuentemente, los invalidan para desarrollar o evaluar sistemas de detección de ataques válidos en escenarios reales modernos [12]. Este trabajo pretende cubrir algunas de estas carencias recreando ataques conforme a la categorización de la matriz MITRE ATT&CK para entornos empresariales [13], la cual agrupa las posibles técnicas empleadas en los ataques en categorías denominadas tácticas que se corresponden con las diversas fases de los ataques.

El estudio se encuentra actualmente en elaboración y en este artículo presentaremos sus resultados preliminares. Las principales contribuciones son: a) generación de un *dataset* con ataques en red categorizados conforme a la taxonomía de la matriz MITRE, y b) evaluación inicial de la capacidad de detección de un SIDS, así como de la influencia sobre la misma del conjunto de reglas empleado y del tipo de ataque realizado. Estas contribuciones permitirán estimar el nivel de seguridad ofrecido por los SIDS en el análisis de tráfico de red. Por otra parte, el *dataset* podría también utilizarse para evaluar sistemas basados en detección de anomalías.

II. Metodología empleada en el estudio

Para realizar el estudio se ha analizado la capacidad de detección de distintas técnicas de ataque de la matriz MITRE ATT&CK Enterprise [13]. Esta matriz es una base para la recolección de información acerca de los diferentes tipos de ataques existentes, a partir de observaciones del mundo real. Las técnicas empleadas en los ataques se clasifican en diversas fases que se ejecutan en un orden secuencial, empezando por la fase de reconocimiento, pasando por las escaladas de privilegios, hasta llegar a las últimas fases de exfiltración o impacto. En la matriz ATT&CK Enterprise se determinan 14 de estas fases, a las que se denomina tácticas. Para cada una de las tácticas, la matriz define diversas técnicas (y subtécnicas), que son las posibles acciones que permitirían lograr el objetivo de esa táctica. Por ejemplo, la técnica de escaneo se encuadraría en la táctica de reconocimiento y en la de movimientos laterales. Para cada técnica se recogen los posibles mecanismos de detección, mitigación y ejemplos de ataques concretos.

Dado que este estudio se centra en la capacidad de detección de ataques basados en tráfico de red, el primer paso del trabajo consiste en hallar qué técnicas de cada categoría de la matriz corresponden a ataques rastreables mediante tráfico de red. Para ello, se ha analizado la información de detección de cada técnica, identificando aquellas que como fuente de información para la detección (*Data Source*) incluyen tráfico de red (*Traffic Network*). Este proceso, realizado de forma automatizada con un script en *Python*, y cuyo resultado ha sido validado de forma manual, ha dado como salida un total de 69 técnicas detectables por el tráfico de red. De estas técnicas, para el presente trabajo se han seleccionado 18 mecanismos concretos de ataque, para cuya implementación existen herramientas gratuitas disponibles. La Tabla I recoge el número de técnicas con rastro en red encontradas por cada una de las tácticas, así como el número de ellas que han sido finalmente implementadas en el presente trabajo.

Tabla I
Técnicas de ataque por categoría basadas en tráfico de red

Táctica	Nº	Implementadas
RECONNAISSANCE	3	2
RESOURCE DEVELOPMENT	2	1
INITIAL ACCESS	4	1
EXECUTION	2	1
PERSISTENCE	6	1
PRIVILEDGE ESCALATION	0	0
DEFENSE EVASION	9	1
CREDENTIAL ACCESS	5	1
DISCOVERY	3	1
LATERAL MOVEMENT	5	2
COLLECTION	2	1
COMMAND AND CONTROL	16	2
EXFILTRATION	7	1
IMPACT	5	3

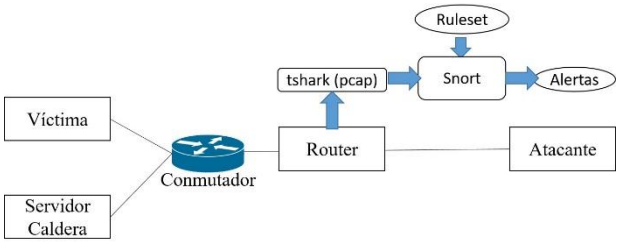


Fig. 1. Escenario de prueba y proceso de detección

La ejecución de cada uno de los ataques seleccionados ha requerido implementar un escenario de prueba con diversas máquinas y sistemas operativos (Ubuntu, Windows 10, Windows Server, Kali Linux). Se ha utilizado CALDERA [14] como gestor de aplicaciones, y se han seleccionado múltiples herramientas de código abierto y gratuitas (*hping3*, *metasploit*, *hydra*, …) que permitan ejecutar los diferentes ataques, capturando su tráfico correspondiente.

Para cada una de las técnicas de ataque realizadas se ha coleccionado el tráfico de red generado en formato *pcap*. Posteriormente, dicho tráfico ha sido inyectado en el SIDS *Snort* configurado con diferentes conjuntos de reglas, analizando a continuación las alarmas generadas a fin de clasificarlas como verdaderos positivos (detección de la técnica) o falsos positivos (Fig. 1).

En la Tabla II se muestran los ataques ejecutados de cada táctica, así como las herramientas utilizadas para su implementación. El *dataset* de tráfico de ataques generado ha sido tomado como entrada por el SIDS *Snort* en su versión 3.0. A fin de evaluar el rendimiento de diversos conjuntos de firmas para *Snort*, se han utilizado los paquetes:

- *Talos* (versiones *Community* y *Registered LightSPD*, 07/06/2022): reglas utilizadas por

Tabla II
DETECCIÓN DE ATAQUES POR EL IDS SNORT

TÁCTICA	ATAQUE (HERRAMIENTA)	TALOS		ETOPEN	
		Nº Alertas	FP	Nº Alertas	FP
RECONNAISSANCE	Escaneo de portal web (*Dirb*)	50	0	9	1
	Transferencia de zona DNS (*command host*)	1	0	0	0
RESOURCE DEVELOPMENT	Fuerza bruta hacia login web (*Hydra*)	2	1	0	0
INITIAL ACCESS	Inyección SQL (*sqlmap*)	7	1	6	0
EXECUTION	Ejecución de *payload* mediante WinRM (*Metasploit*)	2	0	0	0
PERSISTENCE	Web Shell (*Metasploit*)	2	1	0	0
DEFENSE EVASION	Ejecución de scriptlet COM remoto (CALDERA)	0	0	0	0
CREDENTIAL ACCESS	Envenamiento ARP (*Metasploit*)	0	0	0	0
DISCOVERY	Detección de servicios y versiones (CALDERA)	9	1	7	0
LATERAL MOVEMENT	*EternalBlue* (*Metasploit*)	5	0	4	0
	Ejecución código remoto servidor FTP vulnerable (*Metasploit*)	4	1	0	0
COLLECTION	Obtención de información de MIB (*Metasploit*)	2	0	1	0
COMMAND AND CONTROL	Túnel C2 mediante protocolo DNS (*dnscat2*)	1	0	0	0
	Sesión SSH sobre puerto no estándar (ssh)	0	0	0	0
EXFILTRATION	Exfiltración de archivo sobre ICMP (*hping3*)	3	0	0	0
IMPACT	Denegación de servicio de servidor web (*Metasploit*)	1	0	0	0
	Inundación UDP (*hping3*)	0	0	0	0
	Descarga y ejecución de software de minado (CALDERA)	1	1	0	0

defecto por *Snort*, ampliadas con las reglas para usuarios registrados, ambas desarrolladas por el grupo de inteligencia de amenazas Cisco VRT (*Vulnerability Research Team*) [15].

- ETOpen (versión open, 07/06/2022): reglas, a disposición de la comunidad, desarrolladas para *Snort* por el grupo de investigación en ciberseguridad ET (*Emerging Threats*) [16].

Los ataques realizados pueden considerarse clásicos y no explotan ninguna vulnerabilidad tipo '0-day', por lo que la última versión de firmas utilizada debería poder detectarlos.

Cada una de las alarmas generadas por *Snort* ha sido posteriormente verificada a mano para determinar si se corresponde con los ataques realizados o, por el contrario, es un falso positivo.

Esta metodología permite encontrar, para cada una de las tácticas o fases de un ataque, el rendimiento en la detección de las firmas más utilizadas en Snort. A continuación, presentaremos los resultados preliminares

III. RESULTADOS EXPERIMENTALES

La implementación de los ataques ha generado un *dataset* de tráfico de ataques en red, constituido por 1.731.315 mensajes. El tráfico, que incluye sólo los ataques realizados, ha sido clasificado de acuerdo a las distintas categorías y técnicas de la matriz MITRE asociadas a estos ataques. Se encuentra disponible para la comunidad investigadora, junto con los resultados, en https://github.com/javfrumar1/detecciones_con_snort_de_ataques_realizados_con_CALDERA/.

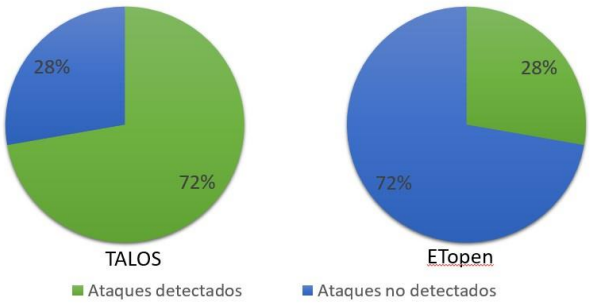


Fig. 2. Capacidad de detección de Snort según las reglas usadas

La Tabla II recoge los resultados de la detección con los distintos conjuntos de reglas empleadas (*Talos* y *ETOpen*), incluyendo el número de alertas diferentes que ha disparado cada ataque, así como cuántas de ellas corresponden a una detección errónea del ataque (FP).

A partir de los datos obtenidos se observa cómo las reglas *Talos* han detectado un 72% de los ataques realizados (13 de los 18 ataques, generándose un total de 90 alertas con 6 FP), mientras que las reglas *ETOpen* sólo han alcanzado una capacidad de detección del 28% (detectando sólo 5 de los 18 ataques, con un total de 27 alertas y 1 FP), lo que pone de manifiesto la relevancia de las reglas utilizadas en la capacidad de detección (Fig. 2). Como contrapartida, se aprecia que las reglas *Talos* han presentado una mayor tasa de FP (6,6% frente al 3,7% de *ETOpen*) (Fig. 3).

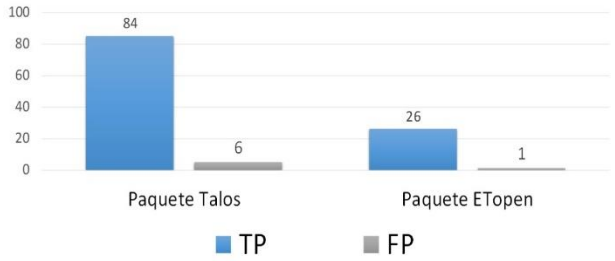


Fig. 3. Tasa de TP (alertas correctas) y FP según las reglas usadas

Si analizamos los resultados en base a las 13 tácticas existentes, se aprecia que las reglas *Talos* ofrecen una capacidad de detección del 84,6% (con ataques detectados en 11 tácticas), frente al 38,46% de las reglas *ETOpen* (con ataques detectados en sólo 5 tácticas, todas también identificadas por Talos). Observamos así que la influencia de la tipología de los ataques sobre la capacidad de detección está fuertemente condicionada por el paquete de reglas empleado.

Analizando la capacidad de detección para cada táctica de ataque (Fig. 4), se identifican dos de ellas que, como cabía esperar por su comportamiento, no han sido detectadas por ninguno de los paquetes de reglas.

IV. Conclusiones y Líneas de Avance

La disponibilidad de *datasets* adecuados es clave para acelerar la investigación en el campo de los IDS. En este trabajo se ha generado un *dataset* básico de ataques que cubre las distintas técnicas aplicables de la matriz MITRE.

La capacidad de detección mostrada por *Snort* está totalmente condicionada por el paquete de reglas empleado. En el mejor de los casos, se ha obtenido una capacidad de detección de ataques cercana al 72%. Así mismo, se identifican técnicas de ataque no detectadas por ninguno de los paquetes de reglas empleados. Todo ello pone de manifiesto que, si bien los SIDS representan una herramienta esencial en la seguridad de un sistema, deben ser mejorados o completados con otros mecanismos cómo IDS basados en host (HIDS) o AIDS, tal como sugieren diferentes estudios [17]. Estos resultados son preliminares y el trabajo continúa en curso. En las siguientes fases del mismo se abordarán retos como la ampliación del tamaño del *dataset* mediante el uso de nuevas herramientas y ataques, así como el análisis de la eficiencia en la detección por tipo de ataque y por herramienta. También se pretende comparar el comportamiento que mostraría un IDS basado en anomalías frente al mostrado por *Snort* mediante firmas.

Como limitaciones de este trabajo pueden destacarse el uso de herramientas gratuitas y la limitación del número de ataques empleado (tan sólo 18 de 69 técnicas) lo que restaría capacidad de generalización a los resultados. El uso de nuevas herramientas y ataques enriquecería la aplicabilidad de los resultados. Estos aspectos deberán ser tratados en posibles ampliaciones al trabajo.

Agradecimientos

Esta publicación es parte del proyecto de I+D+i PID2020-115199RB-I00 financiado por MICIN/AEI/ 10.13039/501100011033.

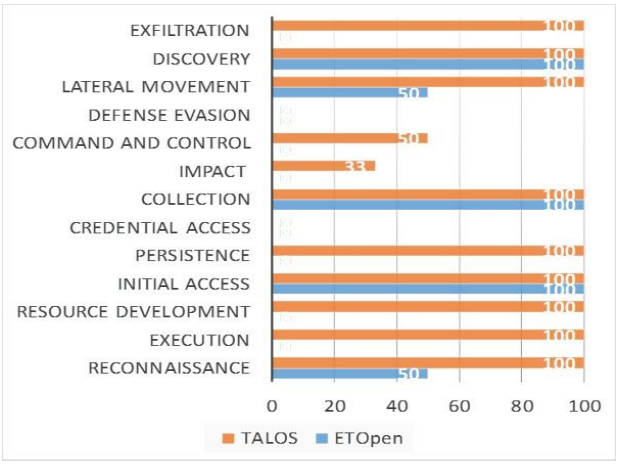


Fig. 4. Capacidad de detección de Snort por táctica

Referencias

[1] N. Moustafa, J. Hu, J. Slay, "A holistic review of Network Anomaly Detection Systems: A comprehensive survey", Journal of Network and Computer Applications, vol. 128, pp. 33–55, 2019.

[2] N.S. Mangrulkar, A.R.B. Patil and A.S. Pande, "Network Attacks and Their Detection Mechanisms: A Review", International Journal of Computer Applications, vol. 90, no. 9, 2014.

[3] N. Hoque, M.H. Bhuyan, R.C. Baishya, D.K. Bhattacharyya, and J.K. Kalita, "Network attacks: Taxonomy, tools and systems", Journal of Network and Computer Applications, vol. 40, pp. 307-324, 2014.

[4] P. Garcia-Teodoro, J. Díaz-Verdejo, J.E. Tapiador, R. Salazar-Hernandez, "Automatic generation of HTTP intrusion signatures by selective identification of anomalies", Computers and Security, vol. 55, pp. 159–174, 2015.

[5] J. Díaz-Verdejo, J. Muñoz-Calle, A. Estepa, R. Estepa, and G. Madinabeitia, "On the detection capabilities of signature-based Intrusion Detection Systems in the context of web attacks", Applied Sciences, vol. 12, no 2, pp. 852, 2022.

[6] "Snort intrusion detection tool", The Snort Team, https://www.snort.org/ (último acceso 01 junio 2023)

[7] S. Hettich, S.D. Bay, "The UCI KDD Archive". Univ. of California, Dep. of Information & Computer Science, http://kdd.ics.uci.edu, 1999.

[8] Cooperative Association for Internet Data Analysis (CAIDA) datasets, 2008.

[9] N. Moustafa, J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems", Military Communications and Information Systems Conference (MilCIS), pp. 1-6, 2015.

[10] A. Ferriyan, A.H. Thamrin, K. Takeda, J. Murai, "Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic", Applied Sciences, vol. 11, no. 17, pp. 7868, 2021.

[11] I. Sharafaldin, A. Gharib, A.H. Lashkari, A.A. Ghorbani, "Towards a reliable intrusion detection benchmark dataset", Software Network, vol. 1, pp. 177-200, 2018.

[12] R. Sommer, V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection", IEEE symposium on security and privacy, pp. 305-316, 2010.

[13] "MITRE ATT&CK Enterprise Matrix", The MITRE Corporation, https://attack.mitre.org/matrices/enterprise/ (último acceso 01 mayo 2023).

[14] "Framework CALDERA", The MITRE Corporation, https://caldera.readthedocs.io/ (último acceso: 01 mayo 2023].

[15] "Ruleset Talos", The Talos team, https://www.snort.org/talos/ (último acceso 01 junio 2023).

[16] "Ruleset ETopen", Emerging Threats, https://doc.emergingthreats.net/ (último acceso 01 junio 2023).

[17] J.E. Rubio, C. Alcaraz, R. Roman, J. Lopez, "Current cyber-defense trends in industrial control systems", Computers & Security, vol. 87, pp. 101561, 2019.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Towards Trustworthy Federated Learning: A privacy-preserving and secure protocol

Idoia Gamiz[1], Cristina Regueiro[2], Eduardo Jacob[1], Oscar Lage[2], Marivi Higuero[1]

[1]Department of Communications Engineering. University of the Basque Country, 48013 Bilbao, Bizkaia, Spain

{idoia.gamiz, eduardo.jacob, marivi.higuero}@ehu.eus

[2]TECNALIA, BRTA. Bizkaia Science and Technology Park, 700, E-48160 Derio, Bizkaia, Spain

{cristina.regueiro, oscar.lage}@tecnalia.com

With the growing demand for collaborative machine learning, ensuring privacy has become not only a legal and ethical responsibility but also a critical factor in preserving the value of each company. Federated Learning has gained significant attention as a decentralized approach for the training of machine learning models with data from various sources. However, it presents new challenges in terms of data privacy and model security. This work introduces a protocol specifically designed for facilitating model exchange during training, with a focus on ensuring resistance against privacy and security attacks. The analysis demonstrates that the proposal has significant advantages in terms of privacy while allowing flexibility in the selection of a suitable method to tackle security attacks.

***Palabras Clave*—Federated Learning, privacy, security, Machine Learning, Artificial Intelligence**

I. INTRODUCTION

Artificial Intelligence (AI) has emerged as a transformative technology across various industries. From predictive maintenance to natural language processing, AI applications continue to reshape the way we live and work. The availability of high-quality and diverse data to train machine learning models is critical to the success of AI. However, AI systems require data from various sources to achieve optimal performance and generalizability. In this context, the unauthorized access to personal information has raised concerns about the potential misuse of sensitive data. Compromising data can arise severe implications including identity theft, financial fraud, or even the loss of value for a company in some way.

Federated Learning (FL) has emerged as a promising solution to collaboratively train machine learning models while protecting data privacy. Nevertheless, it has been shown [1] that the updated parameters of the local models can uncover information about the data (data privacy attacks) and that an adversary could change some of the local models or train the models with manipulated data to avoid the convergence of the global model (model security attacks).

A significant dilemma arises in balancing privacy and security [1]. Ensuring security often implies maintaining complete control over the training process or even accessing the training data. However, this approach can conflict with the use of cryptographic techniques commonly employed to protect privacy.

The primary aim of this study is to find a technique that effectively addresses both issues simultaneously. To do so, this paper proposes a protocol for conducting the exchange of the models in a FL scheme so that the link between each local model and its generator is lost. Furthermore, it presents a privacy and security analysis to demonstrate how the use of this protocol brings significant advantages in both aspects.

II. RELATED WORK

Different works have been presented in the literature regarding data privacy and model security in FL. Most works use cryptographic techniques such as Secure Multi-Party Computation [2], Homomorphic Encryption [3], [4], [5] or Differential Privacy [6] to tackle the privacy issues. The use of those techniques can influence the accuracy of the final model or limit the security measures to specific techniques against particular attacks. Similar to the approach presented in this study, other works [7], [8] also anlyze the concept of losing the connection between an update and its owner. The former [7] tackles the privacy attacks by a protocol that randomly exchanges and mixes fragments of the local updates before sending them to the central authority. To deal with security attacks, they measure the quality of each fragment before the aggregation. The latter [8] presents a protocol for unlinkable anonymity based on the co-utility property to deal with privacy concerns and on reputations to motivate the parties to behave honestly. However, none of these approaches considers collusion between the server and some clients.

III. PROPOSED PROTOCOL

The current work proposes a protocol for the exchange of the local models during a FL scheme, which is called recursively during the iterative process. This section first presents a high-level overview of how the protocol integrates with the general scheme and later it gives an in-depth explanation of its operation for a concrete iteration.

A. High-level overview

A server S proposes a training to a set of n data owners referred to as clients $C = \{C_1, ..., C_n\}$. S initiates contact with each of the clients and provides them with information regarding the algorithm to be trained and the necessary data. After acceptance, the training takes place through an iterative process of T iterations. In each iteration $t = \{1, \dots, T\}$ the process is as follows.

1) S sends the global model ω^t to each of the clients.
2) The clients use the protocol (see Section B) to train and transmit the models to S while maintaining the disassociation between each model and its generator. The clients are represented by C_i^t where the indexes t and i represent that the party holds position i in the cycle at iteration t, for all $i = 1, ..., n$. Each client C_i^t uses a portion of their data for that iteration, obtaining their own local model ω_i^t.
3) When S receives the n local models, S performs the model metrics and the robust aggregation they desire and obtains the global model for the next iteration, ω^{t+1}.

For simplicity, the paper explains the procedure of the protocol for a particular iteration. So from now on, $C_1^t \equiv C_1, \dots, C_n^t \equiv C_n$.

B. Operation

Before starting the iteration, S owns a list with the identifications of all the clients, $L = \{IDC_1, \dots, IDC_n\}$ and the clients own the parameters of the global model ω^t and the algorithm to be trained. The model exchange is done through a cycle of unfixed order and the idea is to prevent S and the clients from knowing the order before starting the protocol. Then, the protocol is conducted in the following manner.

1) S randomly chooses C_1 and sends the list L with the identification of the remaining clients, i.e. $L = L \setminus \{IDC_1\}$.
2) C_1 trains the model with their own data obtaining ω_1^t. C_1 randomly chooses the client C_2 from L and sends ω_1^t and $L = L \setminus \{IDC_1^t\}$.
3) C_i trains the model with their own data obtaining ω_i^t. C_i randomly chooses the client C_{i+1} from L and sends $L = L \setminus \{IDC_i^t\}$. For $i = \{2, ..., n-1\}$, client C_i could hold models from previous clients, so C_i must distribute both their own model and the previous models between S and the next client. If C_i receives $m-1$ models from C_{i-1}, C_i must transmit m models (adding their own local model ω_i^t) either to S or to the client C_{i+1}. The only condition is that the client C_i must send at least one model to C_{i+1}. So the number of options for distributing the m models can be calculated by $\sum_{k=1}^{m} \binom{m}{k}$. Table I shows all the possibilities for an example where $n = 4$.
4) C_n trains the model with their own data and obtains ω_n^t. C_n receives an empty list L, so C_n sends all the remaining models to S and the process ends.

The condition of sending at least one model to the next client is imposed to make sure that C_n receives at least a local model and distributes two models to S. In that way, S cannot know which of the models belongs to C_n. Similarly, none of the intermediate clients is able to know how many clients there were before them, so they cannot know the ownership of the models they receive.

The total number of communications depends on the decisions made by the clients, but the goal is to reach an optimal balance between the minimum and maximum number of rounds. As shown in the previous steps, the only possible communications in each iteration are:

- $S \rightarrow C_1$
- $C_1 \rightarrow C_2$
- $C_i \rightarrow C_{i+1}$, $C_i \rightarrow S$ [optionally] $\forall i = \{2, ...n-1\}$
- $C_n \rightarrow S$

Let c be the number of communications between the parties. Then, the minimum takes place when all the parties perform a single communication by sending the full set of local models to the next client. So, $min(c) = 1+1+1(n-2)+1 = n+1$. Conversely, the maximum takes place when all the clients choose the option of sending some of the local models to the next client and the rest to S. So, $max(c) = 1 + 1 + 2(n-2) + 1 = 2n - 1$. This means that $c \in [n+1, 2n-1]$ as shown in Figure 1.

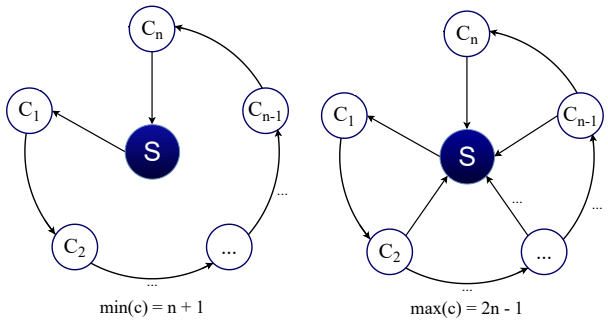


Fig. 1. Minimum and maximum number of communications

IV. PRIVACY AND SECURITY ANALYSIS

This section examines how the proposed protocol responds to data privacy and model security attacks.

A. Privacy

Data privacy attacks exhibit individual characteristics and distinct sets of requirements. Therefore, it is essential to conduct an analysis that can generically address how the protocol would respond to both existing and potential future attacks.

Several attacks against local models tend to focus on specific target clients (TCs), and consequently, on specific

Round 1	Round 2	Round 3	Round 4	Round 5
$S \xrightarrow{L} C_1$	$C_1 \xrightarrow{[\omega_1^t],L} C_2$	$C_2 \xrightarrow{[\omega_1^t],L} C_3$ $C_2 \xrightarrow{[\omega_2^t]} S$	$C_3 \xrightarrow{[\omega_1^t],L} C_4$ $C_3 \xrightarrow{[\omega_3^t]} S$	$C_4 \xrightarrow{[\omega_1^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_3^t],L} C_4$ $C_3 \xrightarrow{[\omega_1^t]} S$	$C_4 \xrightarrow{[\omega_3^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_1^t,\omega_3^t],L} C_4$	$C_4 \xrightarrow{[\omega_1^t,\omega_3^t,\omega_4^t]} S$
		$C_2 \xrightarrow{[\omega_2^t],L} C_3$ $C_2 \xrightarrow{[\omega_1^t]} S$	$C_3 \xrightarrow{[\omega_2^t],L} C_4$ $C_3 \xrightarrow{[\omega_3^t]} S$	$C_4 \xrightarrow{[\omega_2^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_3^t],L} C_4$ $C_3 \xrightarrow{[\omega_2^t]} S$	$C_4 \xrightarrow{[\omega_3^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_2^t,\omega_3^t],L} C_4$	$C_4 \xrightarrow{[\omega_2^t,\omega_3^t,\omega_4^t]} S$
		$C_2 \xrightarrow{[\omega_1^t,\omega_2^t],L} C_3$	$C_3 \xrightarrow{[\omega_2^t,\omega_3^t],L} C_4$ $C_3 \xrightarrow{[\omega_1^t]} S$	$C_4 \xrightarrow{[\omega_2^t,\omega_3^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_1^t,\omega_3^t],L} C_4$ $C_3 \xrightarrow{[\omega_2^t]} S$	$C_4 \xrightarrow{[\omega_1^t,\omega_3^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_1^t,\omega_2^t],L} C_4$ $C_3 \xrightarrow{\omega_3^t} S$	$C_4 \xrightarrow{[\omega_1^t,\omega_2^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_3^t],L} C_4$ $C_3 \xrightarrow{[\omega_1^t,\omega_2^t]} S$	$C_4 \xrightarrow{[\omega_3^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_2^t],L} C_4$ $C_3 \xrightarrow{[\omega_1^t,\omega_3^t]} S$	$C_4 \xrightarrow{[\omega_2^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_1^t],L} C_4$ $C_3 \xrightarrow{[\omega_2^t,\omega_3^t]} S$	$C_4 \xrightarrow{[\omega_1^t,\omega_4^t]} S$
			$C_3 \xrightarrow{[\omega_1^t,\omega_2^t,\omega_3^t],L} C_4$	$C_4 \xrightarrow{[\omega_1^t,\omega_2^t,\omega_3^t,\omega_4^t]} S$

Tabla I
OPERATION. ALL THE POSSIBILITIES FOR THE MODEL EXCHANGE WHEN $n = 4$

target models. However, by disassociating the models from their owners, the connection between the TCs and their models is lost. The design of the protocol prevents S from knowing which clients are the generators of the received models. In addition, the clients are unaware of how many people compose the cycle, so even if they receive the list L, they do not know how many people were in a prior position, preventing them from discovering the generators.

Nevertheless, one of the most significant risks of the protocol lies precisely in reestablishing the link between the model and the client, which would be studied below.

1) Recovering the link: This work outlines two primary strategies that the adversary could follow for reestablishing the link between the TC and their model.

On the one hand, if there is no collusion between the parties, S could apply an inference attack to discover which the model of the TC is. The main advantage of this protocol against these attacks is that even if S knows that a concrete data point belongs to the TC, they would need to perform the attack against all the models and in all the iterations because they cannot know which portion of the data was used in each iteration.

On the other hand, if S colludes with any of the clients they could take advantage of the protocol to link the TC with their local model. Let n be the number of clients and $r < n$ the number of those forming collusion with S. Figure 2 shows all the possible paths that the adversary could take when deciding the order of the cycle for $n = 4$

and $r = 1$. The colour red indicates that the client owning that position is part of the collusion. The members of the collusion are always decision nodes since they have the ability to choose if the next client would be a member or not. However, the non-members are always chance nodes, so the possibility of choosing a member as the next client depends on a probability. The items *TC* and *Path prob.* indicate the position in the cycle that the TC would have in case the adversary follows that path and the probability of following that path, respectively. The column *Success prob.* indicates the likelihood that the TC decides to send their model to a member of the collusion, assuming that the clients will choose any of the options presented in Table I with equal probability. Finally, column *Learning prob.* shows the probability of recovering the link for the TC, supposing that the decision of following a path and sending or not the models to S are independent.

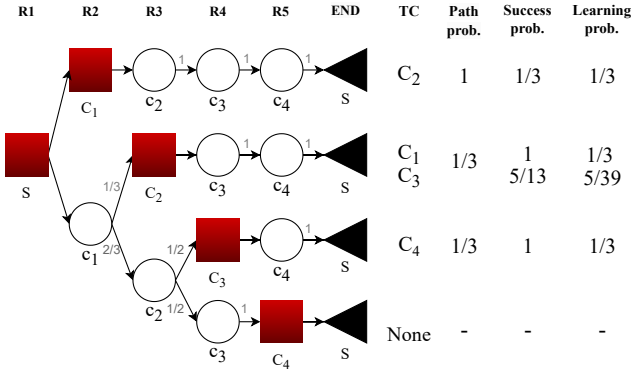


Fig. 2. Analysis of different possibilities for the order of the cycle when $n = 4$ and $r = 1$

As shown in Figure 2, the adversary could study the probabilities and make the correct decisions to follow the best path. Nevertheless, learning which the model of the TC is always depends on a probability.

2) Advantages of the protocol: The effectiveness of the proposed protocol against privacy attacks primarily depends on two factors, whether the attack requires consecutive information about the models and whether information about the target client is necessary for executing the attack. If the attack needs consecutive models, it is very hard to recover the link in multiple iterations. Furthermore, if the attack needs information about TC, the adversary first needs to succeed in recovering the link.

B. Security

Model security attacks could damage the effective operation and prevent the model from converging. The protocol has two main advantages in this respect. S receives the local models in a non-encrypted form, which enables S to apply the model metrics or the robust aggregation they desire. Furthermore, if the adversary aims to change the model of a concrete TC, depending on the number of people forming the collusion and the path they follow, the possibility of changing the model of the TC in more than one iteration depends on a probability. I.e., they could change the local models corresponding to some portions of the TC, but S is likely to receive at least some correct local models of all parties.

To ensure more effectiveness against both privacy and security attacks, it is highly recommended to change the selection of the next client in each iteration.

V. CONCLUSIONS

This work introduces a protocol that facilitates the exchange of models in each iteration of a FL scheme while guaranteeing data privacy and model security. The primary objective is to transmit the local models to S ensuring that each model remains dissociated from its respective generator. The privacy and security analysis shows that the protocol possesses several advantages in mitigating privacy attacks, while also offering flexibility in terms of selecting appropriate security measures.

Several potential steps can be identified as future work within the context of the ongoing investigation, among which the following stand out:

- the development of a more in-depth analysis of the possibilities of success both for privacy and security attacks;
- the exploration of potential strategies to minimize the learning probabilities;
- the integration of the protocol within Blockchain to register the entire process.

ACKNOWLEDGEMENTS

This work was partly supported by Tecnalia and the University of the Basque Country [grant number PIFTEC21/05] and by the Spanish Ministry of Science project AIBioSurv-Tech (Biosurveillance through Artificial Intelligence (AI) in the post-COVID era: Implications for architecture and cybersecurity) TED2021-129975B-C22AEI/10.13039/501100011033/EU/NextGeneration EU/PRTR.

REFERENCES

[1] Lingjuan Lyu et al., "Privacy and Robustness in Federated Learning: Attacks and Defenses", IEEE Transactions on Neural Networks and Learning Systems, pp. 1–21, 2022.
[2] Caiqin Dong et al., "Privacy-Preserving and Byzantine-Robust Federated Learning", IEEE Transactions on Dependable and Secure Computing, pp. 1-16, 2023.
[3] Xu Ma et al., "Privacy-preserving Byzantine-robust federated learning", Computer Standards Interfaces, vol. 30, 2022.
[4] Zhuoran Ma et al., "ShieldFL: Mitigating Model Poisoning Attacks in Privacy-Preserving Federated Learning", IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1639 – 1654, 2022.
[5] Yang Bai et al., "A Method to Improve the Privacy and Security for Federated Learning", 2021 IEEE 6th International Conference on Computer and Communication Systems, ICCCS 2021, 2021.
[6] Xu Ma et al., "Differentially Private Byzantine-Robust Federated Learning", IEEE Transactions on Parallel and Distributed Systems, vol. 33, pp. 3690-3701, 2022.
[7] Najeeb Moharram Jebreel et al., "Enhanced Security and Privacy via Fragmented Federated Learning", IEEE Transactions on Neural Networks and Learning Systems, pp. 1-15, 2022.
[8] Josep Domingo-Ferrer et al., "Secure and Privacy-Preserving Federated Learning via Co-Utility", IEEE Internet of Things Journal, vol. 9, pp. 3988 – 4000, 2022.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Generación sintética de trayectorias mediante aprendizaje profundo con garantías de privacidad diferencial

Victor Rubio Jornet, Javier Parra Arnau, Jordi Forné Muñoz
Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya 08034
victor.rubio.jornet@upc.edu, javier.parra@upc.edu, jordi.forne@upc.edu

Resumen

La generación sintética de trayectorias es crucial para poder realizar estudios y análisis en diferentes campos, como la movilidad urbana, los protocolos de redes móviles, la epidemiología computacional o la simulación de cambios en la movilidad para la planificación urbana. Actualmente, se ha observado que la generación de datos sintéticos sin protección adicional puede poner en riesgo la privacidad de los usuarios, dado que los modelos de aprendizaje profundo actuales son susceptibles a ataques de inferencia de membresía. En este artículo se presenta una actualización de un modelo de aprendizaje profundo ya existente llamado MoveSim, el cual genera datos sintéticos de trayectorias. Este modelo ha sido modificado para ser ϵ-diferencialmente privado y, de esta manera, garantizar la privacidad de los usuarios que han proporcionado sus datos sin perder un rendimiento sustancial del modelo.

***Palabras Clave*—generación de datos sintéticos, trayectorias, privacidad diferencial, aprendizaje profundo, GAN**

I. INTRODUCCIÓN

El uso constante de dispositivos móviles ha permitido que la obtención de datos de trayectorias sea más fácil que nunca. Las trayectorias de usuarios pueden ser de gran utilidad para proyectos de análisis de patrones de movilidad intra-urbana (Wang et al. [1]), para mejorar la conducción de vehículos autónomos (Ferguson et al. [2]), e incluso para evitar contagios durante pandemias a gran escala (Lai et al. [3]). Estos datos son altamente sensibles, ya que pueden dar información sobre la salud, etnia, identificación de género o creencias religiosas, además de poder identificar con relativa facilidad el domicilio o el lugar de trabajo habitual de un usuario. La alta sensibilidad de estos datos evita que, pese a la gran cantidad que se están recogiendo actualmente, se puedan analizar y usar sin añadir un riesgo para la privacidad de los usuarios.

Las trayectorias son un tipo de dato difícil de proteger, ya que son vectores secuenciales, de alta dimensionalidad, con correlación entre los valores y restringidos por el terreno. Por lo tanto, es difícil aplicar técnicas clásicas de enmascaramiento o sustitución de valores sin perder la utilidad de los datos (Miranda et al. [4]).

Para intentar garantizar la privacidad de los datos de trayectoria de usuarios se pueden generar datos sintéticos mediante modelos de aprendizaje profundo (Rao et al. [5], Feng et al. [6]). De esta manera, se puede analizar y compartir con terceras partes una base de datos generada a partir de trayectorias sintéticas con propiedades estadísticas parecidas a los datos reales. Desafortunadamente, este método de securización de bases de datos no es suficiente para garantizar la privacidad de los usuarios cuya información ha sido utilizada para la generación de los datos sintéticos. De esta manera es conveniente añadir una garantía de privacidad a las trayectorias generadas para mitigar el riesgo existente.

II. ESTADO DEL ARTE

En el ámbito de generación de trayectorias sintéticas están ganando peso los modelos basados en aprendizaje profundo. En concreto, los más estudiados en la actualidad son los basados en la arquitectura de redes adversarias generativas (GAN en inglés) (Rao et al. [5], Feng et al. [6]). En esta arquitectura se construyen dos modelos que competirán entre ellos. El generador querrá proporcionar las trayectorias más realistas posibles y el discriminador querrá identificar las trayectorias falsas de las reales. Se entrenarán de manera alternada hasta conseguir el resultado deseado del generador.

Otro enfoque es usar autodecodificadores variacionales (VAE en inglés) (Hueang et al. [7]). Es un método de generación de datos que pretende usar un codificador que reduzca el espacio dimensional de las trayectorias y un decodificador que las devuelva al valor original. Lo que hace que esta arquitectura sea diferente a un codificador-decodificador normal es que entre estos dos se fuerza

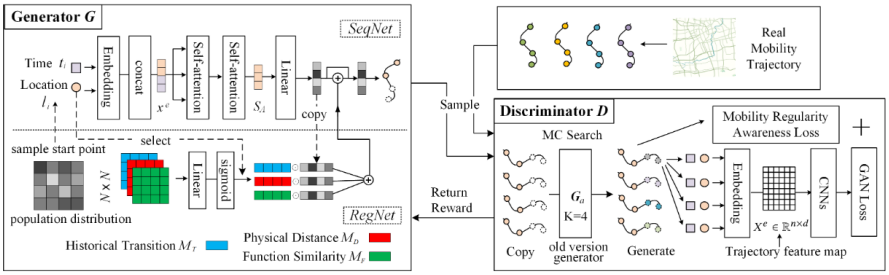


Figura 1. Arquitectura del modelo MoveSim descrita en Feng et al. [6]

que el espacio embebido tenga propiedades especiales. Generalmente, se intenta hacer que tenga propiedades parecidas al ruido gausiano. Una vez entrenado el modelo se puede usar únicamente el decodificador añadiendo ruido gausiano que pertenezca a la dimensión embebida en la entrada y obtendremos datos sintéticos a la salida.

Para ambas soluciones se intenta aprovechar la relación secuencial entre las muestras de trayectorias usando redes basadas en redes neuronales recurrentes (RNN en inglés). Finalmente, se está añadiendo información espacial mediante redes neuronales convolucionales (CNN en inglés) que contengan la topología del terreno e información sobre calles o edificios.

A. *MoveSim*

En este proyecto hemos elegido el modelo MoveSim como base de nuestro trabajo. El modelo de generación de trayectorias sintéticas definido en Feng et al. [6] está basado en una arquitectura de redes adversas (GAN) en la que se define un generador que tiene en cuenta tanto las propiedades del terreno en las que se generan las trayectorias como las propiedades secuenciales de las trayectorias. El modelo ha sido escogido debido a que presenta los mejores resultados en el estado del arte y consigue captar tanto las propiedades secuenciales y periódicas de las trayectorias como obtener información sobre el terreno y la distribución de la población para generar los datos sintéticos.

Como se puede observar en la Fig. 1, el generador está separado en dos redes diferentes, *SeqNet* y *RegNet*. La primera, compuesta por capas de atención y lineales, intenta captar la correlación temporal que existe entre muestras de una trayectoria. La segunda, compuesta por redes convolucionales, prueba de captar los efectos y limitaciones que genera la estructura urbana.

El discriminador está compuesto por redes neuronales convolucionales simples y permite saber qué parte de la trayectoria debe mejorar el generador, ya que se ha aplicado la búsqueda Monte Carlo para saber qué elementos de la trayectoria influyen más en la penalización.

Para realizar el aprendizaje del modelo se ha definido una función de pérdida con regularización de la movilidad, que está dividida en dos partes: L_d, que mide la distancia física, y L_p que mira las variaciones en la periodicidad:

$$L_d = \sum_{i=0}^{n-1} \sqrt{(x_{l_i}, x_{l_{i+1}})^2 + (y_{l_i}, y_{l_{i+1}})^2} \qquad (1)$$

$$L_p = \sum_{i=0}^{n-1} D_I(l_i, l_{i+P}), D_I(l_1, l_2) = \begin{cases} 0, & if\ l_1 = l_2; \\ 1, & if\ l_1 \neq l_2; \end{cases} \qquad (2)$$

La Ec. 1 es la acumulación de la distancia Euclídea de las transiciones en cada trayectoria y la Ec. 2 es la distancia medida con la función característica de las localizaciones con periodicidad fija.

El entrenamiento del modelo se realiza mediante el algoritmo "*reward*"de aprendizaje reforzado [8]. En este caso el resultado del discriminador es usado como un token de premio $R(x)$ en el generador. Para acelerar el entrenamiento se realiza una fase previa de aprendizaje, en la que se actualizan el generador y discriminador en tareas más simples: predecir la siguiente posición de una trayectoria desde el generador y diferenciar entre trayectorias con buenas propiedades espacio-temporales desde el discriminador.

III. NUESTRA PROPUESTA

La generación de datos sintéticos por sí sola no garantiza privacidad, ya que los modelos entrenados para generar datos sintéticos son susceptibles a diferentes ataques.

Los ataques de inferencia de membresía (Hayes et al. [9]) permiten estimar si una muestra pertenece a los datos usados durante el entrenamiento del modelo, poniendo en riesgo la privacidad de los usuarios cuyos datos reales han sido utilizados en el proceso de entrenamiento. Son especialmente vulnerables aquellos modelos que sufren de sobre ajuste (Shokri et al. [10]), y no es necesario tener acceso al modelo para poder realizar este ataque con efectividad. Los ataques de reconstrucción o de inversión de modelo permiten obtener datos de entrada conociendo la salida del modelo y con algún conocimiento parcial de la arquitectura, contenido o características del modelo.

Estas limitaciones no excluyen a los modelos que generan trayectorias sintéticas (Pyrgelis et al. [11]). En este artículo proponemos aplicar la privacidad diferencial (DP) con el objetivo de identificar y reducir el riesgo de que un usuario participe en una base de datos. De manera más concreta, en cualquier escenario con dos bases de datos $\mathcal{D}_1$ y $\mathcal{D}_2$ que difieren en un solo elemento, una función randomizada $\mathcal{K}$ garantiza ϵ-privacidad diferencial según lo descrito en el artículo de C. Dwork [12] si para todo $S \in Rango(\mathcal{K})$ se cumple la Ec. 3:

$$\frac{Pr[\mathcal{K}(\mathcal{D}_1) \in S]}{Pr[\mathcal{K}(\mathcal{D}_2) \in S]} \leq e^{\epsilon} \qquad (3)$$

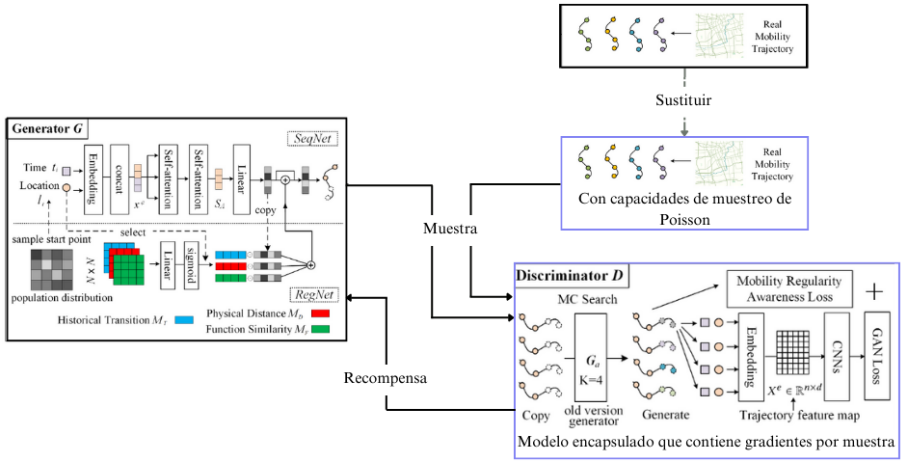


Figura 2. Diagrama de las modificaciones aplicadas en la arquitectura del modelo MoveSim

Aunque existen unas pocas propuestas de generación de datos sintéticos de trayectorias mediante aprendizaje profundo que consideran la privacidad, ninguna de ellas ofrece garantías de privacidad diferencial [4]. Por eso, el ámbito de este proyecto es usar el modelo de generación de trayectorias sintéticas definido en Feng et al. [6] y modificarlo para satisfacer estas garantías. El principal objetivo de este trabajo es estudiar si un buen compromiso entre utilidad y privacidad es alcanzable.

Los principales mecanismos para añadir garantías de privacidad diferencial a modelos generativos de aprendizaje profundo se recogen en tres grandes tipos: los que añaden ruido a la función a optimizar [13], a los pesos del modelo [14], o a los datos sintéticos [15]. Añadir ruido a los datos sintéticos limita el número de llamadas que se pueden hacer al modelo, ya que los datos se degradan con cada iteración que hacemos. Los mecanismos que pretenden añadir ruido a la función a optimizar son costosos y dificultan que se pueda encontrar un mínimo en la función durante el tiempo de entrenamiento del modelo. En este caso hemos optado por añadir ruido en los pesos del modelo para garantizar privacidad diferencial. De esta manera, se modifica el algoritmo de actualización de los pesos, *Stochastic Gradient Descent* (SGD), para (1) usar gradientes por muestra en vez de los gradientes de todo el grupo o mini-grupo, (2) recortar los gradientes para que estén acotados y (3) añadir ruido a los gradientes.

Para añadir las garantías de privacidad diferencial se usa la librería Opacus de Python [16] que permite definir un presupuesto de seguridad ϵ y ceñir el entrenamiento del modelo a ese presupuesto. La librería modifica principalmente tres elementos: el modelo, el optimizador y el cargador de datos. El modelo se cambia para poder mantener un registro de los gradientes por muestra en vez de los gradientes por grupos, el optimizador permite gestionar el recorte de los gradientes y añadir el ruido durante el entrenamiento, y el cargador de datos se modifica para que permita hacer muestreo de Poisson para poder añadir cada muestra a un grupo de manera independiente al resto.

El entrenamiento de una GAN se puede ver como el post-procesado de unos datos que resultan del entrenamiento del discriminador, y la propiedad descrita en Dwork et al. [17] prueba que el post-procesado de datos ϵ-DP no reduce sus garantías de privacidad. Por eso hemos aplicado los cambios únicamente al proceso de entrenamiento del discriminador del modelo MoveSim Fig. 2. Adicionalmente, se ha sobreescrito el cargador de datos por defecto, para tener más autonomía en el proceso de aprendizaje.

IV. RESULTADOS PRELIMINARES

Al ser un trabajo en progreso, no se han podido obtener todos los resultados. Pese a esto, la comparativa que se pretende realizar a nivel de utilidad se mide mediante cuatro métricas, que representan aspectos importantes a tener en cuenta cuando se trabaja con trayectorias:

1. Distancia acumulada por usuario en un intervalo de tiempo fijado.
2. Radio de giro de la trayectoria, que representa el rango del movimiento por usuario.
3. Duración de cada punto por visita.
4. I-Rank, la distribución de la frecuencia de visita de los 100 lugares mas visitados.

Para obtener el resultado numérico se utiliza la divergencia de Jensen-Shannon para comparar las distribuciones de las cuatro métricas mencionadas anteriormente. De esta manera, cuanto más cercano a 0 sea el valor resultante más se parecerán las distribuciones de ambos conjuntos de datos y mejor será el modelo. Se realiza una muestra base en la que se entrena el modelo sin garantías de privacidad diferencial y posteriormente se realizan diversos entrenamientos para diferentes ϵ. De esta manera queremos comprobar cómo cambia la utilidad cuando se incrementa la garantía de privacidad, es decir, cuando se disminuye la ϵ.

Los resultados preliminares quedan representados en la Tabla I, en la que se aprecian las cuatro métricas para cada

Tabla I
RESULTADOS PRELIMINARES, EN NEGRITA LOS MEJORES

Epsilon	Distancia acumulada por usuario	Radio de giro	Duración de cada visita en los puntos	I-Rank
Base	0.0178	0.0663	0.0212	0.0243
20	0.0059	0.0234	0.0261	0.0230
10	0.0059	0.0245	0.0195	0.0211
5	**0.0049**	**0.0223**	**0.0165**	**0.0190**

una de las ϵ medidas hasta la fecha: 20, 10 y 5 junto con el resultado base. Se han realizado 2 medidas para cada una de las ϵ y se ha calculado la media.

Podemos apreciar en la Tabla I que al disminuir la ϵ los resultados mejoran, hasta el punto en que $\epsilon = 5$ obtiene los mejores resultados hasta el momento. También se puede observar que a medida que aumentamos la ϵ los valores convergen a los resultados del modelo base, sin Privacidad Diferencial. Con los resultados preliminares anteriores podemos aventurarnos a decir que añadir ruido de manera controlada en los pesos durante el entrenamiento ayuda al modelo a generalizar mejor, y puede ser una medida más para evitar el sobre ajuste.

V. Conclusiones

Las trayectorias son un tipo de dato crucial en la actualidad, pero su uso plantea riesgos para la privacidad. Si se filtrasen estos datos, se podría identificar usuarios y revelar información sensible sobre ellos. Generar datos sintéticos no es ninguna garantía de privacidad. Los modelos de generación de datos son susceptibles a diversos ataques que ponen en riesgo la privacidad de los usuarios. Por eso es necesario recurrir a la privacidad diferencial.

Este proyecto está en progreso por lo que aún queda trabajo por hacer. Primero obtener más resultados para diferentes ϵ sobre la utilidad del modelo, y segundo evaluar las garantías de privacidad de manera empírica. Queremos obtener resultados realizando ataques de inferencia al modelo para saber hasta qué punto añadir garantías de privacidad diferencial mejora la privacidad real de los usuarios.

VI. Agradecimientos

Javier Parra-Arnau es el beneficiario de una beca Ramón y Cajal (ref. RYC2021-034256-I) financiada por el Ministerio de Ciencia e Innovación de España y la Unión Europea - ”NextGenerationEU”/PRTR (Plan de Recuperación, Transformación y Resiliencia). Este trabajo también ha sido apoyado por el Gobierno español a través del proyecto ”Enhancing Communication Protocols with Machine Learning while Protecting Sensitive Data (COMPROMISE” PID2020-113795RB-C31, financiado por MCIN/AEI/10.13039/501100011033, y a través del proyecto ”MOBILYTICS” (TED2021-129782B-I00), financiado por MCIN/AEI/10.13039/501100011033 y la Unión Europea ”NextGenerationEU”/PRTR y financiado por la Generalitat de Catalunya mediante la beca AGAUR ”2021 SGR 01413”.

Referencias

[1] Zuchao Wang et al. “Visual Traffic Jam Analysis Based on Trajectory Data”. En: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), págs. 2159-2168. DOI: 10.1109/TVCG.2013.228.

[2] Dave Ferguson et al. “Detection, prediction, and avoidance of dynamic obstacles in urban environments”. En: jul. de 2008, págs. 1149-1154. ISBN: 978-1-4244-2568-6. DOI: 10.1109/IVS.2008.4621214.

[3] Shengjie Lai et al. “Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak in China”. En: *medRxiv* (2020). DOI: 10.1101/2020.03.03.20029843.

[4] Àlex Miranda Pascual et al. “SoK: differentially private publication of trajectory data.” En: *Proceedings on Privacy Enhancing Technologies*. Vol. 2023. Abr. de 2023, págs. 496-516. DOI: 10.56553/popets-2023-0065.

[5] Jinmeng Rao et al. *LSTM-TrajGAN: A Deep Learning Approach to Trajectory Privacy Protection*. 2020. arXiv: 2006.10521 [cs.LG].

[6] Jie Feng et al. “Learning to Simulate Human Mobility”. En: New York, NY, USA: Association for Computing Machinery, 2020, págs. 3426-3433. ISBN: 9781450379984. DOI: 10.1145/3394486.3412862.

[7] Dou Huang et al. “A Variational Autoencoder Based Generative Model of Urban Human Mobility”. En: *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. 2019, págs. 425-430. DOI: 10.1109/MIPR.2019.00086.

[8] Richard Sutton y Andrew Barto. “Reinforcement Learning: An Introduction”. En: *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* 9 (feb. de 1998), pág. 1054. DOI: 10.1109/TNN.1998.712192.

[9] Jamie Hayes et al. “: Membership Inference Attacks Against Generative Models”. En: 2018.

[10] Reza Shokri et al. *Membership Inference Attacks against Machine Learning Models*. 2017. arXiv: 1610.05820 [cs.CR].

[11] Apostolos Pyrgelis, Carmela Troncoso y Emiliano De Cristofaro. *Knock Knock, Who’s There? Membership Inference on Aggregate Location Data*. 2017. arXiv: 1708.06145 [cs.CR].

[12] Cynthia Dwork. “Differential Privacy”. En: *Proceedings of the 33rd Intl. Colloquium on Automata* 2 (2006), págs. 1-12. DOI: 10.1007/11787006_1.

[13] Jun Zhang et al. *Functional Mechanism: Regression Analysis under Differential Privacy*. 2012. arXiv: 1208.0219 [cs.DB].

[14] Martin Abadi et al. “Deep Learning with Differential Privacy”. En: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, oct. de 2016. DOI: 10.1145/2976749.2978318.

[15] Shadi Rahimian, Tribhuvanesh Orekondy y Mario Fritz. *Sampling Attacks: Amplification of Membership Inference Attacks by Repeated Queries*. 2020. arXiv: 2009.00395 [cs.CR].

[16] Ashkan Yousefpour et al. “Opacus: User-Friendly Differential Privacy Library in PyTorch”. En: *arXiv preprint arXiv:2109.12298* (2021).

[17] Cynthia Dwork y Aaron Roth. “The Algorithmic Foundations of Differential Privacy”. En: 9.3–4 (ago. de 2014), págs. 211-407. ISSN: 1551-305X. DOI: 10.1561/0400000042.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Correlación de Threat Actors según técnicas antianálisis en muestras de malware

Sebastien Kanj Bongard, Abraham Pasamar, Oriol Rosés, Josep Pegueroles
Departament d'Enginyeria Telemàtica,
Universitat Politècnica de Catalunya
Carrer de Jordi Girona, 1, 08034 Barcelona
sebastien.kanj@upc.edu, apasamar@incide.es, oroses@incide.es, josep.pegueroles@upc.edu

Las muestras de malware (software malicioso) son uno de los Indicadores de Compromiso (IOC) más usados para la atribución de incidentes de ciberseguridad a los distintos actores (Threat Actors, TA) en activo. Es de interés de los atacantes tratar de demorar y dificultar al máximo el trabajo de los analistas de malware y para ello implementan diversas técnicas de evasión o antianálisis. Los desarrolladores de malware venden sus muestras en foros o markets de la Dark Web a distintos perfiles de atacantes y esto genera relaciones entre los distintos grupos. En este trabajo se presenta el estudio en curso, dentro del doctorado industrial realizado con INCIDE, de correlación de las técnicas evasivas presentes en muestras de malware de Threat Actors para obtener un mapa geopolítico de sinergias entre ellos y así perfilar mejor su actividad y mejorar las capacidades de atribución.

***Palabras Clave*—malware, evasión, Threat Actors, antianálisis, geopolítico**

I. INTRODUCCIÓN

Uno de los trabajos en curso del doctorado industrial realizado entre INCIDE Digital Data S.L. y el departamento de ENTEL de la UPC, llevado a cabo por los autores de este artículo, consiste en relacionar en un estudio geopolítico las sinergias entre TA mediante la identificación y el análisis de técnicas evasivas en las muestras de malware usadas. Con ello se pretende perfilar y caracterizar los procedimiento de actuación de los TA y mejorar las capacidades de atribución.

El uso de malware por parte de los distintos Threat Actors, en adelante TA, es una de las mayores preocupaciones de los profesionales de la ciberseguridad. Su uso se incrementa año tras año, al igual que su complejidad y su especialización [1]. Estas mejoras son causa de un incremento del poder adquisitivo de los desarrolladores, que venden su producto a medida y customizado en la Dark Web. Sus compradores, distintos TA especializados, entre otros, en ataques de tipo ransomware, Initial Access Brokers [12], minería de criptomonedas o robo de credenciales e información, compran las muestras en mercados y foros de la Dark Web, generando un mercado clandestino que mueve día tras día mayores cantidades de dinero y de profesionales [2]. Es sabido que algunos grupos de Ransomware As A Service (RaaS) ya disponen de desarrolladores propios que mantienen y mejoran las herramientas usadas y que permiten a los dirigentes de estas bandas ofertar un mejor servicio a sus afiliados [3].

Para proteger las infraestructuras los responsables de ciberseguridad tratan de bloquear y erradicar estas muestras una vez llegan a los equipos de sus organizaciones. Para su detección se confía en los sistemas de EndPoint Detection and Response (EDR) [4] y, en algunos casos, en los ya descatalogados AntiVirus (AV). En algunos casos, cuando estos sistemas fallan o el atacante logra evadirlos es el analista forense o el incident responder el que identifica la muestra mediante el análisis forense y el que realiza su estudio. Normalmente el análisis de esta muestra se realiza con el uso de una SandBox, que automatiza el proceso y reduce el tiempo de investigación. Tan solo algunos especialistas del sector, los malware analyst, se encargan de analizar al detalle estas muestras, permitiendo caracterizar su propósito y su comportamiento. Es tras este análisis que se generan reglas de detección como las YARA [5] rules, que implementadas en los EDR permiten mejorar la capacidad de detección de estos sistemas. Este hecho pone en peligro el negocio de los TA por lo que existe un gran interés en dificultar y demorar el análisis de estas muestras, implementando los desarrolladores de malware una gran variedad de técnicas conocidas como antianálisis o evasión [6]. En este documento se introducen de manera resumida las técnicas más usadas y se presenta el interés de su estudio para nuestra investigación.

II. TÉCNICAS ANTIANÁLISIS

Existen varios proyectos de la comunidad de ciberseguridad que tratan de aglutinar y definir las técnicas

antianálisis más presentes en las muestras de malware usadas por los TA [7][8]. A continuación, se resumen los grandes grupos existentes.

A. Anti-VM y Anti-SandBox

Estas técnicas tienen como objetivo evadir el análisis de las muestras, variando su comportamiento o deteniendo su ejecución cuando detecten que están siendo ejecutadas en un entorno virtualizado [9]. Las técnicas más frecuentes para esta evasión se pueden categorizar en:

- Ficheros y directorios en el sistema: el malware comprueba si existen varios ficheros en el sistema pertenecientes a los sistemas de virtualización más habituales (VMware, VirtualBox, Parallels).
- Artefactos en el sistema: En este caso, el malware revisa si existen artefactos relacionados con la virtualización tales como claves de registro o procesos.
- Características del equipo: los desarrolladores consultan características del sistema tales como el nombre del equipo, los usuarios, la resolución del monitor, la fecha de instalación del sistema operativo, firmware... para identificar equipos con parámetros por defecto o genéricos.
- Hardware del equipo: debido a la virtualización de los componentes del equipo, el malware puede consultar información del hardware para identificar patrones relacionados con esta virtualización. Entre ellos se encuentran los discos duros, los dispositivos de audio, los monitores, la memoria RAM, la CPU o los dispositivos de red.
- Interacción del usuario: dado que los equipos virtualizados no suelen tener un historial de actividad y que la mayor parte de las SandBox no permiten interaccionar con la muestra a analizar, el malware puede sacar provecho, buscando actividad anterior de ficheros, navegación web, software alternativo... o solicitando una interacción para la ejecución. El seguimiento del tiempo de ejecución también permite identificar si la ejecución está siendo controlada.

B. Anti-Debugging

Estas técnicas son usadas por los atacantes para identificar cuándo un binario está siendo analizado mediante técnicas de Debugging, modificando su comportamiento en caso satisfactorio para no poder ser caracterizado. Las técnicas más frecuentes empleadas por los atacantes para evitar este tipo de análisis son las siguientes [10]:

- Uso de funciones de la API de Windows: existen varias funciones nativas de Windows diseñadas para identificar si un programa está siendo depurado. Algunas de ellas son *UsDebuggerPresent*, *CheckRemoteDebuggerPresent*, *NtQueryInformationProcess* o *OutputDebugString*.
- Comprobación de flags del *Process Environment Block* (PBE): la estructura PBE presente en la familia de los sistemas operativos Windows NT, una estructura de datos interna que almacena información sobre un proceso en ejecución, presenta varios indicadores que, en caso de realizarse el debug de un binario, son alterados. Los más usados son *BeingDebugged*, *ProcessHeap* (ForceFlags y Flag) o *NTGlobalFlag*.
- Revisión de los registros: al usar un software de debug, al igual que con la mayoría de software ejecutado en un equipo, se realizan modificaciones en las claves de registro de los sistemas operativos Windows. La comprobación de claves típicas de dubuggers habituales puede permitir identificar su uso.
- Buscando comportamientos de la depuración: cuando se realiza una depuración de un binario, el comportamiento normal de una ejecución es alterado. Estas alteraciones derivadas del análisis pueden ser identificadas mediante la búsqueda de instrucciones INT 3 (causa del uso de breakpoints), calculando el *hash* del propio software o haciendo comparaciones de timestamps o número de instrucciones ejecutadas.

C. Anti-DBI

La instrumentalización dinámica de binarios (DBI) permite orquestar ejecuciones incrustando código en la muestra original, alterando su comportamiento y ejecución. Este tipo de análisis permite recrear condiciones o situaciones necesarias para que el malware continúe su ejecución a interés del analista. Para evitar este tipo de análisis, los atacantes pueden usar, entre otras, estas técnicas:

- Limitaciones funcionales y de recursos: dado que para la instrumentalización del binario se debe emular y virtualizar el entorno y sus recursos, existen varias limitaciones como instrucciones no soportadas, comportamientos incompatibles, limitaciones de procesador, memoria o tiempo.
- Artefactos del entorno: a causa de la instrumentalización, el sistema donde se ejecuta la muestra presenta varios artefactos en la memoria o el disco que, consultados por el malware, pueden permitir identificar su análisis.
- Detección de compiladores JIT: la instrumentalización de binarios usa compiladores *Just-In-Time* (JIT), que dejan trazas en el sistema, las cuales son aprovechadas por el software para identificar su análisis.

En este caso, ya existe una taxonomía de las técnicas realizada en 2022 [11], por lo que nuestro trabajo se está centrando en el desarrollo de reglas de detección.

D. Anti-Disassembly

El análisis por disassembly de malware se utiliza para examinar el código fuente en lenguaje ensamblador de un programa malicioso con el fin de comprender su funcionamiento. Para evitar la correcta traducción de código a ensamblador, los atacantes utilizan diversas técnicas [10], entre las cuales se encuentran las siguientes:

- Instrucciones de salto con *rogue byte*: el desarollador del malware puede usar una instrucción de salto con una condición constante. La rama que nunca se ejecutará contiene un *rogue byte* que al ser interpretado por

el motor de disassembly provoca que todo el siguiente texto no pueda ser interpretado.
- Byte compartido entre instrucciones: en este caso el atacante usa un mismo byte para dos instrucciones distintas, imposibilitando la interpretación por parte de los motores de disassembly.

III. TRABAJOS EN CURSO

Tal y como se ha introducido, el análisis de estas técnicas permite crear reglas de detección YARA. Estas agilizan el análisis de las muestras identificadas durante una respuesta ante un incidente, permitiendo detectar la implementación de las técnicas en el malware.

A continuación se presentan las dos líneas de trabajo que se están llevando a cabo en el proyecto de doctorado industrial INCIDE-ENTEL UPC.

A. Taxonomía de técnicas y sus reglas

Con la intención de categorizar, documentar y generar reglas de detección, se está realizando una taxonomía de las técnicas de evasión y antianálisis conocidas por la comunidad. Actualmente, existen varios proyectos colaborativos que tratan de aglutinar el máximo de técnicas de antianálisis identificadas en muestras de malware. Pese a ello, su categorización y documentación no suele ser completa y el formato de presentación depende de cada fuente en la que se encuentre documentada. De manera adicional, se está realizando una revisión de las reglas YARA existentes [15] que permiten detectar el uso de las distintas técnicas identificadas y, en caso de no existir, se desarrolla dicha regla de detección.

El objetivo de este trabajo es la creación de una documentación estandarizada con todas las técnicas conocidas por la comunidad junto con su regla de detección. Para testear su funcionamiento, se están usando muestras reales presentes en bases de datos y repositorios a los cuales se ha obtenido acceso por estudio académico.

Por el momento, se han identificado reglas YARA para algunas de las técnicas anti-debugging, anti-VM y Anti-SandBox, siendo las técnicas más explotadas por los atacantes y más conocidas por la comunidad. De la primera búsqueda, no se han identificado proyectos existentes dedicados a la creación de reglas YARA para técnicas anti-DBI ni anti-Disassembly, lo cual se plantea como un trabajo a realizar durante nuestra investigación.

B. Correlación de TA

La identificación de las técnicas de evasión y el desarrollo de sus reglas de detección es un primer paso para el objetivo final. Dicho objetivo consiste en tratar de relacionar los distintos desarrolladores de malware de las principales familias con los TA que hacen uso de sus herramientas. Esta relación se realizará mediante la automatización del escaneo de muestras actuales atribuidas a campañas de malware de varias familias y a TA de ransomware, robo de credenciales e información o criptomineros. Este escaneo generará una relación entre técnicas de evasión, familias de malware y muestras usadas por los atacantes, lo que puede llegar a permitir identificar relaciones comerciales entre varios grupos, una información de interés a nivel de Cyber Threat Intelligence.

Es sabido que varios TA tienen relaciones con los gobiernos de distintos países, siendo subvencionados o protegidos por estos [13] [14]. La relación generada a partir del análisis de las muestras que usan podría dar lugar a identificar posibles relaciones entre los distintos TA, tales como muestras compartidas, código común o mismos desarrolladores.

Un ejemplo de este tipo de situaciones se ha dado durante este año, en el que se ha identificado que varios actores de ransomware han reusado código del grupo de atacantes *Babuk*, el cual fue filtrado en junio de 2021 [16]. Se sospecha que hasta diez grupos de atacantes habrían reusado el código publicado para desarrollar sus propias herramientas [17]. Nuestro proyecto en curso permitiría identificar automáticamente actividades similares y usará este tipo de ejemplo para testear la herramienta en desarrollo.

IV. CONCLUSIONES

La profesionalización y especialización de los distintos perfiles de atacantes y el incremento de mercado negro en la Dark Web está implicando una mayor relación entre los distintos perfiles. Identificar estas relaciones es una rama investigativa de la Cyber Threat Intelligence. Con nuestros trabajos pretendemos comprobar si el uso de mecanismos evasivos en las muestras es suficientemente único y particular como para permitir relacionar los desarrolladores de malware con los distintos Threat Actors.

AGRADECIMIENTOS

Este trabajo ha sido posible gracias a la financiación del Ministerio de Ciencia y Educación con el proyecto: TCO-RISEBLOCK (PID2019-110224RB-I00) y de la Generalitat de Catalunya con la subvención 2021-SGR-00594 y 2021 DI 92. Así como la investigación previa realizada por la empresa INCIDE.

REFERENCIAS

[1] M. N. Alenezi, H. Alabdulrazzaq, A. A. Alshaher, and M. M. Alkharang, "Evolution of malware threats and techniques: A review," *International Journal of Communication Networks and Information Security*, vol. 12, no. 3, pp. 326-337, 2020. Kohat University of Science and Technology (KUST).

[2] J. Lusthaus, "Beneath the dark web: Excavating the layers of cybercrime's underground economy," in *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2019, pp. 474-480.

[3] P. H. Meland, Y. F. F. Bayoumy, and G. Sindre, "The Ransomware-as-a-Service economy within the darknet," *Computers & Security*, vol. 92, p. 101762, 2020. Elsevier.

[4] A. Arfeen, S. Ahmed, M. A. Khan, and S. F. A. Jafri, "Endpoint Detection & Response: A Malware Identification Solution," in *2021 International Conference on Cyber Warfare and Security (ICCWS)*, IEEE, 2021, pp. 1-8.

[5] A. Lockett, "Assessing the effectiveness of YARA rules for signature-based malware detection and classification," *arXiv preprint arXiv:2111.13910*, 2021.

[6] Y. Gao, Z. Lu, and Y. Luo, "Survey on malware anti-analysis," in *Fifth International Conference on Intelligent Control and Information Processing*, 2014, pp. 270-275.

[7] Check Point Research, "Evasion Techniques," [Online]. Available: https://evasions.checkpoint.com/. [Accessed: May 31, 2023].

[8] T. Roccia and J.-P. Lesueur, ”Unprotect Project,” [Online]. Available: https://unprotect.it/. [Accessed: May 31, 2023].
[9] A. Issa, ”Anti-virtual machines and emulations,” *J Comput Virol*, vol. 8, pp. 141-149, 2012.
[10] R. R. Branco, G. N. Barbosa, and P. D. Neto, ”Scientific but not academical overview of malware anti-debugging, anti-disassembly and anti-vm technologies,” *Black Hat*, vol. 1, no. 2012, pp. 1-27, 2012.
[11] A. S. Filho, R. J. Rodríguez, and E. L. Feitosa, ”Evasion and countermeasures techniques to detect dynamic binary instrumentation frameworks,” *Digital Threats: Research and Practice (DTRAP)*, vol. 3, no. 2, pp. 1-28, 2022. ACM New York, NY.
[12] S. Tatar, ”Initial Access Brokers,” [Online]. Available: https://arcticwolf.com/resources/glossary/what-are-initial-access-brokers/. [Accessed: June 15, 2023].
[13] Z. Bederrna, ”Financial Perspective Thought Experiment on Russian Cyber Threat Actors” in *International Journal of Economics and Finance*, 2023.
[14] H. Durojaye, R. Oluwaukola, ”Impact of State and State Sponsored Actors on the Cyber Environment and the Future of Critical Infrastructure” in *arXiv preprint arXiv:2212.08036*, 2022.
[15] Yara Rules, ”Yara rules project Github,” [Online]. Available: https://github.com/Yara-Rules/rules/blob/master/antidebug_antivm/antidebug_antivm.yar. [Accessed: September 21, 2023].
[16] Catalin Cimpanu, ”Builder for Babuk Locker ransomware leaked online,” [Online]. Available: https://therecord.media/builder-for-babuk-locker-ransomware-leaked-online. [Accessed: September 21, 2023].
[17] Alex Delamotte, ”Hypervisor Ransomware — Multiple Threat Actor Groups Hop on Leaked Babuk Code to Build ESXi Lockers,” [Online]. Available: https://shorturl.at/fpBF9. [Accessed: September 21, 2023].

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Análisis forense de la herramienta de tunelado de puertos Ngrok

Sebastien Kanj Bongard, Carlos Navarro, Abraham Pasamar, Oriol Rosés, Josep Pegueroles
Departament d'Enginyeria Telemàtica,
Universitat Politècnica de Catalunya
Carrer de Jordi Girona, 1, 08034 Barcelona
sebastien.kanj@upc.edu, cnavaro@incide.es, apasamar@incide.es, oroses@incide.es, josep.pegueroles@upc.edu

Las herramientas de tunelado, como Ngrok, permiten publicar a Internet puertos de interés camuflados en protocolos frecuentes y habitualmente permitidos en infraestructuras corporativas. Usando servidores *proxy* se evita realizar acciones en sistemas de red perimetrales y configuraciones adicionales en equipos finales. Esta funcionalidad es aprovechada por Threat Actors de *ransomware* para usar protocolos de control remoto, como RDP, o protocolos de exfiltración, como FTP, en sus ataques, evadiendo las políticas de seguridad de red. En este articulo se realiza una investigación de los rastros que esta técnica evasiva deja en los equipos a fin de mejorar el análisis forense de incidentes de *ransomware*, proponiendo mecanismos de identificación y subrayando los artefactos útiles para su análisis.

***Palabras Clave*—ngrok, dfir, ransomware, tunelado, puertos**

I. INTRODUCCIÓN

Durante el último año, el equipo de Respuesta a Incidentes de INCIDE Digital Data S.L. ha participado en varios casos de *ransomware* en empresas españolas en los que se ha identificado el uso de herramientas de tunelado de puertos. Estas herramientas suelen ser usadas por los equipos de *Information Technology* (IT) para exponer puertos de distintos servicios a Internet mediante el uso de servidores que actúan como *Proxys*. Existen multitud de herramientas comerciales que permiten exponer los puertos de manera sencilla [1] con tan solo la ejecución de su binario y el uso de un fichero de configuración o una serie de comandos.

Estos últimos meses varios equipos de respuesta ante incidentes han observado [2][3] como grupos de atacantes de ransomware han usado herramientas de este tipo para exponer de manera pública el puerto del servicio de Windows *Remote Desktop Protocol* (RDP) [7] de servidores de la organización atacada. Esta técnica suele ser ejecutada en equipos que no exponen puertos a Internet y que han sido accedidos por el atacante durante la fase de reconocimiento y movimiento lateral [4]. Concretamente, desde INCIDE se ha evidenciado el uso de la herramienta Ngrok [8] por parte del atacante Sparta [5] y de RPort [9] por el atacante LockBit 3.0 [6]. Otras fuentes de Cyber Threat Intelligence confirman que estas herramientas han sido vistas por estos atacantes además de otros Threat Actors (TA) como ALPHV/BlackCat [10], Cobalt Mirage [11] o Fox Kitten [12]. Las herramientas no solo se limitan a Ngrok o RPort, sino que también han sido identificadas otras como por ejemplo Serveo [13], Fast Reverse Proxy (FRP) [14] o Pagekite [15].

En este artículo se realiza una revisión del uso de las herramientas Ngrok, recreando la actividad de un TA en el laboratorio de INCIDE y analizando su comportamiento desde un punto de vista forense. En la sección II se detalla el funcionamiento de Ngrok. A continuación, en la sección III se presenta el escenario recreado en el laboratorio con el detalle de las acciones realizadas y herramientas usadas. Seguidamente, en la sección IV se presentan los resultados obtenidos del análisis forense. El artículo finaliza con una discusión de los resultados en la sección V y las conclusiones alcanzadas en la sección VI.

II. FUNCIONAMIENTO Y USO DE NGROK

Las herramientas de tunelado de Proxy inverso se basan en un concepto muy sencillo. El software de tunelado usará el protocolo HTTPS, permitido y muy frecuente en la mayoría de las infraestructuras corporativas, y establecerá una comunicación cifrada con un servidor, que puede pertenecer a la empresa propietaria del software. Una vez establecida esta comunicación, el software redirigirá el tráfico que reciba al puerto indicado durante la configuración del servicio. De este modo, el puerto indicado, en el caso de la mayor parte de los atacantes RDP, 3389 por defecto, quedará accesible desde Internet. En la figura 1 se muestra un esquema resumen del funcionamiento de la herramienta de Ngrok.

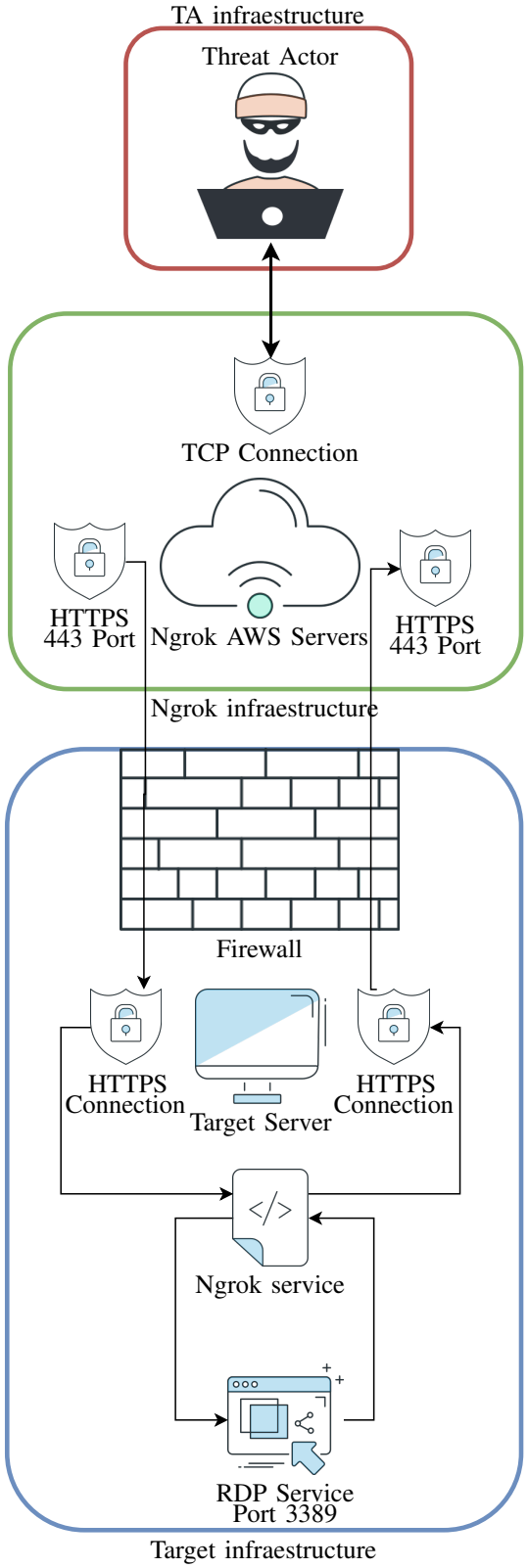


Fig. 1. Ngrok

Esquemáticamente, las comunicaciones que se establecen:

1) Al ejecutar y configurar el software de Ngrok en el equipo objetivo, especificando que el puerto a exponer es el de RDP (por defecto 3389), el software nos facilitará una URL y un puerto, del tipo
   2.tcp.eu.ngrok.io:12345
2) El servicio de Ngrok creará una comunicación HTTPS con el protocolo TCP al puerto 443 de un servidor de AWS propiedad de Ngrok.
3) Para realizar la sesión remota al equipo objetivo, creamos una sesión de RDP desde nuestro equipo al servidor de Ngrok, con destino
   2.tcp.eu.ngrok.io
   y puerto
   12345
4) El servidor de Ngrok redirigirá el tráfico al equipo objetivo mediante la conexión HTTPS establecida durante la iniciación.
5) De este modo, tenemos una sesión RDP tunelada mediante el servidor de AWS, que actúa como *proxy*, al servicio de Ngrok que corre en el equipo objetivo.

III. ESCENARIO

Con el objetivo de simular la actividad habitual de un atacante para posteriormente adquirir los artefactos forenses del equipo se ha usado una máquina virtual limpia con sistema operativo Windows 11 y conexión a Internet, donde fue ejecutado el software de Ngrok y posteriormente usado para realizar una sesión de RDP remota. A continuación se detallan los distintos pasos que se siguieron:

1) Copia del binario de instalación en el equipo.
2) Ejecución del binario con token facilitado en el portal de nuestro usuario en la web de Ngrok:
   ngrok config add-authtoken <token>
   Este comando generó un fichero de configuración en el directorio:
   c:\Usrs\<Usr>\AppData\local\ngrok\ngrok.yml
3) Ejecución del comando de Ngrok para especificar de qué puerto y mediante qué protocolo queremos tunelizar el tráfico. En nuestro caso:
   ngrok tcp 3389
   Una vez ejecutado este comando se inició el tunelado, mostrando por consola las especificaciones del servidor
   tcp://4.tcp.eu.ngrok.io:16745 ->localhost:3389
4) Una vez finalizada la configuración del tunelado, se procede a realizar una sesión RDP desde otra máquina a la dirección y puerto facilitada por consola. Se aprovechó esta sesión para navegar por diversos ficheros y subir y descargar muestras directamente desde la sesión de RDP.

Para la obtención de evidencias forenses se han usado los siguientes sistemas de monitorización y adquisición:

- Sysmon: Para obtener más verbosidad en los eventos de Windows almacenados en los ficheros EVTX del sistema, se instaló la herramienta de Sysmon[16] del paquete de SysInternals de Microsoft.
- Wireshark: Con el objetivo de caracterizar el tráfico generado por el uso del software se ha realizado una

captura del tráfico de red generado por el equipo durante los momentos de la simulación con el software de Wireshark [17].
- GRR: Para realizar una adquisición de los artefactos forenses del equipo para su posterior análisis se ha instalado el agente forense de GRR [18].

IV. ANÁLISIS FORENSE

A continuación se muestran los resultados del análisis de los artefactos forenses del equipo y de la captura del tráfico de red.

A. Análisis artefactos forenses del equipo

El análisis de los artefactos forenses permite caracterizar las distintas fases descritas en el apartado III. El detalle del rastro forense que dejan las distintas acciones es:

1) Copia del binario: Al copiar el ejecutable, al igual que otros ficheros, podemos observar su creación en la base de datos MFT [20] de Windows, observando una entrada de creación con el flag de "*birth (b)*" , y en los eventos de Sysmon, concretamente en una entrada de creación de fichero con Event ID 11.
2) Configuración de token de Ngrok: La primera ejecución de Ngrok en la que se indica el token del usuario puede identificarse mediante varios artefactos forenses. En los eventos de Windows generados por Sysmon se observan los eventos con ID 1 y 5, que nos indican el inicio de un proceso tras la ejecución de un comando por línea de comandos y su posterior finalización. Además estos eventos nos indican el Hash del binario de Ngrok así como su ubicación. La ejecución del binario también puede ser evidenciada en el artefacto forense del *Prefetch* [21], el cual nos muestra como primera entrada la fecha de ejecución del comando. Dado que esta acción crea un fichero de configuración del software, podemos identificar en la MFT la creación del documento:

   C:\Users\User\AppData\Local\ngrok\ngrok.yml
3) Tunelización del puerto 3389 por TCP: Tras la ejecución del comando de tunelización del puerto 3389, puerto por defecto del protocolo *Remote Desktop Protocol*, se identifican varias acciones en los artefactos forenses del equipo. Por un lado se vuelve a observar una ejecución del binario de Ngrok en el *Prefetch* del equipo, actualizando la última hora de ejecución. En los eventos de Windows generados por Sysmon se observan tres tipos de eventos:
   a) Evento con ID 1: Este evento indica la creación de un nuevo proceso de Ngrok mediante la ejecución por linea de comandos de:

      ngrok tcp 3389.
   b) Eventos con ID 22: Estos eventos indican resoluciones DNS realizadas por el equipo. En concreto, se observan tres resoluciones DNS distintas hacia los dominios *tunnel.ngrok[.]com*, *crl.ngrok[.]com* y *update.equinox[.]io*, todos ellos resolviendo a direcciones IP de equipos de *Amazon Web Service*, por sus siglas en inglés AWS. Los dos primeros dominios pertenecen al servicio de Ngrok [19] y en el caso del tercer dominio este pertenece al servicio de gestión de actualizacion de paquetes de aplicaciones programadas en Go y es usado por el propio Ngrok.
   c) Eventos con ID 3: Estos eventos dan información sobre las conexiones de red creadas y se observan seis de ellos generados por el binario de Ngrok. En la siguiente tabla se detallan los parámetros de dichas conexiones: De los eventos resumidos en la tabla anterior

Tabla I
CONEXIONES DE RED CREADAS POR NGROK

IP destino	Puerto destino	IP origen	Puerto origen
3.122.29.226	443	10.0.2.15	50201
18.154.22.118	80	10.0.2.15	50202
54.237.133.81	443	10.0.2.15	50203
::::::: 1	3389	::::::: 1	50207
::::::: 1	3389	::::::: 1	50209
::::::: 1	3389	::::::: 1	50222

   se identifican tres sesiones a direcciones IP externas, dos de ellas con el protocolo HTTPS, puerto 443, y una de ellas HTTP al puerto 80. Estas sesiones contactan con las IP de los dominios previamente resueltos e inician una comunicación con los servidores de Ngrok. Las tres conexiones restantes son generadas también por el binario de Ngrok y tiene como puerto destino el indicado previamente. Estas últimas sesiones son internas del equipo y usan la IP de *loopback* [22] como destino y origen de la sesión. Estas seis sesiones son las que usa Ngrok para comunicarse con sus servidores, los cuales dirigirán el tráfico desde y hacia el equipo del atacante, y las que posteriormente en el equipo tunelarán el tráfico hacia el protocolo RDP.
4) Sesión RDP entrante: Cuando se inicia la sesión RDP entrante mediante el uso de Ngrok se identifican varios eventos de Windows relacionados con el protocolo de control remoto y con el flujo de inicio de sesión. A continuación se detallan los distintos eventos observados:
   - Evento 131 de Windows-RemoteDesktopServices-RdpCoreTS: Este evento indica que se ha aceptado una nueva conexión TCP en el equipo utilizando el servicio de *Remote Desktop Services* (RDS) de Microsoft. Se destaca de este evento que el campo de *"client.ip"* contiene el valor de la dirección IP de *loopback* del equipo, *[::1]:50207* con el puerto origen identificado en una de las conexiones de red creadas anteriormente.
   - Evento 261 de Microsoft-Windows-TerminalServices-RemoteConnectionManager:

Este evento nos indica que un *listener* del servicio de RDP ha recibido una nueva conexión.

- Evento 61 de Microsoft-Windows-RemoteDesktopServices-RdpCoreTS: El evento nos confirma que se ha creado una conexión RDP-TCP en el equipo.
- Evento 4624 de Microsoft-Windows-Security-Auditing: El evento 4624 confirma un inicio de sesión satisfactorio en el equipo. En este caso los parámetros nos indican el usuario, que se trata de un inicio de sesión remoto de tipo 3, es decir, de red, con dirección IP origen *"::1"* y puerto origen *0*.
- Evento 1149 de Microsoft-Windows-TerminalServices-RemoteConnectionManager: El evento en cuestión indica el éxito en la autenticación del usuario mediante el servicio de RDP. Se destaca que en los parámetros de este evento consta como dirección IP de origen *::%16777216*.
- Evento 132 de Microsoft-Windows-RemoteDesktopServices-RdpCoreTS: Este evento nos confirma que se ha establecido un canal entre el equipo y el origen de la conexión.

Por los eventos anteriores se confirma que se ha establecido una conexión RDP satisfactoria entre el equipo y la IP origen de *loopback* y el puerto origen previamente definido por el software de Ngrok.

5) Acciones durante la sesión RDP: La actividad de creación y modificación de ficheros realizados durante la conexión no presenta datos de interés distintos a los que se generan cuando se realizan las acciones en el equipo.
6) Finalización de la sesión RDP: Una vez finalizada la sesión RDP se identifican los eventos de 103, 72 y 102 de Microsoft-Windows-RemoteDesktopServices-RdpCoreTS, los cuales indican la desconexión de la sesión, el evento 4634 de Microsoft-Windows-Security-Auditing, que informa del cierre de sesión de la cuenta, y el evento 5 de Sysmon, que informa del fin de la ejecución del proceso de Ngrok. Tras la finalización de la sesión se puede observar el tráfico generado por el proceso en la base de datos de la System Resource Usage Monitor, SRUM [23] por sus siglas en inglés, la cual nos indica que el proceso de *ngrok.exe* ha consumido 4,264879 Megabytes de trafico enviado y 1,714172 Megabytes de trafico saliente, sumando un total de 5.979051 Megabytes.

B. Análisis captura de tráfico de red

Del análisis de la captura de tráfico de red se puede corroborar la actividad observada en los artefactos forenses del equipo. Concretamente se destacan dos tipos de comunicaciones:

1) Comunicaciones DNS: En las comunicaciones generadas en el equipo se identifica en el momento de la ejecución del tunelado del puerto 3389 las tres peticiones DNS a los dominios de Ngrok y de Equinox. Estas peticiones son resueltas con varias direcciones IP, entre las que se encuentran las identificadas en las conexiones de red generadas en el equipo.
2) Conexiones TCP a las direcciones IP anteriores: Durante el periodo de actividad de Ngrok se observan conexiones con las direcciones IP y puertos siguientes: *3.122.29.226:443*, *18.154.22.118:80* y *54.237.133.81:443*. Observando las volumetrías y la frecuencia de peticiones se identifica que tan solo una IP genera tráfico elevado y de manera constante, siendo esta la sesión con destino *3.122.29.226:443* con resolución inversa al dominio *tunnel.ngrok[.]com*. Esta sesión genera un volumen total de 25.906 paquetes divididos en 4,066 Megabytes enviados y 1,628 Megabytes recibidos en el equipo, un total de 5,694 Megabytes. Este valor se corresponde con el 95,233% del tráfico identificado en el artefacto forense SRUM.

V. DISCUSIÓN

Tras el análisis forense se obtienen varios resultados que nos permite caracterizar esta actividad y así agilizar la detección, el tiempo de análisis y de respuesta ante el incidente de ciberseguridad. Pese a que se han identificado gran parte de eventos en el sistema que permiten trazar la actividad realizada en el equipo por un supuesto atacante que usara Ngrok para tunelizar el tráfico RDP y exponer el servicio de manera publica a Internet, gran parte de estos eventos son generados por Sysmon, una herramienta que pocas veces está presente en los equipos dado que debe ser instalada por los administradores. Descartando estos eventos, la detección y la caracterización de su uso se reduce tan solo a unos pocos artefactos: la MFT permite identificar la creación del binario y de su fichero de configuración, los ficheros de Prefetch nos indican las últimas horas de su ejecución y los eventos de RDP y de inicio de sesión nos indican las sesiones creadas remotamente. Se destaca de estos últimos la característica particular del uso de una dirección IP de *loopback* como origen, la cual nos permite identificar que la sesión ha sido tunelizada por otro proceso corriendo en el equipo. La SRUM de Windows nos permite identificar el volumen de Bytes consumidos por el proceso, lo cual puede ser de ayuda para determinar la duración de las sesión y si se ha realizado exfiltración de información mediante estas. Es en la inspección del tráfico de red donde se puede obtener información complementaria a las comunicaciones realizadas, obteniendo una fecha de inicio del servicio mediante las resoluciones DNS y el detalle de las comunicaciones con las direcciones IP del servicio de Ngrok a lo largo de las sesiones. Por todo lo anterior se plantean dos posibles mecanismos de detección, que podrían ser aplicados a los sistemas de seguridad de la infraestructura para bloquear y monitorizar la actividad relacionada con un atacante:

- Monitorizar actividad de RDP con origen la IP de

loopback. De disponer de un sistema de Security Information and Event Management, SIEM [24], se propone indexar los eventos 4624 y 1149 de Windows. Estos eventos suelen ser auditados por defecto en los equipos y, de ser indexados en un sistema de SIEM, se puede crear una regla de detección que monitorice las direcciones IP origen, detectando y alertando cuando estas conexiones provienen de direcciones IP de *loopback*. Cabe destacar que esta regla de detección puede permitir detectar otro tipo de software que esté tunelando tráfico del protocolo RDP, y no tan solo cuando se use el software de Ngrok.

- Monitorizar el tráfico hacia servidores de Ngrok: Se propone monitorizar la resolución del dominio Ngrok en el servidor DNS de la infraestructura y las comunicaciones con las direcciones IP del dominio *tunnel.ngrok.io* [19]. Adicionalmente, en caso de no usar el software de manera habitual en la infraestructura, se pueden bloquear estas comunicaciones.

VI. CONCLUSIONES

En este articulo se ha analizado el rastro generado por la actividad del software Ngrok para la tunelización del protocolo RDP mediante conexiones HTTPS. Se ha recreado la actividad habitual de un atacante que usa este software, analizando a posterior las evidencias presentes en los artefactos forenses del equipo y el tráfico de red generado. A raíz del análisis se identifica que son pocos los artefactos que permiten caracterizar esta actividad, pero que de disponer de sistemas de monitorización de eventos y de red se puede identificar esta técnica de los atacantes, bloqueando así su uso para mitigar y prevenir el posible compromiso, evitando un potencial incidente de *ransomware*.

AGRADECIMIENTOS

Este trabajo ha sido posible gracias a la financiación del Ministerio de Ciencia y Educación con el proyecto: TCO-RISEBLOCK (PID2019-110224RB-I00) y de la Generalitat de Catalunya con la subvención 2021-SGR-00594 y 2021 DI 92. Así como la investigación previa realizada por la empresa INCIDE.

REFERENCIAS

[1] P. Tavares, "Tunneling and port forwarding tools used during red teaming assessments" [Online]. Available: https://shorturl.at/howY3. [Accessed: June 15, 2023].

[2] J. Hammond, "Abusing Ngrok: Hackers at the End of the Tunnel" [Online]. Available: https://www.huntress.com/blog/abusing-ngrok-hackers-at-the-end-of-the-tunnel. [Accessed: June 15, 2023].

[3] S. Renjith, "HOW THREAT ACTORS STEAL YOUR DATA WITH REVERSE TUNNELLING" [Online]. Available: https://www.hawk-eye.io/2021/11/how-threat-actors-steal-your-data-with-reverse-tunnelling/. [Accessed: June 15, 2023].

[4] T. Dargahi, A. Dehghantanha, P. N. Bahrami, M. Conti, G. Bianchi, L. Benedetto, "A Cyber-Kill-Chain based taxonomy of crypto-ransomware features" in *Journal of Computer Virology and Hacking Techniques*, vol. 15, pp. 277-305, 2019, Springer.

[5] S. Montes, "Los hacktivistas rusos Sparta y KillNet atacan por primera vez a empresas españolas" [Online]. Available: https://shorturl.at/eiwNU. [Accessed: June 15, 2023].

[6] CISA, "StopRansomware: LockBit 3.0" [Online]. Available: https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-075a. [Accessed: June 15, 2023].

[7] Microsoft, "Understanding the Remote Desktop Protocol (RDP) " [Online]. Available: https://learn.microsoft.com/en-us/troubleshoot/windows-server/remote/understanding-remote-desktop-protocol. [Accessed: June 15, 2023].

[8] ngrok, Inc. "ngrok" [Online]. Available: https://ngrok.com/. [Accessed: June 14, 2023].

[9] RealVNC "RPort" [Online]. Available: https://rport.io/. [Accessed: June 14, 2023].

[10] SophosLabs "BlackCat ransomware attacks not merely a byproduct of bad luck" [Online]. Available: https://www.zen-networks.io/blackcat-ransomware-attacks-not-merely-a-byproduct-of-bad-luck/. [Accessed: June 14, 2023].

[11] COUNTER THREAT UNIT RESEARCH TEAM "COBALT MIRAGE Conducts Ransomware Operations in U.S." [Online]. Available: https://www.secureworks.com/blog/cobalt-mirage-conducts-ransomware-operations-in-us. [Accessed: June 14, 2023].

[12] ClearSky Security Ltd. "Fox Kitten Campaign" [Online]. Available: https://www.clearskysec.com/wp-content/uploads/2020/02/ClearSky-Fox-Kitten-Campaign-v1.pdf. [Accessed: June 14, 2023].

[13] Serveo.net "Serveo" [Online]. Available: https://serveo.net/. [Accessed: June 14, 2023].

[14] @fatedier "frp" [Online]. Available: https://github.com/fatedier/frp. [Accessed: June 14, 2023].

[15] The Beanstalks Project ehf. "Pagekite" [Online]. Available: https://pagekite.net/. [Accessed: June 14, 2023].

[16] Microsoft "Sysmon v14.16" [Online]. Available: https://learn.microsoft.com/en-us/sysinternals/downloads/sysmon. [Accessed: June 14, 2023].

[17] Wireshark Foundation "Wireshark" [Online]. Available: https://www.wireshark.org/. [Accessed: June 14, 2023].

[18] GRR team "GRR Rapid Response" [Online]. Available: https://grr-doc.readthedocs.io/en/latest/. [Accessed: June 14, 2023].

[19] ngrok, Inc. "tunnel.json" [Online]. Available: https://s3.amazonaws.com/dns.ngrok.com/tunnel.json. [Accessed: June 14, 2023].

[20] M. Alazab, S. Venkatraman, P. Watters, "Effective digital forensic analysis of the NTFS disk image" in *Ubiquitous Computing and Communication Journal*, vol. 4, no. 1, pp. 551-558, 2009.

[21] N. Shashidhar, D. Novak, "Digital forensic analysis on prefetch files" in *International Journal of Information Security Science*, vol. 4, no. 2, pp. 39-49, 2015.

[22] IFF DAVIS, LLC. "loopback address" [Online]. Available: https://www.pcmag.com/encyclopedia/term/loopback-address. [Accessed: June 15, 2023].

[23] Y. Khatri, "Forensic implications of system resource usage monitor (SRUM) data in Windows 8" in *Digital Investigation*, vol. 12, pp. 53-65, 2015, Elsevier.

[24] M. Cinque, D. Cotroneo, A. Pecchia, "Challenges and directions in security information and event management (SIEM)" in *2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*, pp. 95-99, 2018, IEEE.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Desarrollo y evaluación de técnicas para el análisis de la calidad de la información proveniente de infraestructuras IoT

Laura Martín, Luis Sánchez, Jorge Lanza, Pablo Sotres
Departamento Ingeniería de Comunicaciones,
Universidad de Cantabria
Plaza de la Ciencia s/n, Santander, 3005, España.
lmartin@tlmat.unican.es, lsanchez@tlmat.unican.es, jlanza@tlmat.unican.es, psotres@tlmat.unican.es

Hoy en día, los datos son el motor de la prosperidad económica y la base fundamental de sistemas de toma de decisiones, apoyados por la Internet de las Cosas como una de las tecnologías más representativas. Por ello, la calidad de estos datos se convierte en un aspecto crítico. En este trabajo se presenta una solución para evaluar varias dimensiones de calidad de los flujos de datos IoT a medida que se generan. Además, el enfoque seguido se centra en añadir la información sobre la calidad de los datos como metadatos vinculados a cada elemento del flujo de datos, aprovechando los principios de los datos enlazados y el estándar NGSI-LD para armonizar y enriquecer fuentes de datos heterogéneas. Por último, se evalúa el diseño propuesto, logrando un compromiso sólido entre funcionalidad y sobrecarga, con un rendimiento estable y escalable.

***Palabras Clave*—calidad de los datos, curado de los datos, Inteligencia Artificial, Internet de las Cosas**

I. INTRODUCCIÓN

El creciente número de fuentes de datos asociadas al despliegue de la Internet de las Cosas (IoT, *Internet of Things*), así como aquellas vinculadas a portales de datos abiertos y plataformas de redes sociales están generando una inmensa cantidad de información. Estos datos son sumamente útiles y beneficiosos, tanto para el sector público como privado, gracias al desarrollo de servicios de valor añadido, el aumento de la transparencia y la disponibilidad de las administraciones o el fomento de la eficiencia de los servicios públicos. En febrero de 2020, la Comisión Europea anunció la Estrategia Europea de Datos [1], cuyo objetivo es crear un mercado único para que los datos se compartan e intercambien entre diferentes sectores de manera eficiente y segura dentro de la UE.

Entre las tecnologías que van a desempeñar un papel clave en la futura Economía de los Datos, la IoT es reconocida como una tecnología que cambia las reglas del juego y amplía su aplicabilidad a una enorme variedad de dominios [2]–[5]. Precisamente, los datos generados constantemente por la miríada de sensores incrustados en el entorno pueden transformarse en conocimiento de valor añadido si se procesan de manera inteligente.

Sin embargo, cuanto mayor es la infraestructura IoT desplegada, más probables son los fallos del sistema y de la red [6]. Los errores se deben, principalmente, al uso generalizado de dispositivos de bajo coste, aspecto fundamental para conseguir tener un mayor alcance y adaptación de la IoT en la sociedad. Esto también genera inquietudes en cuanto al uso de estos dispositivos, debido a sus niveles de precisión, fiabilidad, aplicabilidad sobre el terreno y rendimiento. Los errores y fallos introducidos por el uso de dispositivos de coste reducido pueden dar lugar a una mala calidad de los datos (DQ, *Data Quality*), que, a su vez, terminan conduciendo a resultados indeseados en sistemas de toma de decisiones.

Dentro de la literatura se pueden encontrar múltiples definiciones del concepto de DQ. La definición propuesta por [7] es una de las más extendidas, acuñando la filosofía "*adecuado para su uso*". Respecto a las clasificaciones de las métricas para la medición de la calidad de los datos, se identifican numerosas compilaciones y definiciones de las dimensiones de calidad [7]–[10], destacando el conjunto inicialmente formulado por [7] (Categorías Intrínseca, Contextual, Accesibilidad y Representación).

La mayoría de trabajos enfocados al concepto de DQ se han concentrado en el aspecto Big Data de la IoT, investigando, clasificando y discutiendo su gestión para grandes conjuntos de datos [11]. Sin embargo, la IoT posee una naturaleza dinámica fundamental y una parte importante de su potencial proviene de su capacidad para generar datos en tiempo real organizados en los llamados flujos de datos (es decir, series temporales de observaciones sucesivas tomadas por dispositivos IoT). Por lo tanto, es de

gran importancia que la calidad digital se evalúe y mejore a medida que se generan y recogen los datos. Centrando el estudio en la calidad de la información en entornos IoT y flujos de datos, en [12]–[14] se realizan diferentes selecciones de dimensiones de calidad de datos dedicadas a este tipo de escenarios.

Dados los análisis realizados por estas tres publicaciones sobre las dimensiones de calidad enfocadas a entornos IoT, se han priorizado aquellas que proveen información sobre los datos tanto de manera individual como en conjunto, escogiendo finalmente cinco dimensiones: Exactitud —*Accuracy*, Precisión —*Precision*, Puntualidad —*Timeliness*, Integridad —*Completeness* y Usabilidad —*Usabillity*. Esta última propiedad, Usabilidad, se aporta en este trabajo, promoviendo la interoperabilidad de la información y su fácil consumo, de acuerdo con la representación homogénea y estadarizada de los flujos de datos.

Junto con la definición teórica de las dimensiones DQ, se deben especificar los métodos para evaluar la calidad de los flujos de datos en entornos IoT. En [15], los autores se centran en el cálculo de un conjunto de dimensiones DQ y en la aplicación de este tipo de técnicas de mejora y evaluación de la calidad, pero sólo una vez se ha obtenido un conjunto de datos lo suficientemente extenso como para llevar a cabo estas acciones. Por el contrario, en este artículo se propone el modelado y cálculo de las dimensiones DQ seleccionadas, así como varias técnicas de enriquecimiento de la calidad de los datos. Ambos procesos se van a poder aplicar sobre las observaciones recogidas dentro de despliegues IoT de forma continua, es decir, a medida que se van generando.

Asimismo, esta información de calidad obtenida como metadatos DQ se incluirán junto con los flujos de datos evaluados. De este modo, se consigue el máximo potencial del curado de datos al poder acceder directamente a estas características DQ para cada elemento del flujo de datos. Este enfoque no está muy extendido en la literatura, y es más típico evaluar las dimensiones de calidad de manera periódica y enviar alertas en caso de que se superen ciertos límites, tal y como propone [16].

Por otro lado, existen diferentes enfoques sobre la integración del módulo de evaluación de la calidad en las arquitecturas de adquisición y consumo de datos. En [16], se propone una arquitectura y metodología para la evaluación y monitorización de la DQ de forma externa a una plataforma IoT. Sin embargo, en [17] se discute la incorporación de un conjunto de actividades que pueden ser realizadas en paralelo o integradas con el resto de actividades en curso. En este trabajo se propone la integración del Módulo de Curado de Datos IoT en una arquitectura DET (*Data Enrichment Toolchain*) para evaluar la viabilidad e idoneidad de incluir dimensiones DQ en plataformas IoT existentes para mejorar aún más la dimensión Usabilidad que se ha definido.

Así, las contribuciones clave de este trabajo son: (i) identificar y definir las dimensiones DQ más relevantes en los flujos de datos IoT y los mecanismos para evaluar cada una de ellas; (ii) especificar e implementar soluciones de curado de datos que empleen algoritmos de IA (Inteligencia Artificial) para enriquecer los flujos de datos IoT incrementando sus características de calidad; (iii) integrar estos dos tipos de mecanismos (es decir, la evaluación y el enriquecimiento de la DQ de los flujos de datos IoT) en una DET operativa que aproveche los principios de datos enlazados y el estándar NGSI-LD para proporcionar datos enriquecidos semánticamente; y (iv) llevar a cabo una síntesis y reevaluación de los resultados presentados en [18] con el objetivo de resaltar los beneficios conseguidos por el Módulo de Curado de Datos IoT en las diferentes dimensiones de calidad de la información evaluadas.

II. DIMENSIONES DE CALIDAD DE LA INFORMACIÓN

El concepto DQ es crucial para los procesos de minería de datos y análisis de la información, siendo crítico en el uso de la tecnología IoT debido a su gran impacto en los productos finales [14], [19]. DQ define el grado de cumplimiento de los requisitos impuestos por los consumidores de datos, y las dimensiones DQ se refieren a los criterios que deben cumplirse para que los resultados del análisis y el consumo de información sean óptimos y no se vean comprometidos.

Es importante destacar que el planteamiento seguido se enfoca en proporcionar conocimiento sobre las dimensiones DQ que sea útil para que los consumidores sean capaces de comprender los datos que van a recibir. Es decir, no se trata de obtener un valor único que indique si un ítem de un flujo de datos es de alta o baja calidad, sino de incluir metadatos que permitan decidir si ese elemento tiene suficiente calidad o no, enriqueciendo así los criterios de selección del flujo de datos.

A. Definiciones

Siguiendo la clasificación en categorías de las dimensiones de calidad ofrecida por [7], se han escogido las siguientes: Exactitud (*Accuracy*) como dimensión Intrínseca clave; Integridad (*Completeness*), Puntualidad (*Timeliness*) y Precisión (*Precision*) como dimensiones Contextuales; y, una quinta dimensión aportada en este trabajo, Usabilidad (*Usability*), que de alguna manera aúna los aspectos de las categorías DQ de Representación y Accesibilidad.

En los siguientes apartados se exponen las definiciones utilizadas y sus correspondientes métodos de cálculo para cada una de las dimensiones DQ seleccionadas.

1) Exactitud: (*Accuracy*) indica lo cerca que está el valor medido por el dispositivo IoT del valor considerado como verdad absoluta. Esto se representa en la Ec. 1, tomando las unidades del valor de observación y representado por el símbolo ± a su lado.

$$exactitud = |valorObservado - valorReferencia| \quad (1)$$

2) Integridad: (*Completeness*) cuantifica el número de observaciones perdidas en una ventana temporal determinada. El método seguido para el cálculo de este parámetro se muestra en la Ec. 2 y se representa en parte por unidad (ppu). Los parámetros que intervienen en este

método son: $ventana\_temporal$, ventana temporal; n, número de observaciones consideradas perdidas en esa ventana temporal; y $tasa$, frecuencia de llegada media de las observaciones (también conocida como Puntualidad o *Timeliness* en este trabajo).

$$integridad = \frac{ventana\_temporal - n \cdot tasa}{ventana\_temporal} \quad (2)$$

3) Puntualidad: (*Timeliness* o tn) es una dimensión con una gran cantidad de definiciones dentro de la literatura [8], [13], [14], [19]. El enfoque seguido en este trabajo define esta propiedad como la tasa de actualización en el sistema del valor observado y su cálculo se encuentra representado en la Ec. 3 como una media ponderada a través del parámetro α del valor de tn de la iteración anterior (tn_{i-1}) y el calculado con la llegada de la observación que se está evaluando (tn_{bruto}).

$$tn_i = \alpha \cdot tn_{i-1} + (1 - \alpha) \cdot tn_{bruto} \quad (3)$$

4) Precisión: Pese a que la propiedad de Precisión (*Precision*) no pertenece al conjunto original de dimensiones de calidad para IoT encontrado en la literatura, sí que suele aparecer en segundo plano [7], [8], [10], [12]. Tomando como base la definición ISO 5725 de Precisión [20], el objetivo es evaluar la cercanía entre los valores del flujo de datos. El método seguido para su cálculo se muestra en la Ec. 4 y su resultado, incluido con el símbolo ±, toma las unidades de los valores del conjunto de datos. Los parámetros involucrados en este cálculo son: x_i, elemento i del conjunto de datos; μ, media aritmética del conjunto de datos o, en este caso, valor de la observación que se está evaluando; n, número de elementos en el flujo de datos.

$$precisión = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}} \quad (4)$$

5) Usabilidad: A diferencia de las anteriores y dentro del alcance del estudio de este trabajo, la dimensión Usabilidad no está definida, como tal, en la literatura [7], [10], [21]. El concepto más similar podría ser el de "*adecuación para su uso*" descrito en [7]. Ambos conceptos comparten el enfoque de definir una dimensión de calidad polifacética que esté centrada en el consumidor. Sin embargo, se orienta a dimensiones intrínsecas objetivas, sin prestar suficiente atención a las dimensiones subjetivas y contextuales relacionadas con el procedimiento real del consumo de datos.

Usabilidad, al tratarse de una dimensión en la perspectiva subjetiva de DQ, no incluye una expresión cerrada para evaluarla, pero para que los datos puedan calificarse como de alta calidad en sus términos deben cumplir tres aspectos básicos. En primer lugar, los datos deben estar representados de manera uniforme, siguiendo modelos estándar y consiguiendo así coherencia en su representación. En segundo lugar, es importante comprender cuál es la procedencia y linaje de los datos, conociendo cuál ha sido el tratamiento previo al que han sido sometidos, con el objetivo de proveer la confianza necesaria en los consumidores para inyectar estos datos en sus aplicaciones. Por último, la representación de los datos de debe realizarse de manera semánticamente rica, de modo que los datos no sean sólo valores, sino que estén vinculados a cualquier característica asociada a ese valor, en particular las relacionadas con sus dimensiones de calidad.

B. Mecanismos basados en IA para el curado de la información

En [19] se proponen técnicas de mejora de la calidad para datos proporcionados en un entorno IoT: detección de valores atípicos, interpolación, integración de datos, eliminación de duplicados y limpieza de datos. De entre todas estas técnicas, en este trabajo se hace hincapié sobre la imputación de valores, tanto en conjuntos de datos estáticos como en flujos de datos IoT en tiempo real.

La ausencia de valores en conjuntos de datos es un asunto crucial que debe gestionarse, ya que estos valores perdidos pueden sesgar los resultados de la aplicación de otras técnicas o, directamente, reducir la calidad de la información [19]. Dos de los métodos más conocidos para abordar la ausencia de valores son la interpolación y la aplicación de Inteligencia Artificial [19], [21]–[23]. Las técnicas derivadas de este último método se basan, principalmente, en el aprendizaje supervisado, estimando los valores ausentes en función de la información disponible por los valores existentes. El algoritmo kNN (k-Nearest Neighbour) es uno de los más significativos que, utilizado en este enfoque de imputación de valores, implica la estimación del hueco en base a una métrica de distancia entre sus k vecinos más cercanos.

En un escenario IoT, donde la información se distribuye en series temporales y debe procesarse en tiempo real, algunas de las técnicas y dimensiones DQ comentadas anteriormente no se pueden aplicar. Las técnicas que utilizan métodos de Inteligencia Artificial se ven obligadas a reentrenar sus modelos por cada nueva observación recibida de los dispositivos IoT desplegados, ya que la marca temporal o timestamp es crucial. Por tanto, en un entorno IoT con una gestión de los datos en tiempo real se puede determinar que esta situación de reentreno continuo no es escalable. El planteamiento propuesto para abordar este tipo de situaciones se basa en la estimación de valores futuros, aumentando sintéticamente el conjunto de datos y así poder aplicar las técnicas de enriquecimiento sobre este intervalo de tiempo adicional. La estructura de datos más común en un entorno IoT, como ya se ha comentado, es la serie temporal, donde se pueden manifestar características como la estacionalidad. Debido a esto, se ha considerado que el modelo más adecuado para su análisis y estimación es el algoritmo ARIMA Estacional (SARIMA) [24].

III. CURADO AUTÓNOMO DE DATOS IOT

Haciendo referencia a la dimensión de calidad propuesta como Usabilidad, esta propiedad impone requisitos sobre la representación y el uso de los datos tratados. En este trabajo se defiende la idea de que los resultados, tanto de la aplicación de técnicas de enriquecimiento como de la evaluación de las dimensiones de calidad, no tienen valor

por sí mismos, sino que este valor se potencia una vez que se integran en los flujos de datos procesados. De esta manera, se consigue que el consumidor tenga acceso a información de mayor calidad (resultante de las técnicas de enriquecimiento y curado de datos) cuyas características de calidad (producto de las técnicas de evaluación de calidad y de las dimensiones DQ) sean conocidas.

El enfoque seguido para incorporar los resultados de los mecanismos de curación de datos desarrollados e introducir los metadatos asociados a la información proporcionada por los entornos IoT es a través de la integración de estos mecanismos dentro de una DET.

En los siguientes apartados se presenta la arquitectura general de la cadena de enriquecimiento (DET) y los aspectos clave que se han tenido que cumplir para integrar los mecanismos necesarios.

A. *Arquitectura de la cadena de enriquecimiento*

En este trabajo, la DET se define como una cadena de microservicios heterogéneos que tiene como resultado la mejora progresiva de la calidad y valor de la información original. Para ello, se han identificado cinco funciones clave: descubrimiento, con la capacidad de descubrir y solicitar recopilaciones de conjuntos y flujos de datos; formateo, con la transformación de los datos en bruto en conjuntos bien formados y estructurados de acuerdo a los modelos de datos descritos en el estándar NGSI-LD [25]; curado, con la identificación y/o corrección de datos que no reflejan la calidad esperada; enlazado, con la capacidad de relacionar diferentes conjuntos de datos; y, finalmente, enriquecimiento, con la capacidad de comprender y enmarcar las estructuras de datos en situaciones y contextos.

Como se puede ver en la Fig. 1, las fuentes de datos heterogéneas (como por ejemplo conjuntos y flujos de datos, CSV, JSON) se descubren e identifican en el primer paso descrito. A continuación, se aplica la transformación a NGSI-LD con el objetivo de armonizar los datos entrantes y permitir que los procesos posteriores se aprovechen de la semántica y los principios de los datos enlazados. Justo antes de que todos estos datos se almacenen en el Context Broker (pieza central de la API de NGSI-LD [25]), se aplican las técnicas de curado de los datos con el fin de maximizar la DQ, de modo que su resultado sirva como base para las aplicaciones y el resto de bloques funcionales de enlazado y enriquecimiento de la arquitectura.

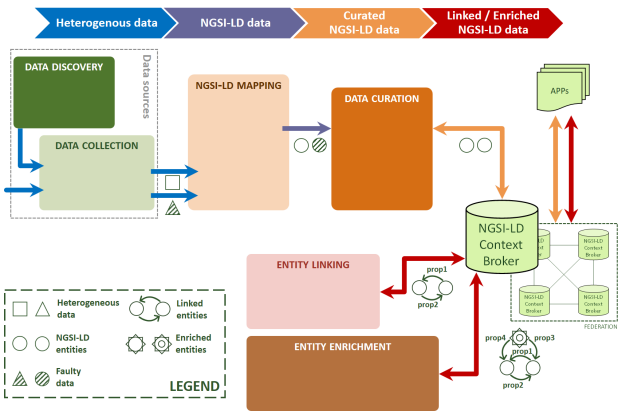


Fig. 1. Arquitectura de la cadena de enriquecimiento (DET) [26].

B. *Modelado de la información de calidad*

Tal y como se ha ido comentado a lo largo del artículo, los resultados de la aplicación de técnicas de enriquecimiento de DQ y el análisis de las dimensiones DQ sólo son valiosos si se incorporan en los flujos de datos como metadatos. De esta manera, los consumidores tienen la capacidad de comprender totalmente el significado de la información, no sólo a partir de su valor en bruto, sino también de sus características DQ.

Por lo tanto, se ha determinado que no sólo las mediciones generadas por los dispositivos IoT, sino también sus dimensiones y propiedades DQ obtenidas por las técnicas de mejora de la calidad, se representarán en el formato definido por el estándar NGSI-LD [25]. La aplicación de este estándar para la representación de los flujos de datos y sus características DQ se alinea con la materialización de los objetivos fijados en la dimensión Usabilidad. Se consigue la representación uniforme de la información de acuerdo con la normativa, la trazabilidad basada en los principios de los datos enlazados y la provisión de valor extra con metadatos a través de los campos de las propiedades incrustadas.

Con anterioridad a la realización de este trabajo, no existía ningún modelo de datos específico para la representación de la información de evaluación de la calidad en el catálogo de la iniciativa Smart Data Models [27]. Es por ello que, como contribución, se ha diseñado y proporcionado un modelo de datos que recoge todas las dimensiones DQ abordadas, así como diferentes propiedades adquiridas con las técnicas de enriquecimiento de datos para la trazabilidad de su aplicación. La documentación relevante se encuentra disponible en [28].

El objetivo principal de este formato de modelado de la información es vincular la observación recibida por los dispositivos IoT (entidad valor) y las propiedades de calidad obtenidas (modelo de datos propuesto).

C. *Módulo de curado de información IoT*

El componente responsable del curado de la información dentro de la arquitectura de la cadena de enriquecimiento (DET) se denomina Módulo de Curado de Datos IoT. Su principal objetivo es aplicar las técnicas de enriquecimiento de DQ y obtener las dimensiones de calidad para poder almacenar la información con la mayor calidad posible en el Context Broker.

El elemento de entrada de este componente se basa en la observación recibida por las fuentes heterogéneas representada ya en el formato definido por el estándar NGSI-LD en el modelo de datos de la iniciativa Smart Data Models que más se ajuste. Como ya se ha comentado, el objetivo de este componente es recopilar las características de calidad de la observación obtenida a la entrada, consiguiendo a su salida dos entidades. La primera de ellas se corresponde con la entidad procesada y, la segunda, con el modelo de

datos propuesto en el apartado anterior que reúne todas las propiedades que representan la calidad de la observación procesada. Es importante comentar que a la entidad valor (o entidad procesada) se le debe añadir un nuevo campo para crear la relación entre ambas entidades de salida, consiguiendo que estén explícitamente vinculadas y así alcanzar los requisitos impuestos por la dimensión de Usabilidad propuesta en este trabajo.

IV. EVALUACIÓN DE LAS TÉCNICAS DE ENRIQUECIMIENTO DE LA CALIDAD DE LA INFORMACIÓN

En esta sección se describe la evaluación llevada a cabo para caracterizar el comportamiento y el rendimiento de las técnicas de mejora de la calidad de la información propuestas en apartados anteriores.

A. Evaluación de las dimensiones de calidad

Tras la definición de las dimensiones DQ seleccionadas y propuestas anteriormente, se caracteriza su comportamiento y rendimiento en un entorno de pruebas.

Se ha simulado la existencia de 100 sensores desplegados que generan observaciones cada 2 minutos, hasta un total de 100 observaciones por dispositivo IoT. Con esto, se consigue un flujo de datos de 10000 elementos de datos como entrada para la cadena de evaluación DQ, por cada una de las 60 iteraciones del método de Monte Carlo [29] realizadas, con el fin de extraer conclusiones de validez general. Todas estas pruebas se han realizado sobre una máquina Ubuntu 20.04.5 LTS (2 núcleos CPU, 2.40 GHz de reloj, 16 GB de RAM) en la que se ha empleado un Broker Scorpio NGSI-LD como Context Broker. Al terminar cada una de las 60 iteraciones, el Context Broker se reinicia para comenzar desde su estado inicial.

La fórmula utilizada para calcular la sobrecarga (*sobrecarga*), en términos de tamaño, introducida por la inclusión de los metadatos generados a través del procedimiento de evaluación de la calidad se presenta en la Ec. 5. En ella, *enriquecida* se refiere al tamaño de la entidad incluyendo información sobre las dimensiones DQ, y *en bruto* se refiere al tamaño original de la entidad (es decir, sólo su valor sin elementos adicionales).

$$sobrecarga(\%) = \frac{enriquecida - en\,bruto}{en\,bruto} \cdot 100 \quad (5)$$

De manera complementaria a la sobrecarga en términos de tamaño, también se introduce cierto retardo en el tiempo de cálculo. Este retardo total se puede dividir en dos fases: el tiempo empleado en obtener toda la información necesaria para realizar el cálculo de la dimensión DQ por cada elemento de datos (denominado, *Retardo de petición*), y el tiempo necesario para efectuar realmente ese cálculo (denominado, *Retardo de procesado*).

En los siguientes apartados, se detallan los procesos de cálculo de cada una de las dimensiones de calidad.

1) Exactitud: Haciendo referencia a la definición propuesta para la dimensión de Exactitud, su valor se obtiene a partir de un valor de referencia. Por tanto, es evidente que se debe realizar una petición a una fuente externa de confianza que provea este valor. A efectos de esta investigación, la fuente externa se corresponde con la Agencia Estatal de Meteorología (AEMET) [30], ya que las observaciones que se están analizando son de tipo Temperatura dentro de la zona de Santander, Cantabria. Una vez que el submódulo recibe esta información de referencia, realiza el cálculo correspondiente (Ec. 1).

2) Integridad: Para realizar el cálculo de la dimensión de Integridad, es necesario que el Context Broker soporte el almacenamiento de valores temporales. De esta manera, el primer paso consiste en solicitar los valores almacenados en una ventana temporal predefinida para cada flujo recibido. Esta ventana temporal predefinida refuerza la correlación temporal para la evaluación de la dimensión de Integridad, lo que significa que se trata de una dimensión que debe tener en cuenta un historial limitado y no el completo del flujo de datos. El submódulo tiene que consultar estos valores históricos tanto para el tipo de entidad evaluado (por ejemplo, Temperatura) como para la entidad DataQualityAssessment [28] vinculada a ella. Tras obtener estos datos, el submódulo es capaz de evaluar la expresión en la Ec. 2. Es importante señalar que *n* sería el número de valores etiquetados como sintéticos en el modelo de datos propuesto y *rate* sería el valor de Puntualidad (*Timeliness*) de la observación actual.

3) Puntualidad: El cálculo de la dimensión de Puntualidad, al igual que las propiedades anteriores, es un procedimiento que se divide en dos fases: solicitud de la información necesaria y procesado para la obtención del valor de Puntualidad. En la fase de solicitud, el submódulo consulta al Context Broker el último valor de Puntualidad registrado de la entidad de evaluación de la calidad vinculada al *id* de la observación recibida. Una vez obtenido este valor, el submódulo continua con la fase de procesado, ejecutando el cálculo definido en la Ec. 3.

4) Precisión: La última dimensión que se procesa es Precisión. En este caso, para cada observación o entidad recibida en el submódulo, se debe realizar una solicitud al Context Broker sobre todas las entidades registradas que se encuentren dentro de un área determinado, tanto para el tipo de entidad evaluado como para la información de calidad vinculada a esta. Una vez el submódulo obtiene toda esta información, es capaz de efectuar la evaluación de la dimensión de Precisión que se describía en la Ec. 4.

5) Análisis de la sobrecarga introducida: Una vez descritos los procesos de cálculo de las dimensiones de calidad, conociendo cuáles son los requisitos de información previa de cada uno de ellos, se presentan los resultados obtenidos en términos de retardo y sobrecarga en las Tablas I y II.

En la Tabla I se muestran los valores medios, tras las 60 simulaciones de Monte Carlo, de los retardos en cada una de las fases necesarias para el cálculo de las dimensiones. Se puede ver que la dimensión Exactitud es la que mayor retardo implica, principalmente ocasionado por esta necesidad de petición a una fuente externa de referencia. En la Tabla II se incluye la sobrecarga introducida en términos de tamaño de cada una de las

Tabla I
RETARDO POR EL CÁLCULO DE LAS DIMENSIONES DE CALIDAD.

	Retardo de petición (ms)	Retardo de procesado (ms)	Retardo total (ms)
Dim. Exactitud	185.1	0.004	185.1
Dim. Integridad	40.7	0.06	40.8
Dim. Puntualidad	12.2	0.3	12.5
Dim. Precisión	64.7	14.8	79.8

Tabla II
SOBRECARGA INTRODUCIDA POR LAS DIMENSIONES DE CALIDAD A LA ENTIDAD BÁSICA.

	Tamaño (bytes)	Sobrecarga (%)
Entidad básica	1205	–
Dim. Exactitud	134	11.1
Dim. Integridad	137	11.4
Dim. Puntualidad	140	11.6
Dim. Precisión	134	11.1

dimensiones de calidad en el momento en el que se integran como metadato a la entidad observada. Este valor de sobrecarga es prácticamente constante en todas las dimensiones, añadiendo alrededor de 136 bytes cada una de ellas, frente a los 1205 bytes de la entidad original.

A la vista de estos resultados, se puede concluir que la inclusión de los metadatos relacionados con DQ impone una sobrecarga no despreciable que podría considerarse un inconveniente para la solución propuesta. Sin embargo, a partir de este análisis, queda claro también que tanto el aumento de los requisitos de almacenamiento como el retraso en el procesamiento, debido a esta integración de las propiedades adicionales, no es un precio tan alto a pagar a cambio de comprender el verdadero significado de los datos disponibles.

El acceso a funcionalidades de valor añadido (i.e. información valiosa sobre la calidad de los datos) siempre conlleva ciertos compromisos. La solución propuesta ofrece un equilibrio entre funcionalidad y sobrecarga.

B. Evaluación de las técnicas basadas en IA

Tras la revisión realizada sobre las técnicas elegidas para incrementar la calidad de la información que se hizo en la Sección II, en este apartado se evalúan algunos de los algoritmos y métodos para llevar a cabo estas técnicas.

En primer lugar, es importante destacar cuál es el conjunto de datos utilizado para la evaluación de estas técnicas. Este conjunto se compone de los registros históricos de temperatura, desde el 01 de enero de 2021 hasta el 13 de junio de 2022, de los sensores pertenecientes a SmartSantander [31]. Como proyecto de Ciudad Inteligente, SmartSantander cuenta con un gran número de sensores que informan de manera periódica cada 1, 2 o 5 minutos, alcanzando un volumen de alrededor de 2 GB de datos, con aproximadamente 16 millones de observaciones, durante el periodo de tiempo mencionado. En línea con la premisa de ser un proyecto de *Smart City*, la calidad de los dispositivos desplegados alrededor de la ciudad de Santander no es elevada, lo que provoca inexactitudes en las observaciones e incluso recurrencia de valores absurdos. Además de estos problemas, también es importante considerar la degradación temporal que sufren los dispositivos, ya que llevan desplegados desde el inicio del proyecto en 2011. En consecuencia, es evidente que el conjunto de datos inicial debe ser procesado antes de la aplicación de las técnicas de mejora de la calidad. De acuerdo con esto, se han aplicado tres métodos relacionados con el conocimiento del entorno: eliminación de absurdos, correlación espacial y correlación temporal.

El primer método, la eliminación de absurdos, se basa en establecer rangos lógicos a los valores de las observaciones recogidas por los sensores. En este caso, al tratarse del fenómeno atmosférico de temperatura en la ciudad de Santander, se han consultado los registros de AEMET [30] en el mismo rango temporal que el conjunto de datos.

El segundo método se corresponde con la correlación espacial. Debido a que en el proyecto SmartSantander se colocaron algunos dispositivos en autobuses y taxis, se han registrado medidas fuera de Santander, como por ejemplo, en Bilbao o Madrid. Por lo tanto, al centrar el estudio de todo el conjunto de datos dentro del área geográfica de Santander, se han eliminado aquellas observaciones tomadas fuera de estas coordenadas.

El último método se centra en la correlación temporal. Para ello, se han agrupado las observaciones en una frecuencia periódica (horaria) y se han promediado sus valores. De este modo, se elimina la posibilidad de que se dupliquen las marcas temporales de las observaciones de distintos sensores y se reduce considerablemente el volumen de datos, sin llegar a perder información.

Una vez preprocesados los datos, el conjunto ya se encuentra listo para ser evaluado bajo las técnicas de enriquecimiento de la calidad propuestas en este trabajo.

Comenzando por la imputación de valores, el objetivo de la evaluación de esta técnica es comparar los diferentes métodos propuestos a través de errores absolutos (MAPE – *Mean Absolute Percentage Error* y MAE – *Mean Absolute Error*). Para ello, se ha dividido el conjunto de los datos en dos grupos: conjunto de entrenamiento y conjunto de validación. El primer grupo se compone de un subconjunto de datos que incluye huecos entre sus observaciones. Estas observaciones "perdidas" son las que componen el subconjunto de validación. Por tanto, los métodos de interpolación y el algoritmo kNN elegidos se aplican sobre el primer subconjunto de datos, obteniendo como resultado los valores de las observaciones ausentes. Tras ello, se

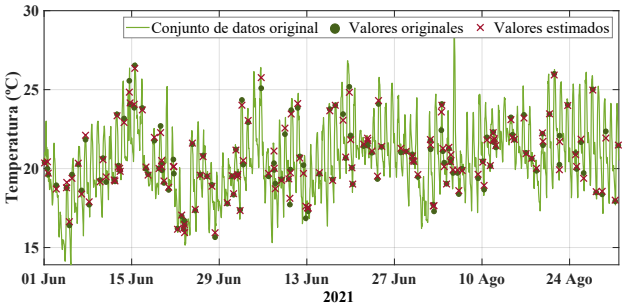


Fig. 2. Imputación de valores utilizando la interpolación polinomial de grado 2. Zoom temporal sobre un periodo de 3 meses.

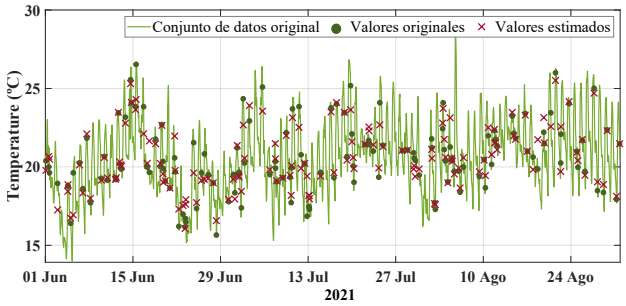


Fig. 3. Imputación de valores utilizando el algoritmo kNN con 5 vecinos. Zoom temporal sobre un periodo de 3 meses.

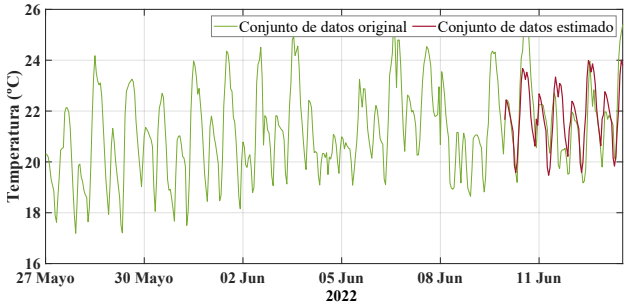


Fig. 4. Estimación del futuro inmediato con el algoritmo SARIMA (configuración (0,1,1)(2,1,0)[24]).

comparan estos valores imputados con los que se habían extraído en el segundo subconjunto de datos, validando y evaluando la exactitud de estas técnicas.

Con el objetivo de facilitar la comprensión sobre los resultados expuestos en las figuras, se han aplicado las técnicas sobre un periodo temporal de 3 meses (junio, julio y agosto de 2021). De esta forma, se obtienen los resultados mostrados en la Fig. 2 y la Fig. 3 de manera más visual, y los resultados numéricos mostrados en las dos primeras filas de la Tabla III de las técnicas aplicadas, interpolación polinomial de grado 2 y algoritmo kNN con 5 vecinos, respectivamente. Las figuras muestran la diferencia mínima en el rendimiento de ambas técnicas, representando los valores originales (conjunto de validación) con puntos verde oscuro y los valores estimados con cruces rojas (resultados de la aplicación de las técnicas). En cuanto a los datos de exactitud mostrados en la Tabla III, se observa que la interpolación polinomial de grado 2 obtiene mejores resultados que el algoritmo kNN, alcanzando valores de exactitud de 99.941% y 99.825% (1-MAPE), respectivamente. La diferencia es insignificante, concluyendo que ambas técnicas son buenas opciones para la imputación de valores.

La otra técnica de enriquecimiento mencionada es la ampliación sintética del conjunto de datos base con la intención de poder aplicar otras técnicas sobre los flujos de datos en tiempo real y así disponer de este rango temporal adicional como apoyo. Dada la naturaleza del conjunto de datos (series temporales con estacionalidad), se ha utilizado el algoritmo SARIMA. En este caso, debido a este aspecto de estacionalidad, se debe acortar el conjunto de datos (a 3 meses) para eliminar la periodicidad anual

Tabla III
EVALUACIÓN DE LAS TÉCNICAS DE IMPUTACIÓN DE VALORES Y ESTIMACIÓN DEL FUTURO INMEDIATO.

	Error porcentual absoluto medio (MAPE)	Error absoluto medio (MAE)
Interpolación pol. de grado 2	0.000592	0.0121
kNN con 5 vecinos	0.001740	0.03541
SARIMA (0,1,1)(2,1,0)[24]	0.040075	0.87221

Tabla IV
MEJORA TRAS LA APLICACIÓN DE LAS TÉCNICAS DE CURADO.

	Pre-procesado (mín-máx)	Post-procesado (mín-máx)	Ganancia (%)
Exactitud ($\pm^oC$)	$0 - 183.1$	$0.0 - 6.7$	95.078
Integridad (ppu)	$0.455 - 1$	$1 - 1$	0.864
Puntualidad (min)	$0.997 - 13.26$	$0.997 - 10.94$	1.124
Precisión ($\pm^oC$)	$0 - 193.07$	$0 - 4.49$	97.642

y mantener la diaria. Teniendo esto en cuenta, se aplica este algoritmo con la configuración más óptima de sus parámetros y el resultado se muestra en la Fig. 4. Además, en la última fila de la Tabla III se muestran las métricas utilizadas para la evaluación numérica del método. Se puede ver que se obtiene una exactitud del 96% (1-MAPE), lo que puede considerarse un buen rendimiento.

Por último, la Tabla IV muestra la mejora conseguida mediante la aplicación de las técnicas de curado ya descritas. Para ello, se han evaluado las dimensiones de calidad en cada uno de los 65 datasets (series temporales correspondientes a cada uno de los sensores utilizadas en la evaluación) y se han comparado los valores de éstas, antes y después de haber sido procesadas en el Módulo de Curado de Datos IoT implementado. Como se puede ver, se consiguen mejoras significativas en todas y cada una de las dimensiones evaluadas. Exactitud y Precisión son en las que se alcanzan mayores ganancias debido a la existencia de una gran cantidad de valores anómalos, los cuales han sido eliminados en el proceso de curado. Integridad y Puntualidad presentan valores de ganancia residuales, ya que dependen del número de observaciones perdidas (marginal en el flujo de datos que se ha empleado en la validación). Aún así, el efecto introducido por el proceso de curado es siempre de mejora. Los parámetros evaluados que se muestran en la Tabla IV son: para la Exactitud y la Precisión, la amplitud del rango de temperaturas (en unidades de variación de grados Celsius, ±ºC); para la Integridad, la cantidad de observaciones no perdidas (en partes por unidad, ppu); y para la Puntualidad, la tasa de actualización (en minutos).

V. CONCLUSIONES

Dado el incremento de la importancia de los datos en la actualidad, garantizar su calidad y ser capaces de comprender todas las dimensiones que tiene el ámbito de la calidad de la información (DQ), se convierten en condiciones ineludibles para cualquier plataforma de gestión de datos. En particular, en las plataformas IoT esto es especialmente importante debido a las particularidades de este tipo de

infraestructuras, que hacen que en ocasiones sea imposible imponer unos requisitos mínimos de DQ.

En este artículo se presenta el trabajo llevado a cabo para evaluar dimensiones específicas de DQ para flujos de datos IoT y compensar, mediante mecanismos de curado de datos habilitados por IA, los valores deficientes en estas dimensiones. La sobrecarga en términos de retardo y de tamaño, debido al cálculo y evaluación de las dimensiones de calidad, ha presentado valores adecuados dentro de un equilibrio entre funcionalidad y rendimiento dado el servicio de valor añadido presentado. En este sentido, se ha evaluado la mejora introducida por el módulo de curado que se ha implementado para este trabajo. Por otro lado, las técnicas de mejora y enriquecimiento de la calidad del conjunto de datos han demostrado un comportamiento apropiado, medido a través de las métricas MAPE y MAE, además de haber expuesto el presente beneficio tras su uso.

Los próximos pasos de esta investigación se basan en el estudio de la posible ampliación del número de dimensiones DQ evaluadas y su elección dinámica por parte del consumidor en la cadena de enriquecimiento (DET). Además, también se plantea investigar diferentes técnicas de enriquecimiento no contempladas hasta el momento e incorporarlas en esta cadena de curado de la información.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Programa CEF de la Comisión Europea a través del proyecto SALTED "Situation-Aware Linked heTerogeneous Enriched Data" bajo el Número de Acción 2020-EU-IA-0274 y por la Agencia Estatal de Investigación (AEI) mediante el proyecto THROTTLE "Mercado de Datos de Movilidad Urbana Confiable" bajo el Acuerdo de Subvención nº TED2021-131988B-I00 (en el marco del Plan de Recuperación, Transformación y Resiliencia) y el proyecto SITED "Semantically-enabled Interoperable Trustworthy Enriched Data-spaces" bajo el Acuerdo de Subvención nº PID2021-125725OB-I00.

REFERENCIAS

[1] European Commission, "A European strategy for data," 2020. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066

[2] P. Bellini, P. Nesi, and G. Pantaleo, "Iot-enabled smart cities: A review of concepts, frameworks and key technologies," *Applied Sciences*, vol. 12, no. 3, p. 1607, 2022.

[3] B. B. Sinha and R. Dhanalakshmi, "Recent advancements and challenges of internet of things in smart agriculture: A survey," *Future Generation Computer Systems*, vol. 126, pp. 169–184, 2022.

[4] Y. Yang, H. Wang, R. Jiang, X. Guo, J. Cheng, and Y. Chen, "A review of iot-enabled mobile healthcare: technologies, challenges, and future trends," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9478–9502, 2022.

[5] K. Garg, C. Goswami, R. Chhatrawat, S. K. Dhakar, and G. Kumar, "Internet of things in manufacturing: A review," *Materials Today: Proceedings*, vol. 51, pp. 286–288, 2022.

[6] P. Sotres, J. R. Santana, L. Sánchez, J. Lanza, and L. Muñoz, "Practical lessons from the deployment and management of a smart city internet-of-things infrastructure: The smartsantander testbed case," *IEEE Access*, vol. 5, pp. 14 309–14 322, 2017.

[7] R. Y. Wang and D. M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5—33, Mar. 1996.

[8] Y. Wand and R. Y. Wang, "Anchoring Data Quality Dimensions in Ontological Foundations," *Commun. ACM*, vol. 39, no. 11, pp. 86—-95, Nov. 1996.

[9] T. C. Redman, *Data Quality for the Information Age*, 1st ed. Artech House, Inc., 1997.

[10] C. Batini, C. Cappiello, C. Francalanci, and A. Maurino, "Methodologies for data quality assessment and improvement," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, p. 16, 7 2009.

[11] I. Taleb, M. A. Serhani, and R. Dssouli, "Big Data Quality: A Survey," in *2018 IEEE International Congress on Big Data (BigData Congress)*, 2018, pp. 166–173.

[12] L. Zhang, D. Jeong, and S. Lee, "Data Quality Management in the Internet of Things," *Sensors*, vol. 21, no. 17, p. 5834, Aug. 2021.

[13] S. Geisler, S. Weber, and C. Quix, "An Ontology-based Data Quality Framework for Data Stream Applications," in *ICIQ 2011 - 16th International Conference on Information Quality*, 01 2011.

[14] A. Klein and W. Lehner, "Representing Data Quality in Sensor Data Streaming Environments," *Journal of Data and Information Quality (JDIQ)*, vol. 1, no. 2, Sep. 2009.

[15] M. Gomez-Omella, B. Sierra, and S. Ferreiro, "On the Evaluation, Management and Improvement of Data Quality in Streaming Time Series," *IEEE Access*, vol. 10, pp. 81 458–81 475, 2022.

[16] C. Batini, D. Barone, M. Mastrella, A. Maurino, and C. Ruffini, "A Framework and a methodology for data quality assessment and monitoring," in *ICIQ 2007 - 12th International Conference on Information Quality*, Sep. 2007, pp. 333–346.

[17] F. de Haro Olmo, A. Valencia, A. Varela Vaca, and J. Alvarez-Bermejo, "Data Curation in the Internet of Things: a Decision Model approach," *Computational and Mathematical Methods*, vol. 3, Sep. 2021.

[18] L. Martín, L. Sánchez, J. Lanza, and P. Sotres, "Development and evaluation of artificial intelligence techniques for iot data quality assessment and curation," *Internet of Things*, vol. 22, p. 100779, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660523001026

[19] A. Karkouch, H. Mousannif, H. Al Moatassime, and T. Noel, "Data quality in internet of things: A state-of-the-art survey," *Journal of Network and Computer Applications*, vol. 73, pp. 57–81, 2016.

[20] "Technical Committee ISO/TC 69, "ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions"." [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:ed-1:v1:en

[21] H. Y. Teh, A. W. Kempa-Liehr, and K. I.-K. Wang, "Sensor data quality: a systematic review," *Journal of Big Data*, vol. 7, no. 1, Dec. 2020.

[22] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, "A survey on missing data in machine learning," *Journal of Big Data*, vol. 8, no. 1, pp. 1–37,, Dec. 2021.

[23] N. Y. Yen, J.-W. Chang, J.-Y. Liao, and Y.-M. Yong, "Analysis of interpolation algorithms for the missing values in IoT time series: a case of air quality in Taiwan," *The Journal of Supercomputing*, vol. 76, no. 8, pp. 6475–6500, Aug. 2020.

[24] R. Adhikari and R. K. Agrawal, "An Introductory Study on Time Series Modeling and Forecasting," Feb. 2013.

[25] "Context Information Management (CIM) ETSI Industry Specification Group (ISG), "NGSI-LD API," 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.04.01_60/gs_cim009v010401p.pdf

[26] Situation-Aware Linked heTerogeneous Enriched Data (SALTED), "D2.1: Report on Data Linking and Enrichment Architecture," Project Deliverable, 2022.

[27] "Smart Data Models – A global program led by FIWARE Foundation, TMForum, IUDX, and OASC." [Online]. Available: https://smartdatamodels.org/

[28] "DataQualityAssessment, Smart Data Model, GitHub Repository." [Online]. Available: https://github.com/smart-data-models/dataModel.DataQuality

[29] J. Von Neumann and S. Ulam, "Monte carlo method," *National Bureau of Standards Applied Mathematics Series*, vol. 12, no. 1951, p. 36, 1951.

[30] "Agencia Estatal de Meteorología - AEMET. Gobierno de España." [Online]. Available: https://www.aemet.es/es/portada

[31] L. Sanchez, L. Muñoz, J. A. Galache, P. Sotres, J. R. Santana, V. Gutierrez, R. Ramdhany, A. Gluhak, S. Krco, E. Theodoridis, and D. Pfisterer, "SmartSantander: IoT experimentation over a smart city testbed," *Computer Networks*, vol. 61, pp. 217–238, 2014, Special issue on Future Internet Testbeds -- Part I.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Experimentación realista sobre la red de acceso desagregada: diseño, implementación y validación de un eNodeB multi-split

Luis Diez[†], Cristian C. Erazo-Agredo[§], Mario Garza-Fabre[§], Javier Rubio-Loyola[§], Ramón Agüero[†]
[†]Departamento Ingeniería de Comunicaciones, Universidad de Cantabria, España
{ldiez, ramon}@tlmat.unican.es
[§]Centro de Investigación y Estudios Avanzados (CINVESTAV), Campus Tamaulipas, México
{cristian.erazo, mario.garza, javier.rubio}@cinvestav.mx

Gracias al uso de técnicas de redes definidas por software y de virtualización de funciones de red es posible habilitar la división funcional (*functional-split*) de las estaciones base, así como la centralización de un conjunto de las mismas, lo que permitiría fortalecer estrategias de colaboración entre elementos de acceso a la red. En este contexto, el 3GPP ha identificado ocho posibles divisiones funcionales, que distribuyen la pila de protocolos de las estaciones base entre las entidades *Central Unit* (CU) y *Distributed Unit* (DU). A pesar de la relevancia de la división funcional, a día de hoy, existen pocas soluciones que implementen dicha división, y en todos los casos el nivel de centralización es fijo y no puede modificarse. En este trabajo se describe una implementación experimental que permite la configuración de varias divisiones funcionales, habilitando la experimentación con hasta cuatro alternativas diferentes. La evaluación de la implementación demuestra su viabilidad práctica de acuerdo a los requisitos de retardo establecidos por el Small Cell Forum para todas las configuraciones, y por el 3GPP para dos de ellas.

***Palabras Clave*—desagregación, RAN, división funcional, implementación**

I. Introducción

Uno de los aspectos claves de las redes 5G, así como del futuro 6G, es el que implica cambios en su arquitectura, haciendo uso de técnicas de *Software Defined Networking* (SDN) y *Network Function Virtualization* (NFV). Estos nuevos paradigmas se están utilizando para realizar la virtualización, división y centralización de las funcionalidades tradicionalmente asociadas a las estaciones base. De este modo, algunas de estas funcionalidades pueden ser llevadas a nodos centrales, típicamente desplegados en servidores en la nube, dando lugar a la llamada *Cloud-RAN* (CRAN) y, más recientemente, a la *Virtual RAN* (VRAN). Este tipo de soluciones proporcionan varias ventajas, tales como la reducción de costes (tanto de despliegue como de operación), o la posibilidad de realizar una coordinación estrecha de los elementos de acceso a la red. Por otro lado, estas arquitecturas también presentan nuevos desafíos, entre los que destaca el cumplimiento de los estrictos requisitos de retardo entre las entidades virtualizadas.

En contraposición con las estaciones base tradicionales, que generalmente requieren *hardware* dedicado, en las soluciones VRAN las funciones se dividen entre las entidades virtualizadas *Central Unit* (CU) y *Distributed Unit* (DU). Las entidades CU se despliegan en servidores en la nube, mientras que las DU se ubican próximas a los elementos de radio-frecuencia, o *Radio Units* (RU). Así, dependiendo de las funcionalidades que se instancian en la CU y DU se definen varias configuraciones o divisiones funcionales (*functional-split*). En la Figura 1 se muestran las diferentes opciones identificadas por el 3GPP, de modo que para cada posible división las funcionalidades de la izquierda se ubicarían en la CU y el resto en la DU. Como se puede observar, las opciones con un índice más elevado implican un mayor nivel de centralización, lo que a su vez se traduce en requisitos más exigentes de capacidad de transmisión y latencia en la red *fronthaul*, que comunica la CU y la DU.

Esta flexibilidad en la arquitectura de la red de acceso presenta nuevos retos de investigación. El primero de ellos, relacionado con la planificación de red, es la ubicación de las CU, y la consecuente conectividad con la DU. Esta decisión se puede realizar en combinación con la configuración del nivel de centralización, que define las funciones ubicadas en cada entidad y, por tanto, modifica los requisitos. Este tipo de problemas se han abordado en algunos trabajos, tanto de forma aislada como combinada [1]. Además, algunos trabajos han propuesto adaptar de forma dinámica el nivel de centralización [2] de tal forma que pueda responder a cambios en el estado de la red, por

lo que esta funcionalidad se enmarcaría en las técnicas de operación y gestión de la red de acceso.

Tanto para el caso de técnicas de planificación como de operación de red se hace necesario el desarrollo de entornos de validación lo más realistas posible. Como respuesta concreta a esta necesidad, se presenta en este trabajo la implementación de una estación base experimental que permite configurar todas las divisiones funcionales que separan protocolos de acuerdo a la Figura 1: O_1, O_2, O_4, O_6. Para ello se ha partido de la implementación srsRAN 4G [3] que puede ser desplegada sobre dispositivos *Software Defined Radio* (SDR), de modo que la opción O_8 está disponible por defecto. La implementación presentada en este trabajo añade elementos intermedios (llamados *wrappers*) entre los protocolos existentes, de forma que se preserva la implementación original del proyecto srsRAN. En este documento no solo se presenta el diseño e implementación de la solución propuesta, sino que se describe su validación ante diferentes configuraciones. También se ha estudiado su rendimiento tanto en términos de retardo, para analizar su viabilidad, como de complejidad computacional (en términos de carga de CPU).

El resto del trabajo se estructura de la siguiente forma. En la Sección II se discute el trabajo relacionado con CRAN y el concepto de división funcional. Seguidamente, en la Sección III se presenta el diseño e implementación de la solución propuesta, mientras que su validación y caracterización se describen en la Sección IV. Finalmente, el documento concluye en la Sección V, en la que se describe el posible trabajo futuro que surge gracias a la solución que aquí se describe.

II. Trabajos relacionados

A pesar de la relevancia de la división funcional y la desagregación de los elementos de acceso en redes móviles, existen pocos trabajos en los que se describan implementaciones. En esta sección se resumirán los pocos ejemplos que se pueden encontrar en la literatura, resaltando cómo el enfoque aquí presentado las complementa.

Los autores de [4] presentan una prueba de concepto de las divisiones funcionales 2 y 5 sobre una red de acceso CRAN basada en TDM-PON, como tecnología que comunica CU y DU, y que usa el concepto de *Mobile-Central Office re-factored Datacenter* (M-CORD). En concreto, los autores presentan el análisis de rendimiento, sobre las divisiones funcionales mencionadas, de protocolos de transporte fiables y no-fiables (TCP y UDP) para los enlaces ascendente y descendente. En este análisis se observa que la configuración 5 permite alcanzar un *throughput* mayor a nivel de aplicación que la 2.

Alfadhli *et al.* presentan en [5] un análisis de la latencia experimentada con las divisiones funcionales 7, 8 y 9 (considera la división a nivel de funciones de radiofrecuencia). Los autores demuestran que la división 8, usada en la comunicación *fronthaul*, incrementa la ganancia de multiplexado estadístico, lo que tiene un impacto positivo en el retardo. En base a este análisis se concluye que la opción 9 reduce el retardo del *fronthaul*, lo que permitiría enlaces con mayor distancia en la comunicación entre la RU y la DU.

Desde otro punto de vista, Martinez-Alba *et al.* presentan en [2] una implementación como prueba de concepto de la división funcional flexible, permitiendo el cambio entre divisiones PDCP-RLC (opción 2) a MAC-PHY (opción 6), y vice-versa, en tiempo real y sin interrumpir el tráfico de usuario. El método que se propone admite tres posibles variaciones: *hard*, en la que se descarta el tráfico almacenado en los *buffers*; *soft*, en la que se asegura que todo el tráfico almacenado se traspasa; y configurable (*custom*), en la que se establecen tiempos máximos de traspaso. La implementación se ha validado de manera experimental considerando diferentes tasas de pérdida y latencia en las comunicaciones, lo que ha permitido verificar que la migración *hard* no induce retardo, pero genera pérdidas. Por el contrario, la opción *soft* produce retardos de hasta 15 ms sin incrementar la probabilidad de pérdida. Finalmente, la opción configurable permite establecer un compromiso entre retardo y pérdida, observándose que tienen una relación lineal.

Otro trabajo de interés es el presentado por Rodriguez *et al.* en [6], donde se analiza e implementa la división 7.3 (intra-PHY), con el objetivo de obtener un alto nivel de centralización, reduciendo los requisitos de capacidad de la red *fronthaul*. En este trabajo el rendimiento de esta configuración se compara con el obtenido usando las otras alternativas intra-PHY (opciones 7.1 y 7.2), demostrando que la opción 7.3 permite utilizar una mayor distancia entre la RU y DU, siempre y cuando sea posible realizar una implementación *multi-thread* de las funciones de codificación y decodificación, gracias a que presenta requisitos más laxos.

Por otro lado, los autores de [7] presentan un prototipo para analizar la influencia de las opciones de división funcional sobre el consumo energético. Otros estudios se han centrado en la reconfiguración de la red subyacente, que comunica las entidades virtualizadas, de modo que se puedan satisfacer los requisitos de cada división funcional. Por ejemplo, en [8], [9] se proponen y evalúan soluciones de reconfiguración de la red *fronthaul* basada en tecnología óptica, mientras que los autores de [10] realizan un trabajo similar, asumiendo comunicaciones basadas en Ethernet. También existen trabajos que combinan varios niveles de configuración, tales como el descrito en [11], donde se propone una solución de selección dinámica de *split* y reconfiguración de la red de transporte, de acuerdo a los recursos radio disponibles.

Asimismo, cabe destacar el esfuerzo conjunto, por parte de fabricantes, operadores y la comunidad investigadora, reflejado en la Open Radio Access Network (O-RAN) Alliance, para definir la arquitectura e interfaces abiertos de la red de acceso desagregada, flexible y virtualizada [12], [13]. En la arquitectura O-RAN se hace uso de las divisiones funcionales 2 y 7.2 (intra-PHY) y se definen entidades análogas a las mencionadas: O-RU, O-DU y O-CU (O-RAN RU, DU, y CU, respectivamente).

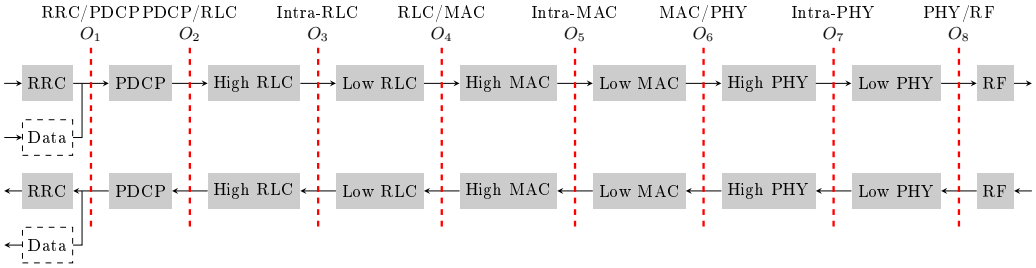


Fig. 1: Posibles divisiones funcionales entre la *CU* y *DU* identificadas por el 3GPP.

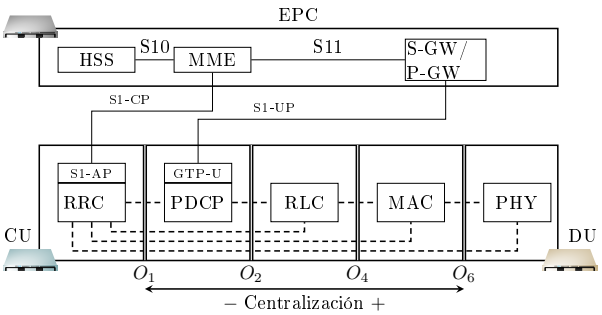


Fig. 2: Descripción funcional de la implementación

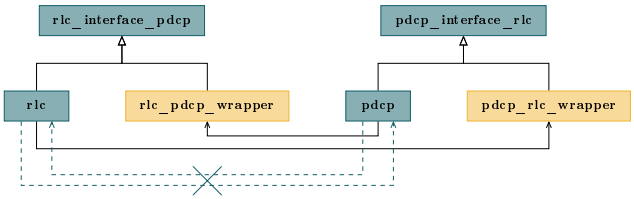


Fig. 3: Ejemplo de aplicación a la división funcional PDCP-RLC (*opción 2*)

El trabajo descrito en este documento complementa las implementaciones mencionadas, permitiendo la configuración de diferentes divisiones funcionales. Por otro lado, esta implementación está diseñada para realizar análisis y caracterizaciones en entorno de laboratorio, a diferencia de las definidas por O-RAN, que se centran en redes comerciales.

III. Diseño e implementación

En esta sección se describe la implementación realizada, sobre una solución de eNodeB 4G. En este sentido, cabe aclarar que el uso de la pila de protocolos 4G se debe a la carencia de una opción 5G en el momento de iniciar el trabajo. Sin embargo, los mismos principios de diseño se podrían aplicar a una implementación 5G. En general, las divisiones mencionadas harán referencia a la separación entre las entidades DU y CU, mientras que se asumirá que la RU es un nodo independiente que alberga sólo el interfaz radio. Sin embargo, algunas configuraciones de *split* pueden representar una separación entre DU y RU. Cabe destacar además que, como se ha mencionado anteriormente, la implementación descrita es experimental, por lo que la comunicación entre las entidades no sigue el estándar F1.

Como se ha mencionado anteriormente, la solución implementada se basa en el proyecto srsRAN 4G. En la Figura 2 se muestra una visión general de la implementación final, donde las líneas a puntos indican los interfaces incorporados, y las continuas los ya existentes en la implementación original. Como se puede observar, en la solución final existe una separación entre las capas de la pila de protocolos, así como entre ellas y la entidad RRC. De este modo, la CU siempre albergará la entidad RRC, mientras que la DU hará lo mismo con la entidad PHY. El resto de capas del protocolo pueden ser instanciadas tanto en la CU como en la DU, en función de la configuración en uso.

La implementación descrita permitiría configurar las siguientes divisiones, todas ellas definidas por el 3GPP [14]:

- *Opción 1* (O_1), en la que el RRC se sitúa en la CU, mientras que el resto de entidades estarían en la DU;
- *Opción 2* (O_2), donde la CU alberga a RRC y PDCP, y el resto se instancian en la DU. Esta división es la adoptada en la arquitectura O-RAN [15] para la separación CU-DU;
- *Opción 4* (O_4), que ubica en la CU las capas RRC, PDCP y RLC, mientras que las capas MAC y PHY residen en la DU; y
- *Opción 6* (O_6), en la que sólo la capa física (PHY) se encuentra en la DU. Esta configuración es la indicada por el Small Cell Forum [16] para la separación DU-RU.

De este modo, la implementación no limita la desagregación de la estación base en 2 entidades (CU y DU), sino que permitiría desplegar las capas de la pila de protocolos en varios nodos. Por ejemplo, se podría realizar una configuración CU-DU-RU usando las separaciones indicadas por el Small Cell Forum, donde la opción 2 se utiliza para separar la CU y DU, y la opción 6 haría lo propio con la DU y RU.

A continuación se describe el diseño *software* que se ha adoptado para modificar el código existente, y posteriormente se describe el flujo de comunicación e interacción entre los módulos implementados.

A. *Diseño* software

En primer lugar se describe el diseño orientado a objetos que se ha adoptado para realizar la implementación de todas las divisiones funcionales. De forma general, el objetio ha sido tener el menor impacto posible sobre el

código existente, de forma que se facilite su adaptación a nuevas versiones.

La pila de protocolos del eNodeB del proyecto srsRAN[1] está implementada en `C++`, donde cada uno de los protocolos se representa mediante una clase. La comunicación entre las clases se realiza a través de interfaces, como se muestra en la Figura 3, donde las clases originales se indican con diferentes tonos de verde, y las añadidas en el marco de este trabajo se indican en color amarillo. La figura también ilustra la herencia entre clases (flechas cerradas), la agregación (terminación en rombo) y la asociación (flechas abiertas).

Como se puede observar, en la implementación original la clase `rlc` hereda de la clase `rlc_interface_pdcp` (leído como interfaz RLC para ser usado por PDCP), que se define como una clase abstracta que indica los métodos que debe implementar el interfaz hacia PDCP. A su vez, la clase `pdcp` se inicializa con un puntero al interfaz hacia RLC (clase `rlc_interface_pdcp`), lo que permite separar la manera de interaccionar con la clase de su implementación. De la misma manera, la clase `rlc` se inicializa con un puntero a un interfaz `pdcp_interface_rlc`, para comunicarse con `pdcp`. En la figura también se muestra la asociación entre las clases `pdcp` y *rlc*, para indicar que existe una interacción entre ambas a través de sus interfaces correspondientes.

Se ha seguido el patrón de diseño *Adapter* para modificar la interacción entre las clases de la pila de protocolos con un impacto mínimo en el código existente. De este modo, para cada par de entidades se han añadido dos clases (llamadas *wrappers*), que interceptan la interacción entre las clases originales. Como se puede observar en la Figura 3, las nuevas clases, `pdcp_rlc_wrapper` y `rlc_pdcp_wrapper`, también heredan de los interfaces existentes. Se ha modificado la inicialización de las clases que instancian los protocolos, `rlc` y `pdcp`, de modo que se almacenen punteros a los interfaces que se correspondan a instancias de los *wrappers*, sin realizar ninguna modificación en las clases originales. Además, las clases *wrapper* también se inicializan con punteros de los dos interfaces involucrados en la comunicación entre capas, para implementar diferentes comportamientos.

Con todo, la clase `pdcp_rlc_wrapper` almacena un puntero a `pdcp_interface_rlc`, que realmente se corresponde con la instancia `pdcp`. Así, cuando la clase `rlc` realiza una llamada a la implementación de PDCP, esta se hace al *wrapper* `pdcp_rlc_wrapper` que, a su vez, la realiza a la instancia de `pdcp`, en ambos casos a través de punteros a `pdcp_interface_rlc`. Este comportamiento no proporciona ninguna ventaja respecto al original cuando las capas de la pila de protocolos se comunican localmente, pero permite interceptar la interacción entre capas. Los *wrappers* también implementan capacidades de comunicación, de modo que las llamadas interceptadas se puedan enviar a entidades remotas, como se describirá en detalle en la Sección B.

[1]https://www.srslte.com/

Así, los *wrappers* permiten recibir llamadas a funciones desde puntos remotos. Para ello, como se ha mencionado, almacenan en la inicialización punteros a interfaces de las dos clases involucradas. Por ejemplo, en la Figura 3 la clase `rlc_pdcp_wrapper` almacena punteros tanto de `pdcp_interface_rlc` como de `rlc_interface_pdcp`, que realmente son instancias de `pdcp` y `rlc`, respectivamente. Esto también se aplica a la clase `pdcp_rlc_wrapper`. Así, si un *wrapper* recibe una llamada remota a través de las funcionalidades de comunicaciones, esta se reenvía a la clase correspondiente. Con todo, las llamadas remotas de `rlc` a `pdcp` implicarían los siguientes pasos:

1) Primero, `rlc` llama a `pdcp_rlc_wrapper`, usando el puntero local a `pdcp_interface_rlc`.
2) Posteriormente, `pdcp_rlc_wrapper` se comunica con su par correspondiente, `rlc_pdcp_wrapper`.
3) Finalmente, `rlc_pdcp_wrapper` realiza la llamada a `pdcp`, usando el puntero local `pdcp_interface_rlc`.

Como se puede ver, las nuevas funcionalidades no requieren ninguna modificación en las implementaciones de las clases, ni de los protocolos, ni de los interfaces, sino que únicamente se modifica cómo estas se inicializan. Teniendo esto en cuenta, la configuración de una CU/DU consiste en indicar las capas que se instancian localmente, y la configuración local o remota de los *wrappers*. En este sentido, merece la pena resaltar que las clases *wrapper* pueden tener punteros no válidos (*dangling pointers*) a ciertos interfaces dependiendo de la configuración. Por ejemplo, la configuración de la CU con división PDCP/RLC (O_2 en la Figura 2) requeriría instanciar RRC y PDCP, así como configurar `rrc_pdcp_wrapper` y `pdcp_rrc_wrapper` para trabajar de forma local, mientras que los otros *wrappers* (`pdcp_rlc_wrapper`, `rrc_rlc_wrapper`, `rrc_mac_wrapper`, y `rrc_phy_wrapper`) operarían de forma remota. Al mismo tiempo, la DU tendría una configuración complementaria (especular): se instanciarían las entidades RLC, MAC y PHY usando los *wrappers* que las comunican (`rlc_mac_wrapper`, `mac_rlc_wrapper`, `mac_phy_wrapper` y `phy_mac_wrapper`) de forma local, y el resto de forma remota.

La Figura 4 muestra el diagrama de clases de la implementación sobre la pila de protocolos completa. Por simplicidad no se incluyen las variables miembro de las clases. Como se puede observar se repite el mismo patrón para cada par de protocolos adyacentes de la pila, así como entre estos y RRC.

B. Flujo de ejecución de la división funcional

Como se ha explicado, la implementación sigue el patrón de diseño *Adapter*, añadiendo clases encargadas de implementar la funcionalidad de separación. De este modo, cada vez que una capa (llamante), `X`, llama a una función de otra capa (llamada) `Y` usando el interfaz

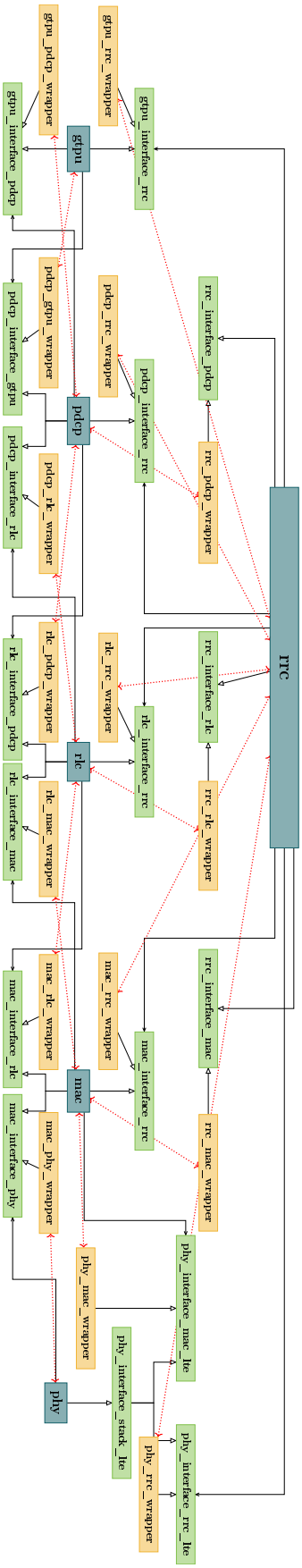


Fig. 4: Diagrama de clases de la implementación. La relación de herencia se representa con flechas cerradas, asociaciones con flechas abiertas y relaciones generales con flechas abiertas y líneas punteadas

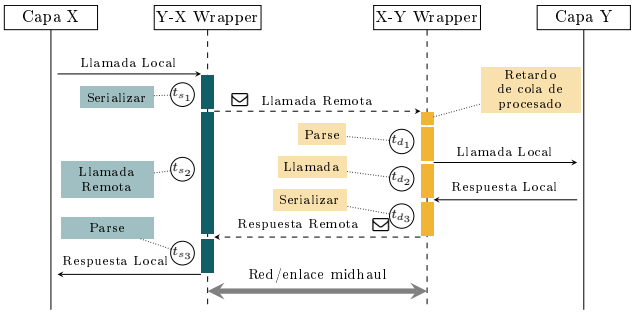


Fig. 5: Secuencia de llamadas de la implementación e identificación de fuentes de retardo.

correspondiente, se llama a la clase *wrapper* involucrada, que comprueba la configuración y decide si la llamada es local o remota. Cuando las instancias `X` e `Y` están en la misma entidad de red (CU o DU), la llamada a la función se re-envía usando la implementación preexistente; en caso contrario, la llamada se envía a un *wrapper* remoto, encargado de hacérsela llegar a la clase correspondiente. En la Figura 5 se muestra la secuencia de interacciones en el caso de llamadas remotas.

Como se puede ver, el *wrapper* en el lado llamante (`X-Y wrapper`) intercepta la llamada a la función y serializa la información requerida para reproducir los argumentos de entrada en el otro extremo. Una vez alcanza a su destino, la información serializada se des-serializa y se realiza la llamada a la función correspondiente de la clase llamada `Y`. Además, en caso de que la función devuelva algún dato (ya sea valor de retorno o a través de parámetros de entrada/salida), estos se envían de vuelta al lado llamante, de forma que el *wrapper* correspondiente se lo entregue a la clase `X` que inició la interacción.

Como se muestra en la Figura 5, se distinguen diferentes periodos de tiempo durante la interacción. En el lado llamante, t_{s1} corresponde con el tiempo de serialización. En el estado actual de la implementación, la mayor parte de la serialización se realiza mediante la copia directa de memoria de las variables involucradas. Sin embargo, hay algunos casos en los que intervienen clases o estructuras anidadas, para las que la serialización y des-serialización es más compleja. En esos casos se ha hecho uso de la implementación de `C++` de *ProtoBuffer*, ya que simplifica el proceso. Tras la serialización, el intervalo t_{s2} indica el tiempo transcurrido desde el inicio del envío de los datos hasta que se recibe el valor devuelto. De forma general, este segundo intervalo incluye transmisión en la red *fronthaul/midhaul*, tiempo de espera, y tiempo de procesado. En este sentido, dependiendo del esquema adoptado, las llamadas a las funciones podrían tener que ser almacenadas a la espera de ser procesadas. Finalmente, el intervalo t_{s3} representa el tiempo que se requiere para procesar los datos devueltos y enviarlos a la clase llamante. Este último tiempo sólo existirá en llamadas a funciones que devuelvan datos.

En el lado del destino, los intervalos t_{d1} y t_{d3} indican los tiempos necesarios para realizar el *parsing* de los datos,

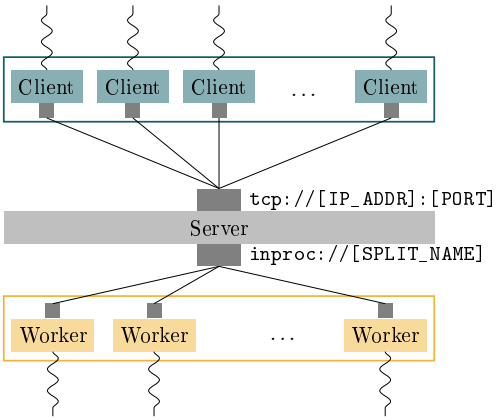


Fig. 6: Diseño del procesado multi-hilo

y la serialización de los valores devueltos. Finalmente, t_{d2} es el tiempo consumido para ejecutar la función original. Como se puede ver, el tiempo t_{s2} no sólo se corresponde con la suma de los tiempos en el receptor (t_{d1}, t_{d2} y t_d3), sino que también incluye retardos en las comunicaciones. Con todo, la eficiencia de la implementación vendría dada por el cociente entre t_{d2} y el tiempo total para realizar la llamada, de modo que cuanto mayor sea la contribución de t_{d2} al retardo total, menor es el impacto de la funcionalidad añadida.

Como protocolo de transmisión se ha usado *ZeroMQ* [17], por diferentes razones. En primer lugar, la librería que permite su uso está incluida en el proyecto srsRAN para emular la comunicación de radio-frecuencia sin necesidad de dispositivos *hardware*, de modo que no hay problemas de integración. Además, *ZeroMQ* se encarga de gestionar el establecimiento de las conexiones, por lo que no es necesario tomar precauciones respecto al orden de instanciación de los *wrappers*. Finalmente, *ZeroMQ* proporciona un conjunto de patrones para computación remota en paralelo, lo que facilita la implementación de la llamada a funciones.

Concretamente, se ha adoptado el patrón *Aysnchronous Client/Server Pattern*, que se muestra en la Figura 6 para una implementación *multi-thread*. Con este esquema, una serie de clientes envían tareas a un servidor común de forma asíncrona. A su vez, el servidor asigna estas tareas entre un conjunto de *workers* locales y envía la respuesta de vuelta a los clientes. La comunicación entre los clientes y el servidor se realiza sobre conexiones TCP, mientras que el servidor usa comunicaciones *inter-trhead* (`inproc`) para interactuar con los *workers*.

IV. Validación y resultados

En esta sección se presenta la validación de la implementación presentada en este trabajo. A fin de simplificar el análisis, todos los resultados se han obtenido emulando las comunicaciones de radio-frecuencia entre el usuario (UE) y la estación base. En primer lugar se analizará el retardo inducido por la implementación. Posteriormente, se usará el entorno desarrollado para la caracterización de la complejidad computacional de la CU ante diferentes configuraciones. En todos los casos los resultados se han obtenido con dos servidores HP ProLiant ML310e Gen8 v2, equipados con 8 CPUs Intel Xeon E3-1241 v3 (4 cores y 2 threads por core) que trabaja a 3.5 `GHz`. Los servidores se comunican mediante un switch Qnap QSW-M5216-1T, que pertecene a la famila 25 `GB` Ethernet, lo que asegura que la red no tiene impacto en la medida de los retardos. Además de una estación base, la configuración incluye el UE y el EPC. En cualquier configuración con división funcional un servidor alberga al UE y al DU, mientras que la CU y el EPC se instancian en el otro. Si no se configura ninguna división, el UE se ejecuta en un servidor y la estación base y el EPC en el otro.

En todos los casos, se envía tráfico en el enlace descendente durante 60 segundos, estableciendo la tasa en función del tamaño de la celda (número de PRBs). Para ello, se ha caracterizado la tasa máxima que se alcanza con la configuración sin división funcional. Con tamaños de celda de 15 y 25 PRBs, se alcanzan 10 y 50 `Mbps` respectivamente. Para tamaños mayores la tasa máxima es de 100 `Mbps`.

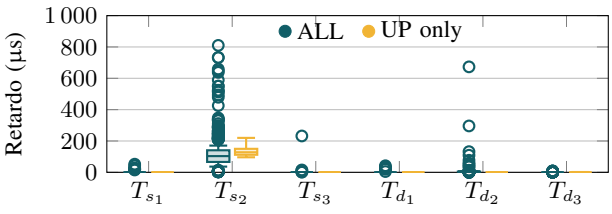
A. Retardos de las divisiones

La Figura 7 muestra la distribución estadística de los intervalos de tiempo indicados en la Figura 5. Los resultados se han obtenido para las diferentes divisiones funcionales y un tamaño de celda de 100 PRBs, con una carga de tráfico que genera situación de saturación. En todos los casos se diferencia entre el retardo de todo el tráfico, y el inducido únicamente sobre los datos de usuario.

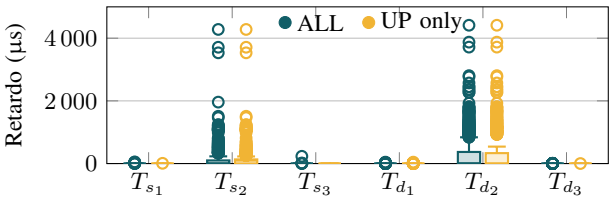
Como se puede observar en la Figura 7a, el retardo inducido con la opción O_1 está por debajo de 800 µs en todos los casos. Además, se puede ver que el retardo en todos los intervalos es prácticamente despreciable, excepto en T_{s_2}, lo que evidencia que la serialización y des-serialización no tienen un impacto significativo en la sobrecarga. Por otro lado T_{s_2} contiene los retardos de comunicaciones, encolado para el procesado, así como los retardos en el receptor (T_{d_1}, T_{d_2} y T_{d_3}). De acuerdo con los resultados, los retardos en el receptor son muy bajos, por lo que los valores de T_{s_2} se corresponden principalmente con el propio diseño de comunicaciones y gestión de procesado (server/workers), por lo que este sería el retardo mínimo que se podría esperar en el escenario configurado.

La Figura 7b muestra un comportamiento diferente para la división O_2. Nuevamente el mayor valor es el de T_{s_2}, pero en este caso se puede observar que esto se debe a la propia llamada a la función (T_{d_2}). En la Figura 7c se observa un comportamiento similar al analizar la opción O_4. En el caso del split MAC/PHY mostrado en la Figura 7d se puede observar un comportamiento más cercano al del RRC/PDCP: el tiempo T_{d_2} es prácticamente 0, por lo que los retardos observados en T_{s_2} serían consecuencia del diseño. Finalmente, se puede observar que no hay un diferencia apreciable entre el retardo del tráfico en conjunto y el observado para el plano de usuario.

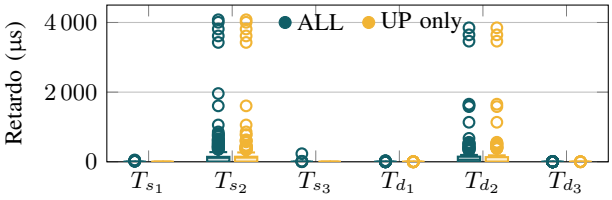
Para concluir la evaluación del retardo, en la Tabla I se muestra la viabilidad de la implementación en com-



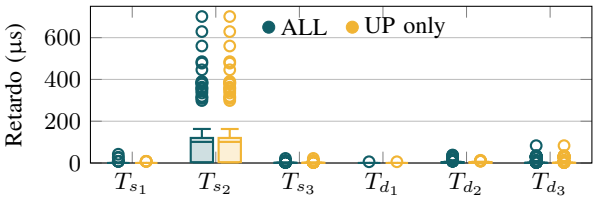
(a) Distribución de retardos en división O_1 - RRC/PDCP



(b) Distribución de retardos en división O_2 - PDCP/RLC



(c) Distribución de retardos en división O_4 - RLC/MAC



(d) Distribución de retardos en división O_6 - MAC/PHY

Fig. 7: Distribución de retardos en cada periodo. Se muestran tanto los retardos en conjunto como los medidos sólo sobre el tráfico del plano de datos de usuario

Tabla I: Viabilidad de la implementación en términos de retardo

	O_1	O_2	O_4	O_6
SCF [16]	30ms ✓	30ms ✓	6ms ✓	(ideal) 250µs ✗ (near-ideal) 2ms ✓
3GPP [14]	10ms ✓	10ms ✓	100µs ✗	250µs ✗

paración con los requisitos propuestos por diferentes organizaciones: Small Cell Forum (SCF) Release 7.0 [16, Section 2.8] y 3GPP Technical Report 38001 [14, Annex A]. En la tabla se indica si el requisito se cumple usando el mayor valor de retardo observado en la Figura 7. Como se puede ver, la solución propuesta cumpliría con los requisitos de las divisiones O_1 y O_2. Asimismo se respetarían las restricciones definidas por el SCF para las opciones O_4 y O_6, con condiciones casi-ideales. Por otro lado, no se cumpliría con los requisitos establecidos por el 3GPP para estaciones base macro usando las opciones O_4 y O_6.

B. Complejidad computacional

Teniendo en cuenta la dificultad de medida de la complejidad computacional de servicios, se ha adoptado una métrica sintética definida en [18] que se basa en el uso de la CPU. Esta métrica define la complejidad computacional en Giga Operaciones Por Segundo (GOPS) como sigue:

$$R = N_{cores} \cdot C \cdot N_{flop} \cdot \mu \tag{1}$$

donde N_{cores} y N_{flop} es el número de cores y flops por ciclo respectivamente; C es la frecuencia nominal del procesador y μ es la utilización del mismo. Esta medida se ha obtenido monitorizando la evolución temporal de la CPU usando la herramienta `top` [2]. En concreto, se ha tomado una muestra cada 100ms durante los 60 de transmisión con cada configuración.

La Figura 8 muestra la distribución de este parámetro con varias configuraciones. En este caso también se incluyen los valores cuando no se divide la estación base. En general se puede observar una tendencia creciente al realizar una mayor centralización, alcanzando los 1000 GOPS con la opción O_1, y los 3000 GOPS de la estación base completa.

Si se presta atención a la división RRC/PDCP, la Figura 8a muestra que el tamaño de la celda tiene un impacto muy limitado. Por otro lado, la influencia de la carga está relacionada con el tamaño de la BS. Con celdas pequeñas se observa un efecto de saturación para la complejidad computacional al 60% de carga, mientras que con celdas mayores este punto se alcanza al 40%.

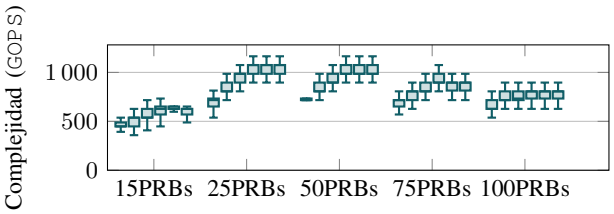
En el resto de las configuraciones se puede ver que el tamaño de celda tiene poco impacto por encima de los 50 PRBs. Además, se puede observar que en las divisiones RLC/PDCP y RLC/MAC la carga de la celda tiene un impacto elevado en la complejidad computacional para todos los tamaños de celda. Por otro lado, en las configuraciones MAC/PHY y sin división la influencia de la carga en la complejidad computacional es menor.

V. Conclusiones

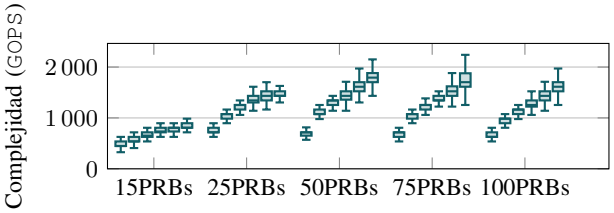
En este trabajo se ha descrito y evaluado una implementación experimental de estación base, tomando como punto de partida el proyecto srsRAN, que permite la configuración de hasta 4 divisiones funcionales. En concreto, la implementación descrita en este documento permite configurar las divisiones RRC/PDCP, PDCP/RLC, RLC/MAC y MAC/PHY, mientras que la opción PHY/RF está intrínsecamente incluida como parte de la solución SDR.

En primer lugar se ha descrito el diseño *software* que se ha planteado, indicando los cambios realizados sobre el proyecto original, con el objetivo de tener el menor impacto posible en el código existente, lo que simplifica su adopción en versiones sucesivas. Asimismo, se ha descrito el diseño de comunicación y procesado entre entidades de la estación base.

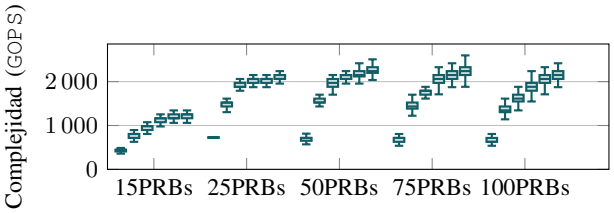
[2]https://man7.org/linux/man-pages/man1/top.1.html



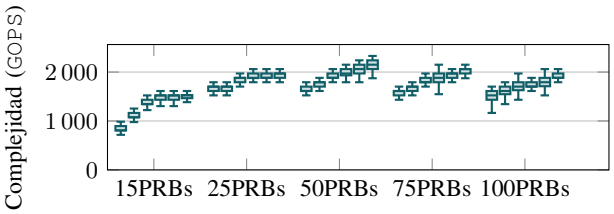
(a) O_1 - división RRC/PDCP



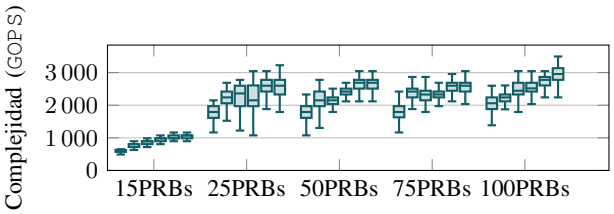
(b) O_2 - división RLC/PDCP split



(c) O_4 - división RLC/MAC split



(d) O_6 - división MAC/PHY split



(e) No-división

Fig. 8: Uso de CPU en la *CU* para diferentes divisiones, tamaños de celda y carga de tráfico: $\{0, 20, 40, 60, 80, 100\}\%$.

Seguidamente, se ha analizado el rendimiento de la implementación en términos de retardo y complejidad computacional. Los resultados permiten validar la viabilidad de la implementación, de acuerdo a los requisitos establecidos por el SCF para celdas pequeñas. Por otro lado, en el caso de macro-celdas sólo se cumple con los requisitos establecidos por el 3GPP para las opciones RRC/PDCP y PDCP/RLC. Como se podría esperar, los resultados muestran un incremento de la complejidad computacional al incrementar la centralización, mientras que el impacto de otros parámetros depende fuertemente del *split* utilizado.

El el futuro se analizará el uso de la solución software propuesta sobre la implementación 5G de srsRAN, que ya implementa las divisiones definidas por O-RAN. Por otro lado, se estudiará la posibilidad de definir arquitecturas en las que varias DUs estén conectadas a una misma CU.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, MINECO-FEDER, a través del proyecto SITED: *Semantically-enabled Interoperable Trustworthy Enriched Data-spaces (PID2021-125725OB-I00).*

REFERENCIAS

[1] C. C. Erazo-Agredo, M. Garza-Fabre, R. A. Calvo, L. Diez, J. Serrat, and J. Rubio-Loyola, "Joint route selection and split level management for 5g c-ran," *IEEE Transactions on Network and Service Management*, vol. 18, no. 4, pp. 4616–4638, 2021.

[2] A. M. Alba, J. H. G. Velásquez, and W. Kellerer, "An adaptive functional split in 5g networks," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2019, pp. 410–416.

[3] I. Gomez-Miguelez, A. Garcia-Saavedra, P. D. Sutton, P. Serrano, C. Cano, and D. J. Leith, "Srslte: An open-source platform for lte evolution and experimentation," in *Proceedings of the Tenth ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation, and Characterization*, ser. WiNTECH '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 25–32. [Online]. Available: https://doi.org/10.1145/2980159.2980163

[4] R. K. Saha, Y. Tsukamoto, S. Nanba, K. Nishimura, and K. Yamazaki, "Novel m-cord based multi-functional split enabled virtualized cloud ran testbed with ideal fronthaul," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–7.

[5] Y. Alfadhli, Y.-W. Chen, S. Liu, S. Shen, S. Yao, D. Guidotti, S. Mitani, and G.-K. Chang, "Latency performance analysis of low layers function split for urllc applications in 5g networks," *Computer Networks*, vol. 162, p. 106865, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128619301343

[6] V. Q. Rodriguez, F. Guillemin, A. Ferrieux, and L. Thomas, "Cloud-ran functional split for an efficient fronthaul network," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, 2020, pp. 245–250.

[7] N. Bartzoudis, O. Font-Bach, M. Miozzo, C. Donato, P. Harbanau, M. Requena, D. López, I. Ucar, A. A. Saloña, P. Serrano, J. Mangues, and M. Payaró, "Energy footprint reduction in 5g reconfigurable hotspots via function partitioning and bandwidth adaptation," in *2017 Fifth International Workshop on Cloud Technologies and Energy Efficiency in Mobile Communication Networks (CLEEN)*, 2017.

[8] Y. Alfadhli, M. Xu, S. Liu, F. Lu, P.-C. Peng, and G.-K. Chang, "Real-time demonstration of adaptive functional split in 5g flexible mobile fronthaul networks," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, 2018.

[9] P. Monti, Y. Li, J. Mårtensson, M. Fiorani, B. Skubic, Z. Ghebretensaé, and L. Wosinska, "A flexible 5g ran architecture with dynamic baseband split distribution and configurable optical transport," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*, 2017.

[10] C.-Y. Chang, N. Nikaein, R. Knopp, T. Spyropoulos, and S. S. Kumar, "Flexcran: A flexible functional split framework over ethernet fronthaul in cloud-ran," in *2017 IEEE International Conference on Communications (ICC)*, 2017.

[11] Y. Li, J. Martensson, B. Skubic, Y. Zhao, J. Zhang, L. Wosinska, and P. Monti, "Flexible ran: Combining dynamic baseband split selection and reconfigurable optical transport to optimize ran performance," *IEEE Network*, vol. 34, no. 4, pp. 180–187, 2020.

[12] S. K. Singh, R. Singh, and B. Kumbhani, "The evolution of radio access network towards open-ran: Challenges and opportunities," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2020, pp. 1–6.

[13] T. D. Tran, K.-K. Nguyen, and M. Cheriet, "Joint route selection and content caching in o-ran architecture," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 2250–2255.
[14] 3rd Generation Partnership Project (3GPP), "Study on new radio access technology: Radio access architecture and interfaces," 2017.
[15] O-RAn Working Group 1 (Use Cases and Overall Architecture), "O-RAN Architecture Description," 2023.
[16] Small Cell Forum), "S-RU and S-DU Test Support," 2021. [Online]. Available: {https://scf.io/en/documents/228_S-RU_and_S-DU_Test_Support.php}
[17] P. Hintjens. (2011) 0mq - the guide. [Online]. Available: http://zguide.zeromq.org/page:all
[18] J. K. Chaudhary, A. Kumar, J. Bartelt, and G. Fettweis, "C-ran employing xran functional split: Complexity analysis for 5g nr remote radio unit," in *2019 European Conference on Networks and Communications (EuCNC)*, 2019, pp. 580–585.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Explorando los L-momentos de orden superior en el análisis y clasificación de flujos de red

Jesús Galeano-Brajones[1], Mihaela I. Chidean[2], Francisco Luna[3,4], Javier Carmona-Murillo[1]
[1]Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos, Universidad de Extremadura, Mérida, 06800
[2]Dpto. de Teoría de la Señal y Comunicaciones, Universidad Rey Juan Carlos, Fuenlabrada, 28942
[3]Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga, E.T.S.I. Informática, Málaga, 29071
[4]ITIS Software, Universidad de Málaga, Málaga, 29071
jgaleanobra@unex.es, mihaela.chidean@urjc.es, flv@lcc.uma.es, jcarmur@unex.es

El despliegue de redes 5G ha llevado a la comunidad investigadora y a la industria a establecer cuáles serán los pilares de la nueva generación de redes móviles. Entre estos pilares se incluyen los sistemas autónomos que proporcionan inteligencia a la red en tareas de gestión y seguridad, como son, entre otros, los mecanismos para caracterizar el tráfico de red y realizar la toma de decisiones. En este contexto, este trabajo explora el uso de L-momentos de orden superior a 3 en el análisis y clasificación de tráfico de red. Basándonos en la metodología propuesta en trabajos previos, este artículo amplía la comparativa directa de un nuevo L-momento de orden superior. Los resultados obtenidos en este caso mejoran la precisión de la clasificación de los flujos, siendo de hasta un 4,44% superior en el conjunto de datos analizado.

***Palabras Clave*—Análisis de tráfico, L-momentos, Gestión inteligente, Machine Learning, 6G**

I. INTRODUCCIÓN

El gran salto hacia la gestión inteligente de redes se debe principalmente a las tecnologías habilitadoras de 5G, como las técnicas de virtualización, la gestión de red basada en software y el *slicing*, entre otras. Estas técnicas permiten virtualizar servicios y aplicaciones en la red de manera que para su gestión se puedan desplegar sistemas inteligentes en forma de aplicaciones. Además, el aumento masivo y continuo del tráfico de red hace que sea aún más esencial su análisis y clasificación. La sexta generación (6G) aún está lejos de estar completamente definida, pero la comunidad investigadora está de acuerdo en que estas tecnologías seguirán siendo cruciales. Además, a medida que los sistemas inteligentes evolucionan hacia la autogestión de redes, la inteligencia artificial (IA) se vuelve mucho más importante para redes de próxima generación, siendo la característica clave de las redes autónomas 6G [1].

La gestión de la red y de los servicios en redes de próxima generación se espera que sea completamente autónoma [2]. Para conseguirlo, las redes deben seguir el paradigma *Zero-touch network and Service Management* (ZSM) definido por el Instituto Europeo de Estándares de Telecomunicación (ETSI) [3], con el que se pretende integrar la IA directamente en la red como una tecnología clave, soportada por la gestión basada en software y la virtualización. En este sentido, las redes serán capaces de autogestionarse, tomando decisiones sin la necesidad de intervención humana [4], y optimizando así los gastos de capital y de explotación [5]. Para lograr esta automatización, el análisis del tráfico de red y las técnicas de clasificación son cruciales para proveer a la red con información relevante que la guíe hacia decisiones precisas.

En este escenario, el análisis del tráfico de red es un tema que despierta gran interés en la literatura, especialmente desde la perspectiva de la seguridad. En este área se han empleado diferentes técnicas para la detección de anomalías mediante el análisis del tráfico y flujos de red: (i) el análisis basado en puertos es la técnica más simple, pero ya no es útil debido a la gran proliferación de nuevos servicios y aplicaciones que no hacen uso de puertos IANA bien definidos [6]; (ii) la inspección profunda de paquetes (DPI, *Deep Packet Inspection*) emerge como una alternativa, pero sus mayores limitaciones relativas a la aplicación en paquetes no encriptados y el problema subyacente con la privacidad de los usuarios, ha propiciado la propuesta de mecanismos de Machine Learning (ML) o Deep Learning (DL) para mitigar esas desventajas [7]; (iii) las técnicas basadas en el mensaje sólo utilizan la información contenida en él de la capa de aplicación, y normalmente se despliega junto a DPI [6]; (iv) las propuestas basadas en mecanismos estadísticos utilizan parámetros independientes de la información contenida en el paquete (por ejemplo, duración del flujo, tiempo entre paquetes, longitud de la cabecera, etc.) que pueden ser

utilizados como entrada de diversos modelos estadísticos, de ML o DL. Finalmente, en los últimos años, ha habido un gran auge de los modelos de DL aplicados a la clasificación de tráfico de red [8].

Los L-momentos han sido ampliamente utilizados en diferentes áreas de investigación desde su propuesta en 1990 [9]. El área con más aplicaciones es el análisis de datos climáticos, especialmente el análisis regional [10]. Algunos ejemplos específicos incluyen el modelado de funciones de distribución de probabilidad de viento y precipitación [11]. Además, los L-momentos han sido también utilizados en otras áreas como la clasificación de objetivos en aplicaciones de radar [12] y en el contexto de la teoría de redes complejas [13]. Algunos ejemplos adicionales incluyen datos financieros y análisis de acciones [14], [15], disciplinas de fiabilidad [16], modelado matemático de procesos mecánicos [17] o datos médicos [18]. Además, los L-momentos de orden superior, como el L-momento estándar τ_6, también han sido utilizados en la literatura [19]. Finalmente, propusimos una metodología basada en los L-momentos para el análisis de flujos [20], validada con conjuntos de datos para la detección de anomalías, donde utilizamos los L-momentos estándar τ_3 y τ_4 para la clasificación de flujos.

En este contexto, este artículo explora las capacidades de los L-momentos estándar de orden superior en la metodología de análisis y clasificación de flujos propuesta en el trabajo previo [20]. El diagrama de L-momentos estándar con τ_3 y τ_4 ha demostrado ser una herramienta especialmente útil para el análisis de los flujos de red, por lo que en este artículo comparamos los beneficios y desventajas del uso de τ_5 en las clasificaciones de modelos ML básicos. Para tener resultados directamente comparables, utilizamos la misma base de datos que en el artículo previo, el CIC-DDoS2019 *dataset* [21], que contiene escenarios con diferentes ataques de denegación de servicio distribuido (DDoS, *Distributed Denial of Service*) y de denegación de servicio distribuido reflejado (DrDoS, *Distributed Reflection/Reflective Denial of Service*) realistas y actualizados.

Este trabajo está estructurado de la siguiente manera. La Sección II se centra en la base teórica de la metodología utilizada. Después, la Sección III muestra los resultados obtenidos de la experimentación con τ_5. Por último, en la Sección IV se exponen las conclusiones y los trabajos futuros.

II. METODOLOGÍA

En esta sección se describe brevemente la metodología propuesta en [20] y cómo se ha ampliado en este trabajo. Para más detalles, se recomienda al lector que acuda a dicho artículo.

A. L-momentos

Al igual que otros momentos estadísticos, los L-momentos caracterizan la geometría de distribuciones y resumen muestras. Son directamente análogos, es decir, tienen interpretación similar, a los momentos centrales (llamados *product moments* o *C-moments* en la literatura [10]). Los L-momentos son combinaciones lineales de diferencias de esperanzas de estadísticos de orden. Esta es la principal diferencia con los momentos centrales, los cuales están basados en potencias de diferencias con la media. Algunos beneficios de los L-momentos frente a los momentos centrales son:

- Los L-momentos son más robustos ante la presencia de datos atípicos, es decir, sufren menos por los efectos de la variabilidad de las muestras.
- Necesitan menos tamaño muestral para estimar con errores bajos, lo que los hace muy útiles para estimaciones en tiempo real y para trabajar con L-momentos de orden alto.
- A diferencia de los momentos centrales, los L-momentos son no sesgados, es decir, no dependen del tamaño de la muestra.
- Están más cerca de su distribución normal asintótica en muestras finitas.

Los estadísticos de orden de una variable aleatoria X para una muestra de tamaño n están formados por el orden ascendente $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$. Los L-momentos teóricos quedan definidos por:

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E[X_{r-k:r}], \quad \forall\ r \geq 1 \tag{1}$$

donde r es el orden del L-momento y $E[X_{r-k:r}]$ es la esperanza del estadístico de orden $r-k$ de una muestra de tamaño r. λ_1 se denomina L-ubicación (*L-location*) y representa la media de la distribución. λ_2 es la L-dispersión (*L-scale*) y representa una medida de la variabilidad de la distribución. Además, las distribuciones útiles tienen una variabilidad distinta de cero, por lo que $\lambda_2 > 0$. Con respecto a λ_2 se definen los momentos estándar teóricos:

$$\tau_r = \lambda_r/\lambda_2, \quad \forall\ r \geq 3 \tag{2}$$

Por analogía a la los momentos centrales en cuanto a la interpretación y también por comodidad, algunos momentos estándar tienen un nombre definido:

$$\begin{aligned} \tau_2 &= \lambda_2/\lambda_1 = \text{coeficiente de L-variación} \\ \tau_3 &= \lambda_3/\lambda_2 = \text{L-asimetría (}\textit{L-skewness}\text{)} \\ \tau_4 &= \lambda_4/\lambda_2 = \text{L-curtosis (}\textit{L-kurtosis}\text{)} \\ \tau_5 &= \lambda_5/\lambda_2 \end{aligned} \tag{3}$$

τ_2 es útil para variables aleatorias positivas ($X \geq 0$) y está acotado en $0 < \tau_2 < 1$. Para $r \geq 3$, todos los L-momentos estándar están acotados en $-1 < \tau_r < 1$. Y, específicamente, las cotas de la L-curtosis son $\frac{1}{4}\left(5\tau_3^2 - 1\right) \leq \tau_4 < 1$. Estas cotas añaden utilidad a los L-momentos estándar dado que permiten comparar distribuciones cuyos rangos son diferentes sin la necesidad de normalizar o reescalar.

Los L-momentos muestrales se calculan para una muestra de estadísticos de orden $x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n}$. Los L-momentos muestrales se definen como:

$$\hat{\lambda}_r = \tfrac{1}{r}\tbinom{n}{r}^{-1} \sum_{i=1}^{n}\left[\sum_{j=0}^{r-1}(-1)^j \binom{r-1}{j}\binom{i-1}{r-1-j}\binom{n-i}{j}\right] x_{i:n}, \quad \forall\ r \geq 1 \tag{4}$$

Y los L-momentos estándar muestrales son:

$$\hat{\tau}_r = \hat{\lambda}_r/\hat{\lambda}_2, \quad \forall\ r \geq 3 \tag{5}$$

$$\begin{aligned}
&\hat{\tau}_2 = \hat{\lambda}_2/\hat{\lambda}_1 = \text{coeficiente de L-variación muestral} \\
&\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2 = \text{L-asimetría muestral} \\
&\hat{\tau}_4 = \hat{\lambda}_4/\hat{\lambda}_2 = \text{L-curtosis muestral} \\
&\hat{\tau}_5 = \hat{\lambda}_5/\hat{\lambda}_2
\end{aligned} \tag{6}$$

B. Framework

El marco de trabajo utilizado en este artículo se puede ver en la Fig. 1 de [20]. La diferencia entre [20] y este trabajo radica en la tipología concreta del diagrama de L-momentos estándar empleado. Mientras en [20] el trabajo se basa en la tupla ampliamente utilizada τ_3-τ_4, en este trabajo vamos un paso más allá incorporando τ_5 y las posibles combinaciones, es decir, τ_3-τ_5, τ_4-τ_5 y τ_3-τ_4-τ_5. Este añadido nos aporta mayor información sobre el comportamiento estadístico de los ataques al mismo tiempo que genera más opciones de clasificación.

C. CIC-DDoS2019 dataset

Existen múltiples bases de datos de tráfico de red disponibles para la comunidad investigadora, cada uno considerando escenarios y aplicaciones específicas, e incluso con variedad de ataques DDoS. Como los ataques evolucionan continuamente y presentan nuevos desafíos, se crean nuevas bases de datos que contienen la información más actualizada. En este trabajo, dado que buscamos realizar una comparativa directa con [20], utilizamos también la base de datos CIC-DDoS2019 [21]. Además, esta es la base de datos que actualmente es considerada como la referencia para cualquier trabajo que analice las amenazas de red, específicamente las amenazas DDoS y DrDoS. Esta base de datos ha sido generado por el Instituto Canadiense de Ciberseguridad (CIC) con el objetivo de remediar todas las deficiencias previas relacionadas con estos tipos de ataques y contiene flujos de tráfico pertenecientes a diferentes tipos de ataques DDoS y DrDoS que se asemejan a los datos reales.

Además del tráfico capturado, el CIC también proporciona archivos CSV etiquetados generados por la herramienta CICFlowMeter-V3. Las características de flujo obtenidas por esta herramienta se basan en la marca temporal, las direcciones IP, puertos de origen y destino, los protocolos, los paquetes, el tiempo de de llegada entre paquetes, etc. El conjunto de datos CIC-DDoS2019 contiene un comportamiento abstracto de 25 usuarios que utilizan los protocolos HTTP, HTTPS, FTP, SSH y correo electrónico. Incluye flujos de red y archivos CSV para 10 ataques DrDoS y 12 ataques DDoS capturados en dos días.

III. EXPERIMENTACIÓN Y RESULTADOS

Al buscar una comparativa en igualdad de condiciones con [20], se emplean los mismos algoritmos ML: kNN (*k-Nearest Neighbors*) con función de pesos uniforme ("kNN-unif") y con función basada en distancia ("kNN-dist"), y SVM (*Support Vector Machine*) con kernel lineal ("SVM-lin"), polinomial ("SVM-poly") y RBF ("SVM-RBF"). Para más detalles, se remite al lector a [20].

La Figura 1 muestra los resultados obtenidos para las diferentes combinaciones de τ_r consideradas, para el escenario (a) de [20]. Por limitaciones de espacio, las gráficas obtenidas para los escenarios (b) a (f) se pueden consultar en el material suplementario[1]. Además, la Tabla I contiene los mejores resultados de la *balanced accuracy* para todos los escenarios (a) a (f) junto a la combinación de τ_r con la que se han obtenido. La última fila contiene los correspondientes valores de [20] para facilitar la comparación. Los valores en negrita indican el mejor valor por cada escenario (por columna).

A partir de los diagramas de L-momentos estándar de todos los escenarios se observa que los clusters obtenidos están, de forma general, igualmente definidos con diferentes combinaciones de τ_r. Como ya comentamos en [20], cada tipo de flujo/ataque requiere de un preprocesado para seleccionar la característica que aporta más información y genera clusters mejor definidos. Ahora, además, es necesario incluir los órdenes r para seleccionar la combinación que mejor se adapte al comportamiento de los flujos.

En cuanto a los resultados de *balanced accuracy*, cabe resaltar que se ha observado mejora con diferentes τ_r en todos los escenarios, o al menos se ha obtenido el mismo valor. Estas mejoras son de hasta un 4,44%, suponiendo una mejora significativa teniendo en cuenta los altos valores de *balanced accuracy* de los que partimos.

IV. CONCLUSIONES

Los sistemas capaces de proporcionar inteligencia a la red para tareas de gestión y seguridad se incluyen entre los pilares de las redes de próxima generación. En este sentido, los L-momentos han demostrado ser una herramienta muy útil para alimentar modelos ML con flujos de red, demostrando además que el uso de L-momentos de orden superior puede suponer mejoras significativas en la calidad de las clasificaciones realizadas. En trabajos posteriores realizaremos un estudio exhaustivo de las implicaciones del comportamiento de los ataques de red en los valores de τ_5, así como profundizar en L-momentos de mayor orden como τ_6.

ACKNOWLEDGEMENTS

Este trabajo ha sido financiado en parte por la Unión Europea NextGenerationEU/PRTR, con el proyecto TED2021-131699B-I00 (AEI/FEDER,UE), por el Ministerio de Ciencia e Innovación, con los proyectos PID2020-112545RB-C54 y PDC2022-133900-I00, y por la Univ. Rey Juan Carlos (Ref. F920 and "AYUDA PUENTE 2022, URJC" Ref. F931). F. Luna también agradece la ayuda de la Univ. de Málaga bajo el II PPIT.

[1] https://doi.org/10.6084/m9.figshare.23309738.v1

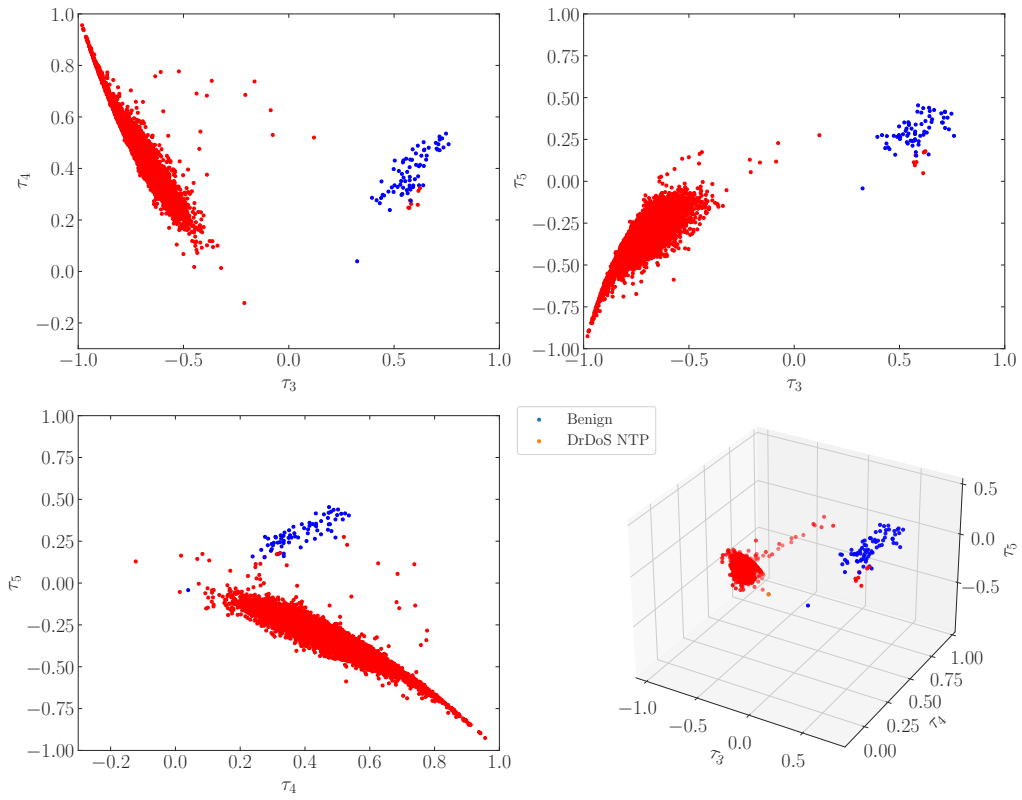


Fig. 1. Diagramas de L-momentos estándar para el escenario (a).

Tabla I

Balanced accuracy PARA TODOS LOS ESCENARIOS. EN NEGRITA SE MUESTRA LA MEJOR PUNTUACIÓN POR ESCENARIO Y EN EL SUBÍNDICE DE CADA VALOR SE MUESTRA LA COMBINACIÓN DE τ_r ASOCIADA. LA ÚLTIMA FILA CONTIENE LOS MEJORES VALORES OBTENIDOS EN [20].

	(a)	(b)	(c)	(d)	(e)	(f)
kNN-unif	$0,9994_{\tau_3-\tau_5}$	$0,5583_{\tau_3-\tau_5}$	$0,9993_{\tau_3-\tau_5}$	$\mathbf{0,9989}_{\tau_3-\tau_4-\tau_5}$	$0,6649_{\tau_3-\tau_4-\tau_5}$	$0,9545_{\tau_3-\tau_5}$
kNN-dist	$0,9995_{\tau_4-\tau_5}$	$0,9291_{\tau_3-\tau_5}$	$0,9985_{\tau_3-\tau_4-\tau_5}$	$0,9800_{\tau_3-\tau_4-\tau_5}$	$0,7548_{\tau_3-\tau_4-\tau_5}$	$0,9670_{\tau_3-\tau_5}$
SVM-lin	$0,9993_{\tau_3-\tau_5}$	$\mathbf{0,9999}_{\tau_3-\tau_4-\tau_5}$	$0,9993_{\tau_3-\tau_5}$	$0,9984_{\tau_3-\tau_4-\tau_5}$	$\mathbf{1,0000}_{\tau_3-\tau_4-\tau_5}$	$0,9946_{\tau_3-\tau_5}$
SVM_RBF	$\mathbf{0,9998}_{\tau_3-\tau_5}$	$0,9875_{\tau_3-\tau_5}$	$0,9993_{\tau_3-\tau_5}$	$0,9795_{\tau_3-\tau_4-\tau_5}$	$0,9999_{\tau_3-\tau_4-\tau_5}$	$\mathbf{0,9995}_{\tau_3-\tau_5}$
SVM-poly	$0,9875_{\tau_3-\tau_5}$	$0,9875_{\tau_3-\tau_4-\tau_5}$	$\mathbf{0,9994}_{\tau_3-\tau_5}$	$0,9800_{\tau_3-\tau_4-\tau_5}$	$0,9973_{\tau_3-\tau_4-\tau_5}$	$0,9921_{\tau_4-\tau_5}$
[20]	$0,9995_{\tau_3-\tau_4}$	$0,9916_{\tau_3-\tau_4}$	$0,9991_{\tau_3-\tau_4}$	$\mathbf{0,9989}_{\tau_3-\tau_4}$	$0,9556\tau_3-\tau_4$	$\mathbf{0,9995}_{\tau_3-\tau_4}$

REFERENCIAS

[1] Z. Zhang *et al.*, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, 2019.

[2] M. Bunyakitanon *et al.*, "End-to-end performance-based autonomous vnf placement with adopted reinforcement learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 2, 2020.

[3] E. GSZSM, "Zero-touch network and Service Management (ZSM); Reference Architecture," Tech. Rep, Tech. Rep., 2019.

[4] C. Benzaid *et al.*, "AI-driven zero touch network and service management in 5G and beyond: Challenges and research directions," *IEEE Netw.*, vol. 34, no. 2, 2020.

[5] M. Bagaa *et al.*, "QoS and Resource-aware Security Orchestration and Life Cycle Management," *IEEE. Trans. Mob. Comput.*, 2020.

[6] H.-K. Lim *et al.*, "Payload-based traffic classification using multi-layer LSTM in Software Defined Networks," *Appl. Sci.-Basel*, vol. 9, no. 12, 2019.

[7] F. Pacheco *et al.*, "Towards the Deployment of Machine Learning Solutions in Network Traffic Classification: A Systematic Survey," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 2, 2018.

[8] S. Rezaei *et al.*, "Deep Learning for Encrypted Traffic Classification: An Overview," *IEEE Commun. Mag.*, vol. 57, no. 5, 2019.

[9] J. R. Hosking, "L-moments: Analysis and estimation of distributions using linear combinations of order statistics," *J. R. Stat. Soc. Ser. B-Stat. Methodol.*, vol. 52, no. 1, 1990.

[10] W. H. Asquith, "Univariate Distributional Analysis with L-moment Statistics using R," Ph.D. dissertation, 2011.

[11] M. Fawad *et al.*, "Multiparameter probability distributions for at-site frequency analysis of annual maximum wind speed with L-moments for parameter estimation," *Energy*, vol. 181, 2019.

[12] R. Ginoulhac *et al.*, "Target Classification Based On Kinematic Data From AIS/ADS-B, Using Statistical Features Extraction and Boosting," in *20th IRS*. IEEE, 2019.

[13] F. Mohd-Zaid *et al.*, "A test on the L-moments of the degree distribution of a Barabási–Albert network for detecting nodal and edge degradation," *J. Complex Netw.*, vol. 6, no. 1, 2018.

[14] J. R. Hosking, *L-Moments and their Applications in the Analysis of Financial Data*. IBM Thomas J. Watson Research Div., 1999.

[15] E. Jurczenko *et al.*, "Efficient frontier for robust higher-order moment portfolio selection," 2008.

[16] N. U. Nair *et al.*, "L-moments of residual life," *J. Stat. Plan. Infer.*, vol. 140, no. 9, 2010.

[17] S. Cao *et al.*, "A novel fourth-order l-moment reliability method for l-correlated variables," *Appl. Math. Model*, vol. 95, 2021.

[18] P. Royston, "Which measures of skewness and kurtosis are best?" *Stat. Med.*, vol. 11, no. 3, 1992.

[19] J. Hosking, "Some theory and practical uses of trimmed L-moments," *J. Stat. Plan. Infer.*, vol. 137, no. 9, 2007.

[20] J. Galeano-Brajones *et al.*, "A novel approach for flow analysis in software-based networks using L-moments theory," *Comput. Commun.*, vol. 201, 2023.

[21] I. Sharafaldin *et al.*, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *IEEE ICCST*, 2019.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Optimización del despliegue de red para la mejora del rendimiento de los protocolos de gestión de la movilidad en redes 6G

Jesús Calle-Cancho[1,2], Jesús Galeano-Brajones [2], Javier Carmona-Murillo[2], David Cortés-Polo[2], Francisco Luna[3]

jesusmanuel.calle@ciemat.es, jgaleanobra@unex.es, jcarmur@unex.es, dcorpol@unex.es, flv@lcc.uma.es

[1]Dpto. Tecnología. CIEMAT. Centro de Tecnologías Avanzadas (CETA-CIEMAT), Trujillo, España.

[2]Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos. Universidad de Extremadura. Cáceres, España.

[3]Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga, E.T.S.I. Informática, Málaga, España. ITIS Software, Universidad de Málaga, Málaga, España.

Con la continua evolución de las tecnologías 5G avanzado y 6G, nuevos servicios y aplicaciones con requisitos estrictos de fiabilidad y latencia han tomado gran importancia. Además, la movilidad del usuario requiere de la participación de protocolos de gestión de movilidad para garantizar la continuidad del servicio. Para responder a estos retos, los operadores necesitan planificar eficientemente el despliegue de red bajo estas condiciones. Este artículo tiene como objetivo optimizar la asignación de estaciones base a los nodos de borde en redes móviles para mejorar el rendimiento del proceso de gestión de la movilidad. Los resultados obtenidos muestran que el mecanismo de optimización propuesto consigue mejoras significativas desde el punto de vista del plano de control y del plano de datos con respecto a otros mecanismos de referencia.

***Palabras Clave*—6G, optimización, gestión de la movilidad, despliegue de red**

I. INTRODUCCIÓN

En los últimos años, se ha producido una evolución considerable de las comunicaciones móviles, debido a la gran proliferación de dispositivos móviles que generan una cantidad de tráfico de datos sin precedentes. La nueva generación de comunicaciones móviles se ha visto obligada a evolucionar para hacer frente a este crecimiento y garantizar servicios y aplicaciones emergentes de acuerdo a demandas específicas de los usuarios, planteando 5G avanzado y 6G como tecnologías que darán respuestas a retos complejos relacionados con despliegues ultradensos con estrictos requisitos de latencia y fiabilidad [1].

Desde el punto de vista de la arquitectura de red, el 3GPP ha realizado importantes esfuerzos para evolucionar el estándar, sentando las bases de lo que se conoce como 5G avanzado (3GPP Rel-18) [2] e introduciendo características para proporcionar mayor flexibilidad y eficiencia a la red y, analizando la evolución de las propias arquitecturas de red para optimimizar sus despliegues. Además, en la 3GPP Rel-18 se definen una serie de áreas de interés como son la eficiencia energética, la cobertura, el soporte a la movilidad y el posicionamiento, entre otros [3]. En los próximos años, se prevé una continua evolución en esta dirección hasta la llegada de redes móviles 6G dando lugar a una red de acceso mejorada con un rendimiento global de red superior [4]. Por tanto, como el soporte a la movilidad es un aspecto clave para las nuevas generaciones de redes móviles, se hace necesario que los operadores de red tengan en cuenta estos aspectos a la hora de planificar sus despliegues. En este sentido, es necesario desarrollar nuevas líneas de investigación que incluyan el uso de técnicas para la asignación de estaciones base a nodos de la red acceso y su optimización para mejorar el rendimiento global de la red desde el punto de vista de la gestión de la movilidad [5]. En este artículo proponemos una optimización basada en un algoritmo genético para resolver este problema de asociación, con el objetivo de mejorar el despliegue de red basado en heurísticos propuestos con anterioridad [6].

El resto del artículo está organizado de la siguiente manera. La sección II presenta el modelado del sistema y la formulación del problema, definiendo el mecanismo de optimización propuesto basado en un algoritmo genético. En la sección III se definen las métricas utilizadas para comprobar el rendimiento del enfoque propuesto. La sección IV muestra los resultados numéricos obtenidos tras la evaluación de rendimiento. Por último, la sección V presenta las conclusiones del trabajo.

II. MODELADO DEL SISTEMA Y FORMULACIÓN DEL PROBLEMA

En esta sección, se presenta el modelado del sistema sobre el que se definirá el problema de optimización para minimizar el impacto de la asignación de estaciones base a los nodos del borde de la red de acceso, sin pérdida de rendimiento.

A. Red de acceso móvil

Se considera una red de acceso representada como un grafo no dirigido $G = (V, E)$, donde V y E denotan los conjuntos de nodos y enlaces (aristas) respectivamente. Sea $K \subseteq V$ el conjunto de los *routers* de acceso que sirven y proporcionan acceso a los nodos móviles a través de un conjunto de estaciones base B, donde cada estación base es denotada por $b_i (1 \leq i \leq |B|)$. Este conjunto B proporciona cobertura total a un área geográfica determinada y cada ubicación viene dada por $\{L_{b_i}\}_{b_i \in B}$, donde $L_{b_i} \in \mathbb{R}^2$ representa las coordenadas donde se ubicarán las estaciones base.

B. Asignación de estaciones base a routers de borde

Cada *router* de acceso $\{k_j\}_{j \in K}$ sirve a un subconjunto de estaciones base $B_k \subseteq B$ dentro de un dominio de red. Los *routers* de acceso (AR) son definidos como los nodos de primer salto que pueden ser considerados como el enlace entre el nivel físico y el nivel de red. Además, N denota el conjunto de nodos móviles (MN) que se mueven por el dominio de red y se conectan a las estaciones base $b_i \in B$; siendo cada MN definido por $N_j (1 \leq j \leq |N|)$.

Adicionalmente, se asume que cada estación base b_i es asociada a un AR del dominio de movilidad, tal y como se muestra en la Figura 1, y cada *router* de acceso k_j gestiona un conjunto determinado de estaciones base.

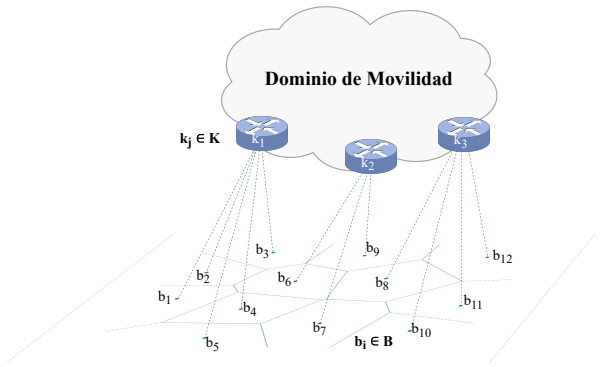


Fig. 1. Asignación de las estaciones base al dominio de movilidad.

Tomando como base la definición previa del modelo de red y su formalización, definimos una variable de decisión x_{mr}^{ps} tal y como se muestra a continuación:

$$x_{mr}^{ps} = \begin{cases} 1 & \text{si la estación base } m \text{ es asignada al } \textit{router} \\ & \text{de acceso } r \text{ y la estación base } p \text{ es asignada} \\ & \text{al router de acceso } s. \\ 0 & \text{en caso contrario.} \end{cases}$$

C. Definición del problema

Teniendo en cuenta esta variable de decisión, la asignación óptima entre la red de acceso y las estaciones base es definida como un problema de optimización [6].

$$\underset{x_{mr}^{ps}}{\text{Min}} \qquad F = \sum_{m \in B} \sum_{r \in K} \sum_{p \in B} \sum_{s \in K} TC x_{mr}^{ps} \tag{1}$$

sujeto a:

$$\sum_{r \in K} \sum_{p \in B} \sum_{s \in K} x_{mr}^{ps} + \sum_{m' \in B} \sum_{r' \in K} \sum_{s' \in K} x_{m'r'}^{p's'} = 1, \quad \forall m = p' \in B, m = p' = 1, \ldots, B \tag{2}$$

$$\sum_{m \in B} \sum_{p \in B} \sum_{s \in K} x_{mr}^{ps} + \sum_{m' \in B} \sum_{r' \in K} \sum_{p' \in B} x_{m'r'}^{p's'} \leq th_j, \quad \forall r = s' = j \in K, r = s' = j = 1, \ldots, K \tag{3}$$

$$x_{mr}^{ps} \in \{0, 1\}, \; r, s \in K \; \& \; m, p \in B \tag{4}$$

Para cada asignación, se define un coste total TC como la suma de dos parámetros críticos relacionados con los protocolos de gestión de la movilidad: coste de señalización (C_S) y coste de entrega de paquetes (C_{PD}).

$$TC = C_S + C_{PD} \tag{5}$$

La Restricción 2 indica que una estación base ($m = p' \in B$) es asignada a un único AR y la Restricción 3 está relacionada con el balanceo de estaciones base entre los diferentes AR. Se asume que un determinado AR ($r = s' \in K$) no puede servir a más de un número específico de estaciones base, que viene determinado por un *threshold* (th_j). En el trabajo previo llevado a cabo por los autores [6], se propuso una nueva estrategia para resolver el problema en tiempo polinomial, definiéndose el algoritmo LNA (*Link-Network Assignment*), basado en la recopilación de información de la red de acceso y en el análisis de distribución de estaciones base con el objetivo de realizar la asignación adecuada, llevando a cabo una evaluación de rendimiento comparativa del mecanismo propuesto con respecto a otros algoritmos de asignación, en términos de costes de señalización y costes de entrega de paquetes. Estos algoritmos serán utilizados como base para la comparativa realizada en el presente artículo.

D. Mecanismo de optimización basado en un algoritmo genético

Con el objetivo de resolver el problema de optimización, en este artículo se propone un mecanismo basado en un algoritmo genético con dos objetivos ponderados, tal y como se muestra en la Ec. 6.

$$\text{Min} \quad F = \alpha C_S + \beta C_{PD} \tag{6}$$

α y β son factores de ponderación que utilizará nuestro algoritmo genético para ajustar la importancia relativa de los costes de movilidad: señalización (C_S) y datos (C_{PD}); debido a la diferencia de magnitudes que existen entre

ambos costes. Además, el algoritmo tiene en cuenta las restricciones definidas anteriormente en la formulación del problema.

A continuación, se realiza una breve descripción de dicho algoritmo genético, así como los operadores utilizados y la metodología seguida.

Los algoritmos genéticos son una técnica de optimización y búsqueda inspirada en la evolución biológica [7] que se basan en los principio de selección natural y la supervivencia del individuo más apto para resolver problemas complejos y encontrar aproximaciones a las soluciones óptimas en amplios espacios de búsqueda. En este trabajo, los operadores genéticos utilizados han sido: selección por torneo binario; cruce en dos puntos con una probabilidad de 0.9; y mutación uniforme con probabilidad $1/B$. Además, la población tiene un tamaño de 100 individuos y la condición de parada se ajusta a la evaluación de 25000 soluciones candidatas por cada ejecución lanzada.

En cuanto a las restricciones del problema, la Restricción 2 queda satisfecha por la propia codificación entera de las soluciones candidatas. La Restricción 3 se implementa directamente en el algoritmo de forma que las soluciones factibles aseguren el balanceo de carga de estaciones base en los *routers*. Por último, dada la naturaleza estocástica de los algoritmos genéticos, se han realizado 15 ejecuciones de 25000 evaluaciones cada una, para la misma instancia del problema, con el objetivo de asegurar la confianza estadística.

III. MÉTRICAS DE EVALUACIÓN

A continuación se definen las métricas de rendimiento utilizadas para evaluar la eficiencia del mecanismo de asignación propuesto analizando el impacto sobre el protocolo de gestión de la movilidad utilizado desde dos puntos de vista: plano de control y plano de datos. Concretamente, en la evaluación llevada a cabo, se ha utilizado como base un protocolo de gestión de la movilidad distribuido basado en la red (*Network-Based DMM*, NB-DMM) [8]. Estas métricas son ampliamente utilizadas con anterioridad en los análisis y evaluaciones de soluciones de movilidad [9], permitiendo determinar el rendimiento global de la red móvil.

A. Plano de control

Con el objetivo de evaluar el plano de control, una métrica relevante es el coste total de señalización relativo a la actualización de vínculos de movilidad durante una sesión (C_S), el cual representa la carga de tráfico acumulada en el intercambio de mensajes de señalización. Este coste depende del tamaño de los propios mensajes de señalización y del número de traspasos de nivel 3 realizados durante el intervalo de tiempo en el que la comunicación en el MN se mantiene activa. En NB-DMM, el coste de señalización durante el movimiento puede ser expresada tal y como se muestra en Ec. 7.

$$C_S = 2s_u + 2s_u \sum_{i=1}^{n-1} h_{GW_i - S_{GW}} \tag{7}$$

donde n define el número de *routers* previos que establecen un túnel con el *router* actual. Además, h_{x-y} representa la distancia en número de saltos desde el nodo x al nodo y en la red y s_u indica el tamaño medio de los mensajes de señalización.

B. Plano de datos

Con respecto al plano de datos, una de las métricas que tienen un mayor impacto sobre el rendimiento global de la red es el coste de entrega de paquetes (C_{PD}). Aparte de la señalización relacionada con el proceso de movilidad, los paquetes de datos tienen que ser enviados desde el CN (*Correspondent Node*) al MN, y viceversa. Este valor está influenciado por el tamaño medio de los mensajes de datos multiplicado por el número de saltos necesarios para reenviar los paquetes desde el CN hasta el MN. En NB-DMM, el coste de entrega de paquetes es definifido tal y como se muestra en la Ec. 8.

$$C_{PD} = (P_n C^d_{PD} + P_h C^i_{PD}) N_{p/s} \tag{8}$$

donde $N_{p/s}$ representa la tasa de transmisión de paquetes por flujo activo. Además, P_n y P_h son, respectivamente, las probabilidades de que un flujo de tráfico sea nuevo o de que un flujo siga abierto después de realizar un *handover*. Por lo tanto, C^d_{PD} y C^i_{PD} representan los costes de entrega de un paquete para los modos de funcionamiento directo e indirecto de NB-DMM, respectivamente.

$$C^d_{PD} = s_d h_{CN-S_{GW}} + s_d h_{S_{GW}-MN} \tag{9}$$

$$C^i_{PD} = s_d h_{CN-GW} + (s_t + s_d) h_{GW-S_{GW}} + \\ + s_d h_{S_{GW}-MN} \tag{10}$$

s_d es el tamaño medio de los paquetes de datos y s_t es el tamaño medio de la cabecera de tunelización IPv6.

IV. RESULTADOS NUMÉRICOS

Esta sección tiene como objetivo llevar a cabo un análisis del impacto de los costes de movilidad sobre el rendimiento general de la red, así como la evaluación del mecanismo propuesto de asignación óptima basada en un algoritmo genético. La Fig. 2 muestra la topología de red utilizada para llevar a cabo la evaluación [10].

El escenario de simulación es una región cuadrada de 10×10 km^2, donde las estaciones base están distribuidas siguiendo un proceso llamado *Poisson Point Process* (PPP), cuya intensidad (λ_{BS}) coincide con el número medio de estaciones base (N_{BS}) por unidad de área (A) y es obtenido como $\lambda_{BS} = N_{BS}/A$. Además, el área de cobertura de las estaciones base es modelada como una teselación de *Poisson-Voronoi*, donde cada usuario

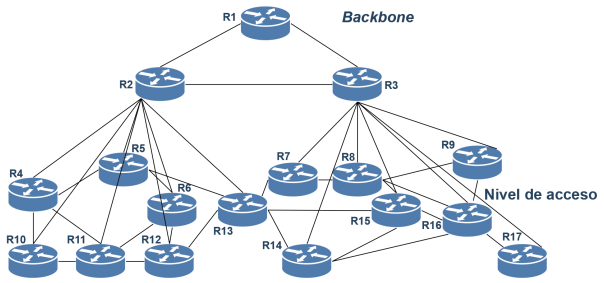


Fig. 2. Topología de red utilizada en las simulaciones.

móvil se conecta a la estación base más cercana. La movilidad de usuario es definida a través del modelo de movilidad *Random Waypoint* con una velocidad uniformemente distribuida entre 1 y 20 m/s. Los usuarios móviles gestionan un conjunto de sesiones durante el tiempo de simulación, asumiéndose que el número de sesiones entrantes por usuario móvil sigue un proceso de Poisson con una tasa media $\lambda = 0.01$, y la duración de una sesión está exponencialmente distribuida con parámetro $\mu = 10$.

El rendimiento de la propuesta es evaluado sobre este escenario de red comparándose con las soluciones LNA y K-Means presentadas en [6] y, calculando los costes de movilidad: coste de señalización y coste de entrega de paquetes. La Fig. 3 muestra la comparativa en promedio entre K-Means (verde), LNA (naranja) y las soluciones obtenidas por el mecanismo propuesto en este trabajo (azul). Por un lado, el mecanismo basado en un algoritmo genético mejora un 10.35% el C_{PD} y un 22.62% el C_S, con respecto al mecanismo basado en K-Means. Por otro lado, con respecto al mecanismo LNA, el mecanismo propuesto en este artículo mejora un 9.71% el C_{PD} y un 16.98% el C_S.

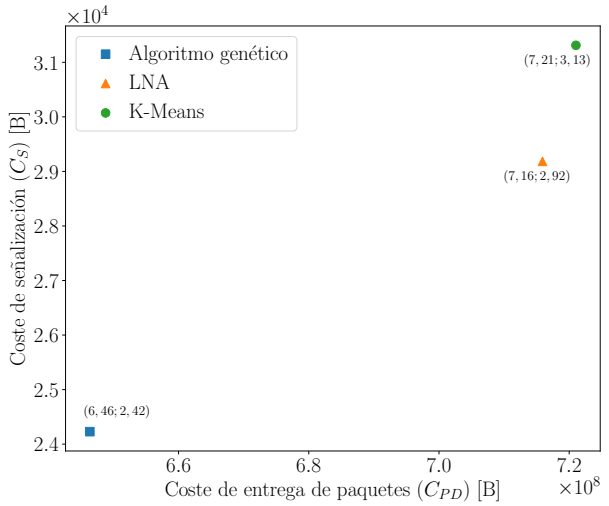


Fig. 3. Comparativa del desempeño promedio (C_S;C_{PD}) del algoritmo genético (azul), la solución de LNA (naranja) y K-Means (verde).

V. CONCLUSIONES

En este artículo se propone un mecanismo de optimización basado en un algoritmo genético que realiza la asignación entre las estaciones base y la red de acceso de manera eficiente, llevando a cabo una evaluación comparativa de rendimiento del mecanismo propuesto con respecto a otros algoritmos de asignación de referencia, en términos de costes de señalización y costes de entrega de paquetes, permitiendo evaluar el rendimiento global de la red móvil.

Los resultados obtenidos demuestran que la propuesta permite reducir con éxito los costes asociados al plano de control hasta un 22.62% y los costes asociados al plano de datos hasta un 10.35%, en comparación con los algoritmos de referencia.

En definitiva, el mecanismo propuesto se trata de una primera aproximación de un trabajo que se encuentra en proceso actualmente y, en el que seguiremos trabajando en dos líneas principalmente: optimización multiobjetivo basada en Pareto y ampliación de la formulación del problema incluyendo otros aspectos como la sobrecarga de tráfico de las estaciones base o la eficiencia energética, entre otros.

AGRADECIMIENTOS

Este trabajo se ha llevado a cabo haciendo uso de la infraestructura de computación facilitada por el CETA-CIEMAT, financiado por el Fondo Europeo de Desarrollo Regional. El CETA-CIEMAT pertenece al Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas y al Ministerio de Ciencia e Innovación. La publicación es parte del proyecto TED2021-131699B-I00, financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea "NextGenerationEU"/PRTR.

REFERENCIAS

[1] X. Xu, X. Tang, Z. Sun, X. Tao, and P. Zhang, "Delay-oriented cross-tier handover optimization in ultra-dense heterogeneous networks," *IEEE Access*, vol. 7, pp. 21 769–21 776, 2019.

[2] 3GPP-RP-213468, "Summary for RAN Rel-18 package," 2021, Último acceso: 14 junio 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/

[3] X. Lin, "An Overview of 5G Advanced Evolution in 3GPP Release 18," *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 77–83, 2022.

[4] Ömer Bulakçı, X. Li, M. Gramaglia, A. Gavras, M. Uusitalo, P. Rugeland, and M. Boldi, *Towards Sustainable and Trustworthy 6G: Challenges, Enablers, and Architectural Design*. Boston-Delft, 2023.

[5] J. Martín-Pérez, L. Cominardi, C. J. Bernardos, A. de la Oliva, and A. Azcorra, "Modeling mobile edge computing deployments for low latency multimedia services," *IEEE Transactions on Broadcasting*, vol. 65, no. 2, pp. 464–474, 2019.

[6] J. Calle-Cancho, J. Carmona-Murillo, J.-L. González-Sánchez, and D. Cortés-Polo, "A Novel Link-Network Assignment to Improve the Performance of Mobility Management Protocols in Future Mobile Networks," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.

[7] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, pp. 8091–8126, 2021.

[8] H. Ali-Ahmad, M. Ouzzif, P. Bertin, and X. Lagrange, "Performance Analysis on Network-Based Distributed Mobility Management," *Wireless Personal Communications*, vol. 74, no. 4, p. 1245–1263, 2014.

[9] E. M. O. Fafolahan and S. Pierre, "A Seamless Mobility Management Protocol in 5G Locator Identificator Split Dense Small Cells," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2019.

[10] G. Zheng, A. Tsiopoulos, and V. Friderikos, "Optimal VNF Chains Management for Proactive Caching," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6735–6748, 2018.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

SareQuant: Towards a quantum-based communication network

Ane Sanz[1,2], David Franco[1], Asier Atutxa[1], Jasone Astorga[1,2], Eduardo Jacob[1,2]
[1]Department of Communications Engineering, University of the Basque Country (UPV/EHU). 48013 Bilbao, Spain.
[2]EHU Quantum Center, University of the Basque Country (UPV/EHU). 48940 Leioa, Spain.
{ane.sanz, david.franco, asier.atutxa, jasone.astorga, eduardo.jacob}@ehu.eus

***Abstract*—This paper presents the SareQuant project, which aims to evolve the Basque NREN (National Research and Education Networks) into a quantum-based communication infrastructure. SareQuant focuses on the network design and on the integration of quantum technologies into real-world scenarios and applications. Therefore, this paper provides insights into the opportunities and challenges regarding the integration of quantum technologies, thus paving the way for a secure and advanced Quantum Internet.**

***Keywords*—Quantum Internet, Quantum Key Distribution, National Research and Education Networks**

I. INTRODUCTION

In today's rapidly evolving digital era, National Research and Education Networks (NRENs) play a crucial role in supporting different research and education activities. NRENs encompass high-speed networks that interconnect multiple research and education institutions, such as universities and research centers, to provide high-speed connectivity and access to advanced resources. NRENs also act as the fundamental framework for knowledge sharing, enabling seamless collaboration, resource sharing, and use of advanced technologies among researchers and students. As an example, the Global P4 Lab (GP4L) is a high-performance communication network for research and education that leverages resources of the European NRENs' for their interconnection. Considering the continuous and rapid evolution of technology, there is a strong need for NRENs to enhance their capabilities by adopting and integrating emerging technologies.

Among the emerging technologies, Quantum Technologies stand out as one of the most promising for next-generation services and applications, including Quantum Computing and Quantum Communications. Such technologies, although still in early stages of development and lacking full maturity, are expected to have a major impact and challenge conventional systems. Quantum Computing, for instance, promises to solve highly complex operations that are beyond the capabilities of classical computing. Quantum communication, on the other hand, offers information-theoretic security by leveraging the principles of quantum mechanics. In addition, the integration of such technologies, as well as the interconnection of quantum devices, implies the deployment of quantum networks that must coexist with classical ones, forming hybrid classical-quantum networks that support new capabilities [1]. It is envisaged that these networks, which will be deployed gradually based on the availability and maturity of the technology, will culminate in a network commonly referred to as the Quantum Internet (QInternet).

In this context, it is important to note that, due to the intrinsic characteristics of quantum technologies, the architecture and operation of the QInternet vastly differs from the classical one, which means that the transition to a quantum-enabled network is not immediate and needs to be studied in more detail.

Therefore, this paper introduces the current status of the SareQuant project, which aims at analysing the requirements and implications for the evolution of I2Basque, the NREN of the Basque Country, towards a quantum-based network architecture. On the other hand, SareQuant also seeks the integration of quantum technologies into real-world scenarios, employing key technologies identified for the initial stages of the Quantum Internet.

II. QUANTUM INTERNET

This section describes the concept of QInternet, exploring the identified opportunities and challenges. It also provides a description of the Quantum Key Distribution (QKD) technology, identified as a key enabler for the early stages of development of the QInternet.

A. Concept, opportunities and challenges

The QInternet is expected to enhance classical Internet by enabling quantum communications between any two points in the world, which will in turn enable new services and applications beyond the scope of classical networks, such as the realisation of complex optimisation problems. According to the Internet Research Task Force (IRTF), which has published several RFC and drafts regarding the QInternet [2], [3], the deployment of the QInternet involves the definition of a new quantum network stack that accounts for fundamental principles of quantum mechanics such as superposition, entanglement or measurement.

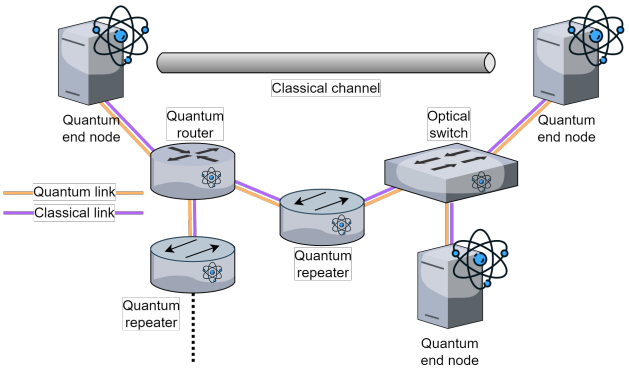


Fig. 1. Quantum Internet generic architecture.

Additionally, the RFC 9340 defines the architectural principles of the QInternet, identifying, among others, the elements that comprise the network. According to this RFC, the main elements that should make up the network are quantum routers, quantum repeaters, quantum end-nodes, and passive elements such as optical switches. In addition, most quantum devices require of both quantum and classical links to properly perform their processes and tasks. Fig. 1 shows a generic quantum network architecture where all the above elements are represented. This scheme represents a scenario where two applications running on two end nodes with quantum capabilities need to communicate with each other, which requires the establishment of an end-to-end connection through different quantum devices.

However, it is important to outline that the deployment of quantum networks is conditioned by the availability of the involved technologies and devices. In fact, quantum repeaters, which are considered essential for establishing long-range links, remain unavailable at the moment. This means that it is not possible to completely stick to the presented generic architecture, and network configurations must be adapted to the maturity and availability of the technology at any given time. For instance, the interconnection of different quantum devices presently relies on a combination of both optical switches and trusted relays that enable connections over medium-large distances. The eventual availability of quantum repeaters will presumably increase this capability, further enhancing the interconnection of quantum devices.

Therefore, there exists a need for a progressive evolution of conventional networks towards a quantum-based infrastructure, in line with technological advances. In this context, studies such as [4] concur with the notion of developing the QInternet in different stages, with the first stages focused on the implementation of QKD systems.

B. Quantum Key Distribution as an early stage

QKD technology enables the generation of a symmetric key between two endpoints in a information-theoretic secure manner, ensuring the utmost privacy and confidentiality of such key. The underlying principle of QKD, which combines the use of both quantum and classical channels, is that any attempt to observe the transmitted photons in the quantum channel disturbs the transmission in a way that induces detectable errors at both communication ends. Thus, considering the ability of establishing and distributing secure keys, QKD emerges as one of the most promising quantum communication technologies for achieving ultra-secure communications [5].

There are two main implementation options available for QKD systems, based on the information encoding method: Discrete-Variable (DV-QKD) and Continuous Variable (CV-QKD). In DV-QKD systems, discrete quantum states such as polarization or phase are used for encoding information. Conversely, CV-QKD systems employ continuous variables of quantum states, such as the quadrature components of the electromagnetic field for information encoding. Additionally, QKD systems can employ two main approaches for the transmission and measurement of quantum signals, namely prepare-and-measure and entanglement-based. In the prepare-and-measure approach, the transmitter is responsible for preparing and transmitting the quantum states, while the receiver undertakes the task for measuring them. On the other hand, entanglement-based approaches entail the use of an external source that generates and distributes entangled photons between both communication ends for subsequent measurement.

Therefore, QKD is currently one of the quantum technologies with the highest level of maturity, supported by the availability of commercial equipment from multiple international manufacturers. This accessibility of QKD devices greatly facilitates research in the area and the implementation of the technology. Consequently, the advanced stage of development and availability of QKD, as well as its inherent advantages in terms of unconditional security, position it as the most suitable technology for early-stage deployment in the QInternet.

III. QUANTUM INITIATIVES IN SPAIN

The quantum technologies sector has gained significant interest in the last years from both companies and public institutions within the international community. This growing attention can be reflected in the multitude of initiatives that have been launched with the aim of promoting the advancement of such technologies.

In the Spanish context, the *Plan Complementario de Comunicación Cuántica* launched by the Ministry of Science and Innovation, has earmarked 54 million EUR to fund research projects that foster the development of quantum digital technologies to enforce the cybersecurity in Spain. This initiative aligns with other European initiatives with similar objectives, such as the EuroQCI or the OpenQKD. Specifically, within the scope of these projects, some work has already been done over local NRENs, as the MadQCI infrastructure [6], a quantum-based metropolitan network, which is undergoing its development.

At the regional level, the Basque Country has also launched the IKUR strategy that prioritises four specific fields, including the Quantum Technologies. This strategy aims at attracting and generating quantum-related talent, fostering the development of novel infrastructures and positioning the Basque Country as an international leader.

Similarly, the University of the Basque Country (UPV/EHU) has recently created the EHU Quantum Center. The objective of this center, which is considered itself an actor of the IKUR strategy, is to coordinate members, groups, and activities, to contribute to scientific excellence, and to participate in public and private quantum initiatives.

Therefore, all these initiatives demonstrate the growing interest of local, national, and international institutions in fostering the development of quantum technologies. In addition, the SareQuant project presented in this work, is realised in the context of the aforementioned initiatives, specifically as a part of the IKUR initiative, contributing to the collective efforts aimed at advancing quantum technologies and their applications.

IV. SAREQUANT: TOWARDS A QUANTUM-BASED NETWORK

SareQuant is a project aimed at proposing and designing an infrastructure compatible with the existing I2Basque, the basque NREN, to allow experimentation with quantum technologies, taking a first step in the evolution towards a QInternet. This infrastructure would extend the network's current functionalities, transforming it into a practical testbed for using and evaluating quantum technologies and supporting experimentation with already accessible quantum technologies in the initial phases. Therefore, SareQuant encompasses two main lines of work. The first one is focused on the progressive evolution of the actual infrastructure design towards a QInternet, while the second concentrates on the integration of quantum technologies, specifically QKD, into real-world scenarios.

A. Current state of I2Basque

I2Basque is responsible for providing advanced network infrastructure and services to the research and education communities of the region. This network, as depicted in Fig. 2, consists of a central ring with four main Points of Presence (PoPs) in Donostia-San Sebastián, Arrasate, Leioa and Vitoria-Gasteiz, connected through 10 Gbps lambda links. This interconnection leverages Dense Wavelength Division Multiplexing (DWDM) technology to allow simultaneous transmission of multiple wavelengths over a single optical fiber. It also includes two metropolitan rings with dark fiber in Bizkaia between Bilbao-Leioa-Zamudio, and in Donostia between Ibaeta-Miramon-H.Donostia, as well as different dark fiber links in Gasteiz and Arrasate. The I2Basque network topology comprises several nodes connected by links of different distances, all of them within an acceptable range in terms of feasibility with current quantum technology, as shown in Fig. 2. This feature, which holds significant importance and often becomes critical in the context of quantum communications, positions this network as an ideal platform for initiating the transition towards a quantum-based network.

Furthermore, this infrastructure also has PoPs to the networks of RedIris, the Spanish NREN, and GÉANT, the collaboration of European NRENs. This PoPs offer several benefits to the I2Basque network, including improved and direct collaboration opportunities with other international research and education institutions present in GÉANT, thereby expanding the potential for experimentation.

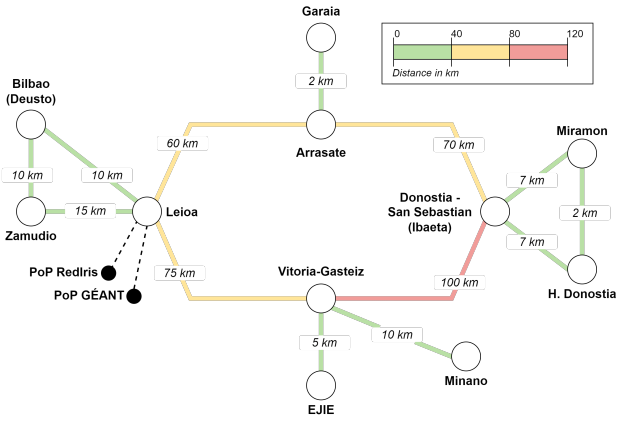


Fig. 2. I2Basque network nodes and links

B. Evolution of I2Basque towards a Quantum Internet

The transition from classical networks to quantum-based communication networks is a challenging task that requires careful analysis of the existing infrastructure. In this case, the current I2Basque network must be studied in detail, with special emphasis on those aspects that may condition the proper integration of quantum technologies. Important aspects requiring careful consideration encompass the number of nodes, link distances, availability of a dark fiber infrastructure, and the multiplexing technology employed to enable the coexistence between classical and quantum signals.

The conclusions of our initial studies reveal that the topology and technologies employed in the I2Basque network present certain attributes that make it an optimal platform for the transition to a quantum-based infrastructure, as described as follows. Considering the current unavailability of more advanced quantum devices that enable long distance connections, most of the **link distances** within I2Basque allow the deployment of a quantum network with end-to-end connections. Consequently, the need for intermediate trusted relays, for example, can be avoided. Regarding the transmission of quantum signals, it is worth mentioning the **absence of a dark fiber infrastructure** in the main ring of I2Basque, so multiplexing techniques must be employed. These techniques enable the transmission of quantum signals alongside the classical ones, without compromising the performance of either communication. In this particular case, as **DWDM** is the technology employed in the I2Basque network, the feasibility and implications of transmitting quantum signals using this technique must be studied in detail, in order to assess its compatibility, limitations and requirements.

In addition, there are some other attributes that also require further consideration in the design of the new quantum-enabled infrastructure. This includes exploring the need for developing **new protocols or interfaces** specifically tailored to quantum networks that address their

Table I
KEY FACTORS FOR THE INTEGRATION OF QKD INTO REAL-WORLD APPLICATIONS

Factor	Description	Example
Definition of use cases.	Identify different use cases and applications suitable with QKD which may benefit from its use.	Securisation of data centers supporting virtualised environments, securisation of 5G/6G infrastructures, etc.
Use of standardised interfaces.	Implement standardised APIs for key exchange, control and management to ensure security, efficiency and compatibility.	Implementation of ETSI 004, ETSI 014 or ETSI 015 APIs.
Adaptation of standard protocols.	Adapt the protocols used in the selected application, if needed, in order make use of QKD keys.	Make changes in protocols such as TLS, IPsec, SSH, etc. to adapt them to the specific features of the QKD keys.
Secure key storage and management.	Implement proper methods to ensure secure storage and management of QKD keys.	Use of secure enclaves, strong authentication methods in interfaces, etc.
Secure and efficient use of keys.	Implement proper methods to ensure that applications make use of keys in a secure and efficient manner.	Implementation of key-synchronisation methods, definition of re-keying processes, definition of key derivation or re-utilisation politics, etc.

unique requirements. Additionally, **scalability** is another critical attribute, as it is essential to establish strategies that support possible future growths and evolutions of the network, considering that more advanced quantum devices and capabilities are expected to be integrated, and that the network itself may also be expanded.

Therefore, the assessment of the main requirements for quantum network implementations according to the State of the Art, show that the I2Basque network has favorable characteristics to support the establishment of a quantum network in its current state and in coexistence with the existing infrastructure. However, more research needs to be done to evaluate all requirements and to design an infrastructure that ensures successful integration of quantum technologies in the I2Basque infrastructure.

C. Deployment of QKD-based secure services

Considering that QKD has been identified as the key enabling technology for the initial phases of QInternet deployments, the integration of QKD into real-world scenarios and applications stands as an important subject within the SareQuant project. Enabling QKD-based services in current applications as an initial step, in addition to bringing numerous security-related advantages, allows the acquisition of essential knowledge and the laying of the foundations for future quantum services. This strategic approach allows leveraging currently available quantum capabilities, built upon existing systems, while simultaneously paving the way for the adoption of more advanced quantum technologies.

Consequently, several factors have been identified as crucial to be considered and analysed in any integration of QKD into real applications. Table I presents an overview of these factors, complemented with a short description and some examples for enhanced comprehension. The assessment of all these factors enables a successful integration of QKD into practical applications and the development of robust services.

V. CONCLUSIONS AND FUTURE WORK

Quantum Technologies are expected to revolutionise current networks and services, with more advanced and promising capabilities. In this context, being NRENs crucial infrastructures for the research and education community, it is essential for them to embrace these technologies in order to enhance their capabilities and contribute to the development of the QInternet.

This paper presents the SareQuant project, focusing on the design requirements for the evolution of the current I2Basque network into a quantum-based one. The conclusions of the performed analysis up to the date show that the features of I2Basque pose it as an optimal infrastructure to be integrated with quantum technologies. In addition, considering QKD the technology for initial stages of the QInternet, this paper also presents an analysis of requirements and key factors to be considered when integrating QKD into real-world scenarios. This contributes to complete successful and robust integration of QKD with real applications, thus setting the ball rolling towards the adoption of more advanced future quantum technologies.

Building upon these first steps, there is still more work to be done. Further research is required to explore full potential of quantum networks and its applications, as well as to define all requirements in order to achieve full integration of classical and quantum networks.

ACKNOWLEDGEMENTS

This work was supported in part by Basque Government through the SareQuant project from IKUR (IKUR-DIPC-PRTR-23/02) and the QFirst project (KK-2022/00062), and in part by the European Commission through the GN5-1 HORIZON-INFRA-2022-NET-01-SGA project.

REFERENCES

[1] R. Van Meter, *Quantum networking*. John Wiley & Sons, 2014.
[2] W. Kozlowski, S. Wehner, R. V. Meter, B. Rijsman, A. S. Cacciapuoti, M. Caleffi, and S. Nagayama, "Architectural Principles for a Quantum Internet," RFC 9340, Mar. 2023.
[3] C. Wang, A. Rahman, R. Li, M. Aelmans, and K. Chakraborty, "Application Scenarios for the Quantum Internet," Internet Engineering Task Force, Internet-Draft draft-irtf-qirg-quantum-internet-use-cases-16, May 2023, work in Progress.
[4] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, no. 6412, p. eaam9288, 2018.
[5] Y. Cao, Y. Zhao, Q. Wang, J. Zhang, S. X. Ng, and L. Hanzo, "The evolution of quantum key distribution networks: On the road to the qinternet," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 839–894, 2022.
[6] D. Lopez, J. P. Brito, A. Pastor, V. Martín, C. Sánchez, D. Rincon, and V. Lopez, "Madrid quantum communication infrastructure: a testbed for assessing qkd technologies into real production networks," in *Optical Fiber Communication Conference*. Optica Publishing Group, 2021, pp. Th2A–4.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Mejorando la escalabilidad de la replicación de datos en sistemas distribuidos shared-nothing mediante la Internet Cuántica

Agustín Zaballos, Joan Navarro
Departamento de Ingeniería,
La Salle Campus Barcelona - Universitat Ramon Llull
Quatre Camins, 30. 08022, Barcelona.
agustin.zaballos@salle.url.edu, jnavarro@salleurl.edu

La replicación de datos en sistemas distribuidos *shared-nothing* es uno de los mecanismos fundamentales para mejorar la disponibilidad y tolerancia a fallos en estos entornos. Sin embargo, la implementación de los algoritmos tradicionales de replicación de datos presenta limitaciones en cuanto a escalabilidad debido a la congestión—tanto a nivel de red de comunicaciones como a nivel de nodos de computación y almacenamiento—asociada a la sincronización distribuida de datos. En este trabajo se propone estudiar, desde un punto de vista analítico, si la Internet Cuántica podría contribuir a mitigar estas limitaciones y mejorar la escalabilidad de la replicación de datos en sistemas distribuidos. Concretamente, se explorará cómo las propiedades de las redes cuánticas (superposición y entrelazamiento), pueden contribuir a reducir la congestión y el procesamiento eficiente de los mensajes de replicación de datos. Este estudio busca abrir nuevas perspectivas en la replicación de datos en entornos cuánticos.

Palabras Clave*—Internet Cuántica, replicación de datos, escalabilidad, sistemas distribuidos, *shared-nothing

I. INTRODUCCIÓN

En un sistema distribuido *shared-nothing* cada nodo tiene su propio conjunto de recursos y no comparte memoria ni almacenamiento con otros nodos, por lo que los distintos servidores se deben comunicar entre ellos a través de una red de comunicaciones. En este tipo de entornos, la replicación de datos—la cual consiste en mantener múltiples copias de los datos en diferentes ubicaciones—es uno de los mecanismos clásicos para mejorar algunas de las prestaciones del sistema en cuanto a tolerancia a fallos y disponibilidad [1] permitiendo una mayor redundancia y acceso local rápido a los datos. Así, si, por ejemplo, un nodo puede servir k peticiones por segundo, cuando los datos de este nodo están replicados en otros $j-1$ nodos idénticos, se podrían llegar a servir, idealmente, hasta $j*k$ peticiones por segundo [2].

Sin embargo, es ampliamente conocido que los algoritmos de replicación tradicionales presentan limitaciones importantes en cuanto a su escalabilidad [3]. El (alto) número de mensajes que estos algoritmos necesitan [2] para sincronizar los datos datos entre los distintos nodos y lleguen a una situación de consenso, hace que tanto la red como los propios nodos se congestionen rápidamente cuando el número de réplicas crece [4]. Hasta la fecha, estos problemas de escalabilidad se están intentando soslayar mediante modelos relajados de consistencia [5], reducir o modificar dinámicamente el número de réplicas [6], [7] que participan en el proceso de replicación, o delegando parte del rol del protocolo de replicación a los dispositivos de red [8] entre otras. Estas aproximaciones comprometen otros parámetros del sistema como por ejemplo un incremento en la complejidad lógica de las aplicaciones que se alimentan de los datos replicados o incluso en la tolerancia a fallos (o *k-safety*) del sistema [9].

El propósito de este trabajo en curso es abordar el desafío de la replicación escalable de datos en sistemas distribuidos *shared-nothing* explorando el potencial de la Internet Cuántica. Concretamente, defendemos que el consenso cuántico puede mejorar la eficiencia en el proceso de sincronización en la replicación de datos gracias a la coordinación instantánea que nos ofrece la Internet cuántica. Una de las ventajas clave del consenso cuántico es la capacidad de lograr una coordinación instantánea entre los nodos. En el caso específico de la replicación de datos, esto significa que los nodos replicados pueden alcanzar un acuerdo casi al instante sobre el estado cuántico que se debe replicar. Esto elimina la necesidad de esperar por largos periodos de tiempo para lograr la sincronización entre los nodos. De esta forma los algoritmos de consenso cuántico podrían reducir la complejidad de comunicación

necesaria para lograr la sincronización en la replicación de datos y, mejorar así, su factor de escalabilidad.

El resto del trabajo se estructura de la siguiente manera: la Sección 2 hace una breve revisión de los fundamentos teóricos sobre los que se apoya esta investigación. A continuación, la Sección 3 presenta la adaptación de un modelo analítico que permite estimar el factor de escalabilidad (*ScaleOut*) de la replicación de datos en entornos distribuidos *shared-nothing*. Finalmente, la Sección 4 discute los efectos que tendría la Internet Cuántica sobre dicho modelo y apunta posibles líneas de investigación futuras.

II. FIJANDO LA TRABAZÓN ENTRE LAS DOS INTERNETS

En el contexto de la replicación de datos, un algoritmo de consenso cuántico debe lograr que un grupo de participantes en una computación distribuida llegue a un acuerdo sobre un valor o estado común, a pesar de la presencia de nodos defectuosos o comportamientos maliciosos [10], [11]. Estos algoritmos utilizan principios y propiedades de la mecánica cuántica, como el entrelazamiento y la superposición, para lograr un consenso más eficiente y seguro en comparación con los algoritmos clásicos [12], [13], [14]. De hecho, el uso de un algoritmo de consenso cuántico en el contexto de la Internet cuántica permite aprovechar las características únicas de la mecánica cuántica, como la privacidad inherente y la coordinación instantánea, para lograr acuerdos confiables y reducir la complejidad de comunicación en comparación con los enfoques clásicos. Sin embargo, hay que tener en cuenta que los algoritmos de consenso cuántico aún se encuentran en desarrollo y son objeto de investigación activa [15], [16].

En [17] experimentan con un algoritmo de consenso cuántico que permite a los autores alcanzar un acuerdo general con una complejidad de comunicación esperada de O(1). Este algoritmo de consenso cuántico permite tolerar hasta un número óptimo de jugadores defectuosos. Dicho algoritmo se presenta en este trabajo como un ejemplo de un algoritmo de consenso cuántico que logra un acuerdo general eficiente con una complejidad de comunicación reducida. El consenso cuántico puede desempeñar un papel importante en la replicación distribuida al garantizar la consistencia y la integridad de los datos replicados en un entorno distribuido. De hecho, la utilización de la superposición y el paralelismo cuántico podría permitir a los nodos replicados realizar múltiples cálculos en paralelo y explorar diferentes estados cuánticos simultáneamente.

En este trabajo en curso nos focalizamos en aprovechar la mejora del rendimiento potencial que la Internet cuántica nos ofrece en este caso de uso. Aunque los algoritmos de consenso pueden requerir una comunicación significativa entre los nodos replicados, la utilización de la Internet cuántica puede reducir la complejidad de las comunicaciones en comparación con la Internet clásica. Esto puede conducir a una mejora en el rendimiento de la replicación distribuida al reducir la sobrecarga de comunicación y minimizar la congestión del tráfico. Otras ventajas que nos proporciona la Internet cuántica como la integridad de la información (si se produjera una actualización en uno de los nodos, el consenso cuántico facilita la propagación de esa actualización a los demás nodos, manteniendo así la consistencia de los datos replicados), la detección de nodos defectuosos (los algoritmos de consenso cuántico pueden identificar discrepancias o comportamientos anómalos en los nodos replicados) o la mayor seguridad (gracias a las propiedades cuánticas, como la privacidad inherente y la imposibilidad de clonación) quedan fuera del estudio actual.

Si la sincronización se lleva a cabo en la Internet cuántica aprovechando la coordinación instantánea entre los sistemas, se espera que la Internet clásica tenga un papel secundario en el proceso de replicación distribuida. La Internet clásica seguirá desempeñando su función habitual de transporte de datos, pero la sincronización y la coordinación se realizarán a través de la Internet cuántica. En este escenario, se espera que la Internet clásica experimente una disminución en la carga de trabajo y en la complejidad de la comunicación en relación con la replicación distribuida. Dado que la coordinación instantánea y la sincronización eficiente se logran a través de la Internet cuántica, la Internet clásica se libera de la necesidad de manejar múltiples mensajes y protocolos complejos para lograr la sincronización entre los sistemas replicados. Esto puede conducir a una replicación distribuida más eficiente y confiable, ya que los sistemas pueden alcanzar un acuerdo más rápido y con menos sobrecarga de comunicación.

III. MODELO DEL SISTEMA

Suponiendo un modelo de sistema distribuido replicado como el que se propone en [2], con N nodos (o servidores), donde cada nodo $site_i$ tiene una capacidad máxima de procesamiento C (es decir, la cantidad de transacciones por segundo (tps) que puede procesar) y que cada nodo realiza una cantidad de trabajo local L_i, el factor de escalabilidad ($ScaleOut$) de la replicación de datos se calcula analíticamente cómo sigue [18]:

$$ScaleOut = \frac{\sum^{\forall site_i} L_i}{C}. \quad (1)$$

Este factor de escalabilidad permite determinar cómo evoluciona el rendimiento global o neto de un sistema distribuido a medida que se añaden (o quitan) nodos en el sistema. De esta forma, este parámetro da una noción del número equivalente de nodos que están dando servicio en un sistema de replicación de datos; valores de $ScaleOut$ próximos a N significa que los nodos están siendo capaces de dar todo el servicio que se les pide mientras que valores de $ScaleOut$ próximos a 0 significan que el sistema está congestionado y, por ende, no puede dar todo el servicio que se le pide.

Dado que idealmente la cantidad de trabajo local realizado en el nodo i debería ser igual a su capacidad máxima (C) ($L_i = C$), el factor de escalabilidad en un sistema totalmente replicado [19] con N nodos debería aumentar

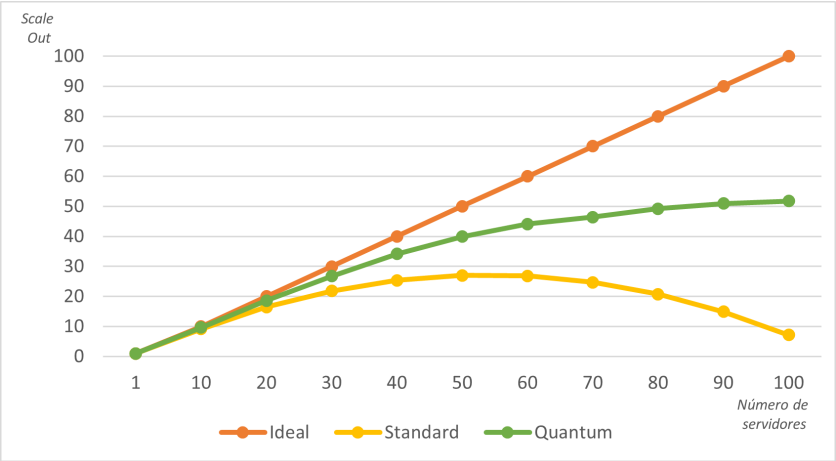


Fig. 1. Evolución del factor de escalabilidad en un sistema replicado en el que cada nodo puede ejecutar 4K transacciones por segundo y está expuesto a una carga constante de 150 transacciones por segundo que contienen operaciones de actualización.

linealmente según la Ecuación 2.

$$ScaleOut_{ideal} = \frac{\sum_{i=1}^{N} C}{C} = N. \quad (2)$$

Esto tiene sentido porque, idealmente, los datos almacenados en un sistema distribuido totalmente replicado con N nodos deberían estar disponibles N veces más que si los datos no estuvieran replicados. Es decir, si por ejemplo un nodo puede servir 100 peticiones por segundo, 7 nodos en un sistema totalmente replicado podrían servir, idealmente $7 * 10 = 70$ peticiones. Sin embargo, la Ecuación 2 no considera el costo de aplicar el protocolo de replicación en la base de datos (también conocido como sobrecarga del proceso de replicación) que permite sincronizar las réplicas de cada dato alojadas en distintos nodos. Por lo tanto, la cantidad total de trabajo realizado en el sitio i ($work_i$) se debe calcular de la siguiente manera:

$$work_i = LocalTransactions_i + RemoteTransactions_i. \quad (3)$$

De hecho, la cantidad de trabajo local en el nodo i (L_i) será la suma de (a) el trabajo derivado de las transacciones de lectura y/o actualización locales emitidas para los elementos almacenados en el nodo i ($LocalTransactions_i$) más (b) el trabajo de sincronización derivado de las transacciones de actualización remotas asociadas al protocolo de replicación y realizadas contra los datos almacenados en el nodo i ($RemoteTransactions_i$). Recordemos que en un sistema totalmente replicado con consistencia fuerte [5] se puede asumir que las operaciones de lectura de datos no necesitan sincronizarse con otros nodos y por lo tanto no generan más sobrecarga (i.e., sólo afectan a la carga local del nodo). Por lo tanto, podemos modelar el costo de aplicar las operaciones emitidas contra el nodo $site_i$ con los siguientes parámetros:

- R_i: Cantidad de transacciones por segundo de sólo lectura emitidas contra el nodo $site_i$ desde los clientes.
- LU_i: Cantidad de transacciones por segundo de actualización emitidas contra el nodo $site_i$ desde los clientes.
- RU_i: Cantidad de transacciones por segundo asociadas a la sincronización de operaciones de actualización emitidas contra el nodo $site_i$ desde los otros nodos.
- w_o: Este parámetro modela el coste tanto a nivel de red (p.ej., congestión) como a nivel de nodo (p.ej., aplicar el resultado de transacciones ya ejecutadas por la replica delegada en replicación pasiva [2]) de ejecutar las operaciones de sincronización remotas enviadas por los otros nodos del sistema.
- R_d: Número de nodos que almacenan réplicas remotas de cada dato.
- C_i: Cantidad máxima de transacciones por segundo que el nodo $site_i$ puede procesar.

Por lo tanto, la Ecuación 3 puede particularizarse de la siguiente manera:

$$work_i = \underbrace{R_i + LU_i}_{LocalTransactions_i} + \underbrace{w_o \cdot (R_d - 1) \cdot RU_i}_{RemoteTransactions_i} \leq C_i. \quad (4)$$

Recuérdese que el factor de escalabilidad solo considera la cantidad de trabajo local neta realizado en cada nodo (es decir, transacciones locales de lectura/actualización que se pueden ejecutar después de ejecutar las operaciones de sincronización con otros nodos). Suponiendo que todos los nodos funcionan al máximo rendimiento ($C_i = work_i$), de acuerdo con la Ecuación 4, la cantidad de trabajo local que se puede ejecutar en el nodo $site_i$ se puede expresar de la siguiente manera:

$$L_i = C_i - w_o \cdot (R_d - 1) \cdot RU_i. \quad (5)$$

A partir de las Ecuaciones 1 y 5, se puede obtener la fórmula final para calcular el factor de escalabilidad en un sistema distribuido replicado:

$$ScaleOut = \frac{1}{C} \sum_{i=1}^{N} (C_i - w_o \cdot (R_d - 1) \cdot RU_i). \quad (6)$$

Esta es una fórmula general que se puede aplicar a distintos escenarios si se ajusta adecuadamente. En la siguiente sección se discute como afecta el uso de la Internet cuántica a este modelo analítico.

IV. DISCUSIÓN

En la Fig. 1 se muestra la evolución del factor de escalabilidad en un sistema replicado según el modelo analítico propuesto en la sección anterior. Los valores característicos del sistema (C, C_i, R_d, w_0, RU_i) se han elegido al azar para poder enfatizar el comportamiento del modelo utilizado en esta investigación (i.e. mejorar la escalabilidad de la replicación de datos gracias a la Internet cuántica). Así, se ha elegido una situación de replicación total [2] en la que todos los nodos almacenan una copia de cada dato, cada nodo puede ejecutar hasta 4,000 transacciones por segundo y está expuesto a una carga constante de 150 transacciones por segundo que solo contienen operaciones de actualización.

Por un lado, en la Fig. 1 se puede ver (en naranja) la curva ideal del factor de escalabilidad. Idealmente, añadir más nodos al sistema no genera más tráfico ni congestión en la red así como tampoco se sobrecarga al resto de nodos con las operaciones de sincronización de datos ($w_0 = 0$ en la Ecuación 6). Este comportamiento es el que se obtendría si todas las transacciones fueran de lectura. Por otro lado, en la Fig. 1 también se puede ver (en amarillo) la evolución típica del factor de escalabilidad para un protocolo de replicación de datos. A medida que el número de réplicas crece, el número de mensajes de sincronización que deben intercambiar los nodos también crece, lo que añade más congestión a la red y sobrecarga de trabajo a los nodos. Cuando la pendiente del factor de escalabilidad es negativa, significa que el sistema está congestionado y que los nodos están más tiempo procesando mensajes de sincronización que no transacciones de los clientes.

Finalmente, en la Fig. 1 también se puede apreciar (en verde) que gracias al eficiente mecanismo de sincronización del algoritmo de consenso cuántico el comportamiento del procedimiento de escritura en el sistema de replicación se acerca al óptimo que se obtendría en el procedimiento de lectura. La razón principal es que la red de datos no llega a congestionarse en este punto y el límite teórico alcanzado se debe en exclusiva a la limitada capacidad de procesamiento de los nodos. A partir de los 60 nodos esta curva pone en evidencia que el sistema se satura por el tráfico debido a la propia replicación congestionando finalmente la red de datos aunque se haya evitado la sobrecarga producida por el mecanismo de sincronización entre los participantes en primera instancia.

V. TRABAJO FUTURO

En el escenario descrito, el cuello de botella podría estar en la transición entre la Internet clásica y la Internet cuántica. Para ello se debería enriquecer el modelo con más propiedades y características de la red. Aunque la Internet cuántica ofrece ventajas en términos de coordinación instantánea y eficiencia en la sincronización, aún estamos en las primeras etapas de desarrollo de la tecnología cuántica y de la infraestructura necesaria para su implementación a gran escala. La transmisión de información entre los sistemas de replicación de datos a través de la Internet clásica hacia la Internet cuántica puede requerir interfaces y protocolos específicos que aún están en proceso de desarrollo.

REFERENCIAS

[1] R. Guerraoui and A. Schiper, "Fault-tolerance by replication in distributed systems," in *Reliable Software Technologies—Ada-Europe'96: 1996 Ada-Europe International Conference on Reliable Software Technologies Montreux, Switzerland, June 10–14, 1996 Proceedings 1*. Springer, 1996, pp. 38–57.

[2] M. Wiesmann and A. Schiper, "Comparison of database replication techniques based on total order broadcast," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 551–566, 2005.

[3] J. Gray, P. Helland, P. O'Neil, and D. Shasha, "The dangers of replication and a solution," in *Proceedings of the 1996 ACM SIGMOD international Conference on Management of Data*, 1996, pp. 173–182.

[4] B. Kemme, "Database replication for clusters of workstations," Ph.D. dissertation, ETH Zurich, 2000.

[5] W. Vogels, "Eventually consistent," *Communications of the ACM*, vol. 52, no. 1, pp. 40–44, 2009.

[6] S. Bottoni, S. Braghin, A. Trombetta, and S. Venugopal, "Adaptive replication strategy in highly distributed data management systems," in *2022 IEEE International Conference on Cloud Engineering (IC2E)*. IEEE, 2022, pp. 273–274.

[7] X. Zhou, X. Yu, G. Graefe, and M. Stonebraker, "Lotus: scalable multi-partition transactions on single-threaded partitioned databases," *Proceedings of the VLDB Endowment*, vol. 15, no. 11, pp. 2939–2952, 2022.

[8] G. Kim and W. Lee, "In-network leaderless replication for distributed data stores," *Proceedings of the VLDB Endowment*, vol. 15, no. 7, pp. 1337–1349, 2022.

[9] E. A. Lee, S. Bateni, S. Lin, M. Lohstroh, and C. Menard, "Quantifying and generalizing the cap theorem," *arXiv preprint arXiv:2109.07771*, 2021.

[10] M. Marcozzi and L. Mostarda, "Quantum consensus: an overview," *arXiv preprint arXiv:2101.04192*, 2021.

[11] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, "Towards a distributed quantum computing ecosystem," *IET Quantum Communication*, vol. 1, no. 1, pp. 3–8, 2020.

[12] A. S. Cacciapuoti, M. Caleffi, R. Van Meter, and L. Hanzo, "When entanglement meets classical communications: Quantum teleportation for the quantum internet," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3808–3833, 2020.

[13] A. Singh, K. Dev, H. Siljak, H. D. Joshi, and M. Magarini, "Quantum internet—applications, functionalities, enabling technologies, challenges, and research directions," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2218–2247, 2021.

[14] P. Botsinis, D. Alanis, Z. Babar, H. V. Nguyen, D. Chandra, S. X. Ng, and L. Hanzo, "Quantum search algorithms for wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1209–1242, 2018.

[15] A. Zaballos, A. Mallorqui, and J. Navarro, "Unboxing trustworthiness through quantum internet," *arXiv preprint arXiv:2210.10687*, 2022.

[16] S. Tani, H. Kobayashi, and K. Matsumoto, "Exact quantum algorithms for the leader election problem," *ACM Transactions on Computation Theory (TOCT)*, vol. 4, no. 1, pp. 1–24, 2012.

[17] M. Ben-Or and A. Hassidim, "Fast quantum byzantine agreement," in *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, 2005, pp. 481–485.

[18] D. Serrano, M. Patiño-Martínez, R. Jiménez-Peris, and B. Kemme, "Boosting database replication scalability through partial replication and 1-copy-snapshot-isolation," in *13th Pacific Rim International Symposium on Dependable Computing (PRDC 2007)*. IEEE, 2007, pp. 290–297.

[19] A. R. Nasibullin and B. A. Novikov, "Replication in distributed systems: Models, methods, and protocols," *Programming and Computer Software*, vol. 46, pp. 341–350, 2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Reducción de la latencia en redes 5G-NR

Felix Delgado-Ferro, Jorge Navarro-Ortiz, Natalia Chinchilla-Romero, Juan Manuel Lopez-Soler
Department of Signal Theory, Telematics and Communications
University of Granada
Av. del Hospicio, 1, 18012 Granada.
felixdelgado@ugr.es, jorgenavarro@ugr.es, nataliachr@ugr.es, juanma@ugr.es

Este trabajo presenta un estudio que tiene como objetivo reducir la latencia en función de los parámetros más relevantes de la capa física 5G *New Radio (NR)*. En primer lugar, se explican los conceptos teóricos fundamentales, relacionados con el *subcarrier spacing (SCS)* y la estructura de las tramas. Después se describe el entorno de experimentación, compuesto por una red 5G basada en *Amarisoft* y un equipo de usuario con modem Quectel. Por último, se presentan resultados preliminares que muestran cómo afectan los parámetros seleccionados. Como conclusión principal, el *subcarrier spacing* y el uso de tramas flexibles son los parámetros fundamentales para reducir la latencia en 5G.

Palabras Clave*—5G New Radio, latencia, reducción, numerología y *subcarrier spacing

I. INTRODUCCIÓN

La ITU (*International Telecommunication Union*) definió las recomendaciones IMT-2020 (*International Mobile Telecommunications*) [1]. Considerando estas recomendaciones, el 3GPP (*3rd Generation Partnership Project*) definió tres posibles servicios que incluyen todos los casos de uso conocidos actualmente [2]. Estos servicios son: *enhanced Mobile Broadband (eMBB)*, *massive Machine Type communications (mMTC)* y *Ultra-Reliable Low Latency Communications (URLLC)*.

Entre los principales retos de las redes 5G, destacan los servicios URLLC debido a sus estrictos requisitos. Estos requisitos consisten en proporcionar una transmisión de datos con una fiabilidad del 99,99 % y una latencia por debajo del milisegundo. Existen múltiples casos de uso que deben ser soportados por este tipo de servicio [3]. Además, los avances actuales, que pretenden satisfacer estos requisitos, se han centrado en aumentar el ancho de banda -generalmente centrado en el enlace descendente-. Sin embargo, las necesidades del tráfico están cambiando. Esto implica que las nuevas aplicaciones requieren una mayor flexibilidad en la asignación de recursos debido al uso de más recursos del enlace ascendente. En consecuencia, la duplexación por división en el tiempo (TDD) es cada vez más popular y significativa, ya que permite aumentar la flexibilidad y la eficiencia espectral [4]. Por estas razones, en nuestro trabajo nos centraremos en la configuración TDD para reducir la latencia, siguiendo las especificaciones técnicas de la *Release 15* [5].

El resto del documento se organiza como sigue. La Sección II proporciona una revisión de los trabajos relacionados. A continuación, en la Sección III se presentan los fundamentos teóricos de la capa física 5G NR. La Sección IV describe el montaje y explica el diseño experimental, mientras que la Sección V muestra el análisis en términos de latencia. Finalmente, la Sección VI presenta las conclusiones de este trabajo.

II. TRABAJOS RELACIONADOS

Esta sección presenta las contribuciones relacionadas con nuestro trabajo. Principalmente, estas contribuciones se centran en el análisis del rendimiento utilizando mecanismos de optimización sobre la capa física 5G NR.

Ali *et al.* [6] presentan un estudio del *Physical Uplink Shared Channel (PUSCH)* teniendo en cuenta parámetros como los diferentes *Subcarrier Spacing (SCS)*, las técnicas de modulación (QPSK o QAM) y el número de antenas de la *Base Station (BS)* y del *user equipment (UE)*. Mhedhbi *et al.* [7] se centraron en aprovechar las características NR de las tramas y evaluar el impacto de las configuraciones radio y el ancho de banda reservado para el tráfico crítico en los servicios URLLC. Chinchilla-Romero *et al.* [8] evaluaron la ganancia de rendimiento al incluir 5G a un escenario con multi-conectividad y lo extrapolaron a un posible aumento en redes 5G, basándose en sus requisitos y numerología. Por último, Mohamed *et al.* [9] abordaron la gestión de los recursos radio empleando mini-slots definidos para las redes 5G NR. Estos permiten reducir la latencia de retransmisión a costa de reducir el rendimiento del enlace.

En cambio, nuestro trabajo presenta un análisis de la latencia en función del ancho de banda dedicado, el *Subcarrier Spacing (SCS)*, la estructura de las tramas y *slots*. Además, la experimentación se realiza empleando

una estación base real; mientras que los trabajos citados anteriormente se realizan en escenarios simulados.

III. CAPA FÍSICA DE 5G NR

La pila de protocolos de 5G NR consta de varias capas que trabajan juntas para permitir la comunicación entre dispositivos y equipos de red. Las principales capas son: *Physical (PHY), Medium Access Control (MAC), Radio Link Control (RLC), Packet Data Convergence Protocol (PDCP), Radio Resource Control (RRC)* y *Service Data Adaptation Protocol (SDAP)* [10]. Además, existen protocolos de capa superior que permiten comunicaciones extremo a extremo.

La capa física de 5G NR está diseñada a partir de la capa física de 4G LTE. Del mismo modo, ambas utilizan OFDM con prefijo cíclico (CP) para el enlace descendente (DL). En 5G NR, OFDM también puede utilizarse para el enlace ascendente (UL), aunque se establecen dos mecanismos nuevos: *Peak to Average Power Ratio (PAPR)* y OFDM con precodificación por transformada discreta de Fourier (DFT-s-OFDM) [10]. Además, NR admite dos rangos de frecuencias: FR1 (410 MHz - 7125 MHz) y FR2 (24250 MHz - 52600 MHz). Cada rango de frecuencias admite anchos de banda de canal específicos [11].

El ancho de banda (BW) de NR se compone de un conjunto de *Resource Blocks (RB)*. Un RB es un conjunto de 12 subportadoras. Una subportadora -en el dominio de la frecuencia- o un símbolo OFDM -en el dominio del tiempo- se denomina *Resource Element (RE)*.

A. Numerología

En 5G [12], la distancia entre portadoras se especifica como $\Delta f_{ref} = 15KHz$. NR define múltiples numerologías OFDM. Cada numerología está determinada por el valor μ, que está directamente relacionado con el *subcarrier spacing* ($SCS = 2^{\mu} * \Delta f_{ref}$) para un enlace ascendente o descendente y el prefijo cíclico. La tabla I muestra las múltiples numerologías disponibles para FR1.

Tabla I
NUMEROLOGÍA EN NR PARA FR1 [11][12].

μ	SCS $[KHz]$	CP	$N_{slots}^{subframe,\mu}$	N_{symb}^{slot}
0	15	Normal	1	14
1	30	Normal	2	14
2	60	Normal/Extended	4	14/12

Cabe destacar que la numerología se introduce en 5G NR, mientras que en 4G LTE el *subcarrier spacing* es siempre $SCS_{LTE} = 2^{\mu} * \Delta f_{ref} = 15KHz$ (caso $\mu = 0$).

B. Estructura de la trama

Las transmisiones en NR se organizan en tramas flexibles. En el dominio del tiempo, la duración de la trama se define como $T_{\mathrm{f}} = (\Delta f_{\mathrm{máx}} N_{\mathrm{f}}/100) \cdot T_{\mathrm{c}} = 10$ ms [12]. Cada trama se subdivide en diez subtramas con duración $T_{\mathrm{sf}} = (\Delta f_{\mathrm{máx}} N_{\mathrm{f}}/1000) \cdot T_{\mathrm{c}} = 1$ ms

Además, cada subtrama consta de $N_{slots}^{subtrama,\mu} \in \{1, 2, 4, 8, 16\}$ *slots*, o ranuras temporales, dependiendo de la SCS o numerología seleccionada (ver Tabla I). La Fig. 1 muestra los casos de valores de SCS de 15; 30 y 60 de numerologías posibles en NR para FR1. Además, cada *slot* consta de 14 símbolos OFDM (independientemente del SCS) precedidos de un prefijo cíclico.

C. Configuración del Slot

NR admite el funcionamiento tanto FDD como TDD empleando la misma estructura de trama. En TDD, que permite la adaptación flexible del tráfico, cada símbolo OFDM de un *slot* se puede clasificar como *"downlink", üplink"* o *"flexible"*.

El gNB debe proporcionar estas configuraciones al *UE* mediante el parámetro *TDD-UL-DL-ConfigurationCommon*. A continuación, el UE establece el formato *slot-by-slot* para un número especificado de *slots*. La información proporcionada por el parámetro *TDD-UL-DL-ConfigurationCommon* es la siguiente:

- Configuración del *subcarrier spacing* de referencia μ_{ref} mediante el parámetro *referenceSubcarrierSpacing* (see Table I).
- *Pattern 1* y opcional *Pattern 2* proporcionan la periodicidad (P) del patrón DL-UL en [ms], el número de *slots* DL completos y consecutivos al principio de cada patrón DL-UL (n_{slots}^{DL}), el número de símbolos DL consecutivos al principio del *slot* que sigue al último *slot* DL (n_{symb}^{DL}), el número de *slot* UL completas y consecutivas al final de cada patrón DL-UL (n_{slots}^{UL}); y número de símbolos UL consecutivos al final del *slot* que precede al primer *slot* UL (n_{symb}^{UL}).

En [5] se presentan los formatos de *slot* para prefijo cíclico normal que se establecen mediante *Pattern 1*. Si la configuración sólo utiliza este patrón, entonces la periodicidad (P) incluye $n_{slots} = P{*}2^{\mu_{ref}}$ *slots*, donde μ_{ref} es la configuración del *subcarrier spacing*. La periodicidad (P) del patrón DL-UL cambia en función de la numerología de referencia (μ_{ref}). El cálculo P/20 identifica la posición del primer símbolo de una trama par que debe coincidir con el primer símbolo de cada periodo para que estén sincronizados. Esto se resume en una tabla II.

El *Pattern 2* puede utilizarse opcionalmente. En este caso, la periodicidad es ($P' = P + P_2$). El número de *slots* para el *Pattern 2* seguirá la misma lógica que el *Pattern 1*. El uso de dos patrones permite transmitir diferentes cantidades de *slots* de enlace descendente y ascendente en

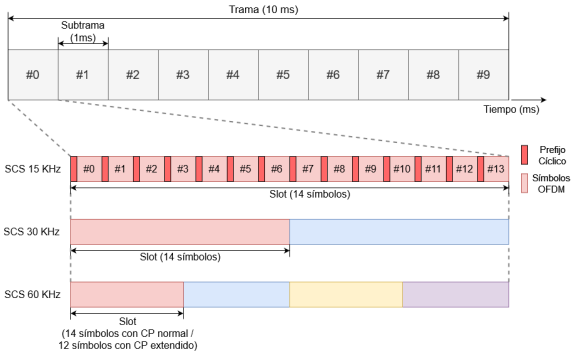


Figura 1. Estructura de la trama en NR para FR1

tramas consecutivas -sin necesidad de modificar el patrón del UE-. El usuario espera que esto se transmita en dos tramas consecutivas ($P' = P + P_2 = 20\text{ms}$) y que el ajuste del *subcarrier spacing* μ sea mayor o igual que el valor de referencia μ_{ref} para ambos patrones. La Fig. 2 muestra un ejemplo de la configuración cuando se utilizan ambos patrones.

Tabla II
PERIODICIDAD DE TRANSMISIÓN APLICABLE (BASADA EN [5])

$P[ms]$	μ	SCS $[KHz]$	$P/20$
0,625	3	120	32
1,25	2/3	60/120	16
2,5	1/2/3	30/60/120	8
5	-	Not described	4
10	-	Not described	2

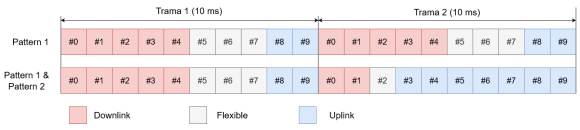


Figura 2. Ejemplo de configuración del *Pattern 1* y *Pattern 2*

Además, mencionar que a partir de la *Release 16* se ha definido un nuevo parámetro *tdd-UL-DL-ConfigDedicated* que mejorará las prestaciones en servicios URLLC [5].

IV. EXPERIMENTOS

La experimentación para este trabajo se ha realizado en el *testbed* que se muestra en la Fig. 3. Concretamente, se han realizado los experimentos en la conexión inalámbrica 5G NR. Esta conexión, que aparece destacada en la Fig. 3, se establece empleando una gNB, desplegada con el software de *Amarisoft* con version 2023-03-17, y un módem Quectel RM500Q-GL (v1.6) que permite la conexión mediante 5G, conectado al Intel NUC que actúa como *Customer Provided Equipment (CPE)*.

Figura 3. Testbed para conectividad de baja latencia

Cabe mencionar que la red 5G NR está deplegada en modo *standalone (SA)*. Por otro lado, debido a tener que transmitir utilizando bandas licenciadas, este radioenlace estará contenido dentro de una jaula de Faraday para no contaminar el espectro radioeléctrico.

Previo al proceso de experimentación, se decidieron los parámetros de configuración de la capa física 5G NR que se modificarían para comprobar su impacto y poder caracterizar la conexión en términos de latencia. Estas configuraciones aparecen en la Tabla III.

Para los valores correspondientes a la estructura de la trama se ha optado por emplear configuraciones definidas en [5]. Estos se describen en la Tabla IV.

Tabla III
CONFIGURACIONES DE LOS EXPERIMENTOS

Parámetro	Acrónimo	Valores
Ancho de banda	BW	20/40 MHz
Subcarrier Spacing	SCS	15/30 KHz
Estructura de la trama	FS	1/2/3/4

Tabla IV
CONFIGURACIONES DE LAS TRAMAS

FS	Patterns	n^{DL}_{slots}	n^{DL}_{symb}	n^{UL}_{symb}	n^{UL}_{slots}
1	*Pattern 1*	7	2	2	2
2	*Pattern 1*	7	6	4	2
3	*Pattern 1*	6	2	2	3
4	*Pattern 1*	3	6	2	4
	Pattern 2	4	0	0	0

Este trabajo se ha centrado en el estudio de la latencia en función de cada uno de los parámetros citados previamente. Debido a las múltiples configuraciones posibles se ha establecido una configuración por defecto para poder comparar de forma coherente. Esta configuración es: $BW = 20$ MHz; $SCS = 30$ KHz y $FS = 2$. A parte de estas configuraciones, se han dejado estáticas el número de antenas (MIMO 2x1) y la banda de frecuencia empleada para TDD (n41).

V. RESULTADOS PRELIMINARES

Como se ha comentado, los resultados se centran principalmente en el análisis de la latencia, dado que es unos de los indicadores críticos para el despliegue en el futuro de servicios eMBB y URLLC.

A. *Ancho de banda*

La comparativa de la latencia al emplear las configuraciones de ancho de banda establecidas en la Tabla III se muestran en la Fig. 4. Observando esta *Función de Distribución Acumulada (CDF)*, podemos comprobar como el aumento del BW mejora la latencia de la comunicación en media, pero el rango de latencias obtenidas son similares [6-7.5 ms].

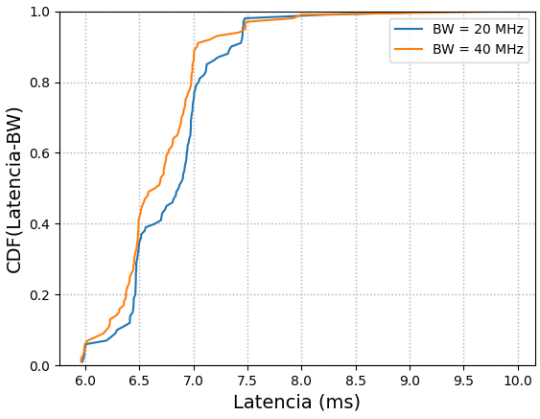


Figura 4. CDF de la latencia en función del BW

B. Subcarrier Spacing (SCS)

El *Subcarrier Spacing (SCS)* es un parámetro muy significativo a la hora de mejorar la latencia, como se aprecia en la Fig. 5. Además, la numerología y el uso de tramas flexibles en 5G, empleando diferentes periodicidades y *subcarrier spacing*, conlleva una mejora significativa de la latencia frente a la configuración estándar de 4G o 5G *Non Standalone (NSA)*. Esta mejora implica una reducción de la latencia de aproximadamente el 50 % y existen SCS superiores que deben tener un comportamiento aún mejor.

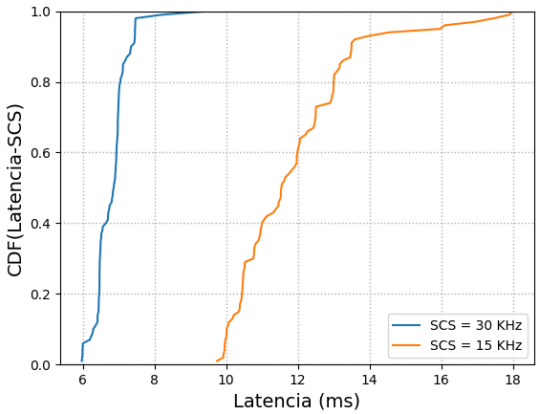


Figura 5. CDF de la latencia en función del *Subcarrier Spacing*

C. Estructura de Trama

Las estructuras de las tramas presentan un rango de latencias comprendido entre 6 y 7.5 ms. Cabe destacar que el caso con $FS = 3$ reduce significativamente la latencia frente al resto y su rango se encuentra entre 6 y 7 ms, como se puede comprobar sobre la Fig. 6. Esto implica que la latencia en este caso es más constante y menor debido a la asignación de más recursos al UL y menos al DL.

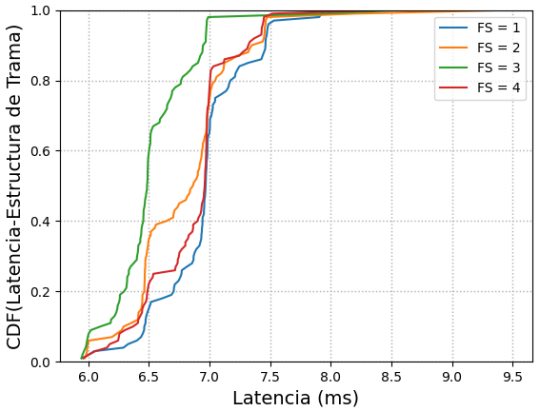


Figura 6. CDF de la latencia en función de la estructura de trama

VI. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se presenta un estudio de la latencia en la capa física 5G NR. De acuerdo a los resultados, se concluye que el *subcarrier spacing* y la estructura de trama afectan significativamente a la latencia; también concluimos que al emplear tramas flexibles, usando mayor numerología, la latencia puede reducirse abruptamente (≈50 %). Por tanto, se considera que, aunque actualmente en este trabajo no se presentan latencias inferiores al milisegundo, estas latencias serán posibles con las futuras mejoras introducidas en Release 16.

Respecto a los trabajos futuros, se están considerando emplear dispositivos que nos permitan aumentar el *subcarrier spacing*, emplear FR2 para aumentar el ancho de banda y reducir la periodicidad de las tramas. Y, por último, realizar un estudio de cómo estos factores afectan también al *throughput* en DL y UL.

AGRADECIMIENTOS

Este trabajo está financiado por el Ministerio español de Economía y Transformación Digital (TSI-063000-2021-28); el Ministerio español de Ciencia e Innovación (PID2019-108713RB-C53 y PID2022-137329OB-C43); y el Ministerio de Universidades español (Beca FPU: 20/02621)

REFERENCIAS

[1] Recommendation ITU-R M.2083: IMT Vision - "Framework and overall objectives of the future development of IMT for 2020 and beyond", Sept. 2015.

[2] Technical Report TR38.912 V15.0.0, "5G; Study on scenarios and requirements for next generation access technologies,"3GPP, Sept. 2019.

[3] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz and J. M. Lopez-Soler, .A Survey on 5G Usage Scenarios and Traffic Models,ïn *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 905-929, Secondquarter 2020, doi: 10.1109/COMST.2020.2971781

[4] GSMA, "5G TDD Synchronisation Guidelines and Recommendations for the Coexistence of TDD Networks in the 3.5 GHz Range,"Technical document, 2020.

[5] Technical Specification TS38.213, V15.3.0, "5G; NR; Physical layer procedures for control"3GPP, Oct. 2018.

[6] M. Ali and Y. Kabalci, "Throughput Analysis over 5G NR Physical Uplink Shared Channels,"*2020 2nd Global Power, Energy and Communication Conference (IEEE-GPECOM-2020)*, Nov. 2020, doi: 10.1109/GPECOM49333.2020.9247906

[7] M. Mhedhbi, M. Morcos, A. Galindo-Serrano and S. Elayoubi, "Performance Evaluation of 5G Radio Configurations for Industry 4.0,"*15th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2019)*, Oct 2019, Barcelone, Spain. HAL id: hal-03687385

[8] L. Chinchilla-Romero, J. Prados-Garzon, P. Muñoz, P. Ameigeiras and J. M. Lopez-Soler, ÜRLLC Achieved Data Rate through Exploiting Multi-Connectivity in Industrial Private 5G Networks with Multi-WAT RANs,"2023 IEEE Wireless Communications and Networking Conference (WCNC 2023), Glasgow, United Kingdom, 2023, pp. 1-6, doi: 10.1109/WCNC55385.2023.10119085.

[9] A. Mohamed, A. Quddus, P. Xiao, B. Hunt and R. Tafazolli, "5G and LTE-TDD Synchronized Coexistence with Blind Retransmission and Mini-Slot Uplink,"*2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, 2020, pp. 1-5, doi: 10.1109/VTC2020-Spring48590.2020.9128796.

[10] Technical Specification TS38.300, V16.4.0, "5G; NR and NG-RAN Overall description; Stage-2"3GPP, Jan. 2021.

[11] Technical Report TR38.912 V15.0.0, "5G; Study on scenarios and requirements for next generation access technologies,"3GPP, Sept. 2019.

[12] Technical Specification TS38.211, V16.2.0, "5G; NR; Physical channels and modulation"3GPP, Jul. 2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Despliegue Óptimo de Servicios IoT en Redes Inalámbricas basadas en Enjambres de UAV para Minimizar la Latencia

Santiago García Gil*, Juan Manuel Murillo†, Jaime Galán-Jiménez*

*Departamento de Ingeniería de Sistemas Informáticos y Telemáticos, Universidad de Extremadura, España.

† Cénits-COMPUTAEX, España.

santiagogg@unex.es, juan.murillo@cenits.es, jaime@unex.es

Los enjambres de vehículos aéreos no tripulados (*Unmanned Aerial Vehicles*, UAVs), el paradigma del IoT, las arquitecturas orientadas a servicios (*Service Oriented Architectures*, SOAs) y las redes celulares 5G desvelan formas sin precedentes de proporcionar conectividad a dispositivos situados en lugares remotos donde las infraestructuras de red tradicionales no son viables. Además, permiten situar los elementos de computación más cerca que nunca de los usuarios. Sin embargo, estas nuevas tecnologías traen consigo nuevos problemas. Ofrecer conectividad a dispositivos y personas situados en lugares de difícil acceso sin cumplir las restricciones de calidad de servicio (*Quality of Service*, QoS), impuestas por la naturaleza de las tareas que deben realizar, sería un esfuerzo en balde. Este trabajo se centra en la distribución óptima de microservicios en redes desplegadas mediante enjambres de UAVs con recursos limitados con el objetivo de minimizar la latencia de aplicaciones IoT. Existen aproximaciones basadas en heurísticas y otras basadas en aprendizaje profundo por refuerzo, pero éstas pueden acabar conduciendo a soluciones subóptimas. Para evitarlo, se sigue una formulación del problema basada en programación lineal entera mixta (*Mixed Integer Linear Programming*, MILP). Los resultados de los experimentos muestran la efectividad del modelo a la hora de desplegar servicios en escenarios IoT con redes inalámbricas implementadas mediante enjambres de UAV de forma que la latencia sea mínima.

***Palabras Clave*—UAV, IoT, latencia, microservicios, MILP**

I. INTRODUCCIÓN

Gracias a la transición hacia redes móviles 5G, que cada vez toma más inercia, la visión de una sociedad totalmente conectada está más cerca que nunca. Además, con la proliferación de los vehículos aéreos no tripulados (*Unmanned Aerial Vehicles*, UAVs), se abren nuevos caminos hacia despliegues de infraestructura de redes móviles. Gracias a otras técnicas, como las redes definidas por software (*Software Defined Networks*, SDNs), las funciones de red virtualizadas (*Virtual Network Functions*, VNFs) y las arquitecturas orientadas a servicios (*Service Oriented Arquitectures*, SOAs), la computación puede acercarse a los usuarios finales, lo que permite la ejecución ubicua y en tiempo real de servicios que pueden consumir muchos recursos y tiempo. Las posibilidades que plantean estas tecnologías son infinitas, pero también hay una larga lista de problemas que deben abordarse.

El uso de UAVs para la mejora de las redes celulares es uno de los temas de investigación emergentes en el área de las redes de comunicaciones [1]. Esta estrategia resulta eficaz para reducir la latencia global de la red si los UAV están bien situados. Este mecanismo puede desempeñar un papel importante en escenarios con estrictas exigencias de calidad de servicio (*Quality of Service*, QoS). Las aplicaciones en tiempo real acríticas, como el *streaming* de audio o vídeo, los videojuegos en línea o la voz sobre IP (*Voice over IP*, VoIP), constituyen excelentes ejemplos. Sin embargo, es en las aplicaciones de tiempo real críticas dónde se hace evidente la importancia de este enfoque. Los dispositivos de control sanitario o la conducción autónoma de vehículos agrícolas, por ejemplo, dependen de una conectividad rápida y fiable para ser eficaces en sus tareas.

El posicionamiento de los UAVs desempeña un papel crucial en el rendimiento de la red, pero no es el único aspecto a tener en cuenta. Dada una distribución óptima de los UAVs para un escenario concreto, uno de los siguientes problemas que se plantea es el despliegue de los servicios que se van a ofrecer. Este tipo de problema se conoce como Problema de Distribución Computacional Descentralizada (*Decentralized Computation Distribution Problem*, DCDP) [2] [3].

La contribución de este trabajo es un modelo novedoso cuyo objetivo es desplegar de forma óptima, en términos de minimización de latencia, un conjunto dado

de microservicios sobre una red inalámbrica con recursos limitados compuesta por un enjambre de UAVs. Para ello, el problema se formula matemáticamente y el modelo se implementa mediante Programación Lineal Entera Mixta (*Mixed Integer Linear Programming*, MILP). Además, utilizando Kathará [4], se proponen y evalúan emulaciones de soluciones a diferentes escenarios.

El resto del artículo se organiza de la siguiente manera: las motivaciones y el escenario base se describen en la Sec. II. La formulación del problema se define en la Sección III. Los resultados de los experimentos se evalúan en la Sección IV. En la Sección V se analizan trabajos relacionados. Por último, la Sección VI concluye el artículo.

II. MOTIVACIÓN Y ESCENARIO

En esta sección, se analiza el escenario que propicia la creación del modelo. Este debe proporcionar un despliegue óptimo de microservicios en redes implementadas mediante enjambres de UAVs con recursos de computación limitados.

Para avanzar hacia una sociedad totalmente conectada que aproveche al máximo las ventajas del paradigma IoT, se necesita una conectividad fiable para cada dispositivo que desempeñe un papel en la escena. Esta premisa es más fácil de cumplir en los grandes núcleos de población, donde el despliegue y mantenimiento de las infraestructuras tradicionales de redes celulares, que implican múltiples estaciones base (*Base Station*, BS) para cubrir toda la zona, resulta rentable. Para los usuarios de aquellas redes de centros densamente poblados, las restricciones de QoS pueden cumplirse con facilidad. Sin embargo, ¿qué ocurre con los lugares donde la población es reducida y está dispersa sobre un gran área, como las zonas rurales, y por tanto esta premisa no se cumple?

En núcleos de población dispersos puede que sólo haya una única BS compartida entre ellos. Esto lleva a la siguiente situación: para los usuarios que estén cerca de la BS se cumplirán las restricciones de QoS, pero a medida que los usuarios se alejen, la conectividad empeorará.

Este no es el único problema que hay que tener en cuenta. Para ciertas aplicaciones, la ventana de tiempo de respuesta de sus peticiones a los microservicios son muy restringidas e incluso cuando los usuarios o dispositivos están cerca de la BS, el tiempo de servir sus peticiones desde la nube puede no ser lo suficientemente rápido. Colocar la computación más cerca de los usuarios finales es la única solución real a este problema. Esto se conoce como computación en la niebla o en el borde (*fog computing* y *edge computing*), en función de lo cerca que estén los recursos de los usuarios finales. En el trabajo realizado por Tang et al. [5], se propone una solución para desplegar UAVs como servidores *Fog* para reducir la latencia minimizando el número de nodos necesarios siempre que se cumplan las restricciones de QoS. Sin embargo, formulan el problema de tal forma que pierden la linealidad haciendo que no sea factible obtener soluciones óptimas en tiempos razonables.

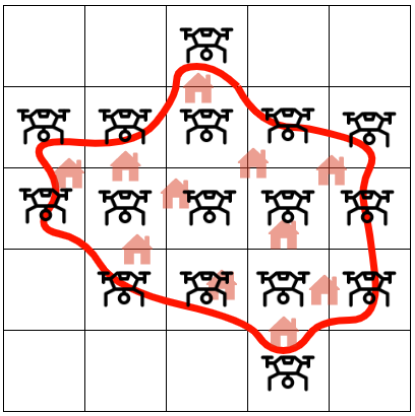


Fig. 1: Despliegue de un ejambre de UAVs sobre una población ficticia.

En esta situación, entran en juego despliegues de red no convencionales más flexibles. El uso de enjambres de UAVs parece encajar a la perfección. Pueden utilizarse para ampliar la zona de cobertura de las BS [6]. Además, la computación puede descargarse de la nube hacia los UAVs mediante el despliegue de microservicios en ellos, proporcionando así mejores tiempos de respuesta para los usuarios [7].

El planteamiento de utilizar enjambres de UAVs no está exento de limitaciones y hay que tenerlas en cuenta. La más obvia es que la capacidad de cálculo de estos dispositivos está severamente restringidas, no sólo porque el hardware de computación está diseñado para hacer un uso moderado de la batería del dispositivo, sino por las razones que se van a exponer a continuación. En primer lugar, tienen que ceñirse a una misión de vuelo [8], que es una aplicación en tiempo real. Además, si se van a utilizar para ampliar la cobertura de la red, deben desempeñar el papel de nodos de red que ejecutan algunas VNFs. Sólo después de realizar las tareas anteriores, las capacidades de computación restantes del UAV pueden asignarse al despliegue y ejecución de microservicios.

El correcto posicionamiento de los microservicios es una tarea delicada ya que, debido a las limitaciones expuestas, no hay mucho margen para replicarlos. Para ilustrar este hecho, se propone el escenario sintético mostrado en la Figura 1. En este escenario, un núcleo de población, cuyo perímetro está delimitado por la línea roja cerrada, se muestra sobre una cuadrícula. Cada celda de la cuadrícula denota el área que podría cubrir un UAV si se desplegara uno allí. Hay UAV situados en cada celda que contiene una parte del centro de población. También hay suficientes UAV listos para sustituir a los que proporcionan conectividad una vez que se queden sin batería, por lo que es seguro suponer que siempre hay un UAV cubriendo cada celda. Cada uno está, como máximo, a 900 m de cada uno de sus vecinos, esto permite una comunicación UAV-a-UAV fuerte y estable bajo la premisa de que se organizan en una topología ad hoc en la cual se emplea tecnología de comunicación basada en radiofrecuencia. A su vez, se maximiza el área de cobertura[8].

Una vez que los UAV se han colocado y actúan como

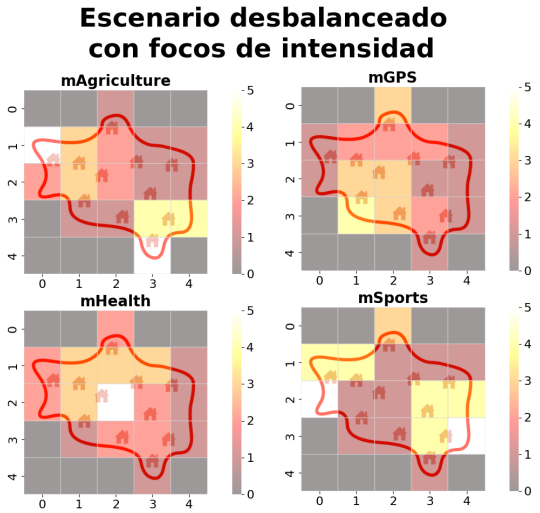


Fig. 2: Peticiones de microservicios categorizadas en 5 niveles. Se puede ver una distribución uniforme con ciertos focos en los que se llevan a cabo más peticiones.

nodos de red que amplían el área de cobertura de la BS, desempeñando a su vez ciertas VNFs, la siguiente tarea es determinar dónde (en qué UAVs) desplegar los microservicios que se van a ofrecer.

Para enriquecer el escenario, supongamos que cada UAV puede alojar, como máximo, un microservicio y que sobre el enjambre deben desplegarse cuatro microservicios: *mAgriculture*, que está relacionado con la agricultura de precisión, *mGPS*, un servicio de geoposicionamiento, *mHealth*, relacionado con Internet of Medical Things (IoMT) y *mSports*, un proveedor de resúmenes de actividad física. Las opciones son, para cada UAV, desplegar uno de los cuatro microservicios o no desplegar nada. Además, se puede suponer que cada microservicio tiene una distribución espacial de peticiones específica en la red en lugar de que se soliciten de forma uniforme. La Figura 2 contiene un mapa de calor para cada microservicio que ilustra su distribución. Cuanto más clara es la celda, mayor es el números de peticiones que realizan los usuarios situados en ella al microservicio.

Con la información sobre los UAVs, la lista de los microservicios y sus mapas de calor, se debe proponer un despliegue óptimo de los servicios en términos de latencia. Para minimizar la latencia extremo a extremo, los esfuerzos se van a dedicar a reducir el número de saltos necesarios para servir los microservicios a los dispositivos finales teniendo en cuenta sus mapas de calor. Se emplea el número de saltos debido a que existe una fuerte correlación, demostrada experimentalmente, entre estos y la latencia de un sistema distribuido [9], [10]. Teniendo esto en cuenta, en la siguiente sección se muestra cómo se formula el problema en un modelo MILP.

III. FORMULACIÓN MATEMÁTICA DEL MODELO

Dada la descripción del escenario de la sección anterior, se propone un modelo MILP que puede resolver cualquier escenario dada la información de entrada necesaria. Los parámetros del modelo son los siguientes elementos:

- El conjunto de UAVs. $D = \{d_1, d_2, ..., d_i\}$, donde cada UAV tiene una capacidad máxima indicada como $capacity(d_n)$.
- El conjunto de los microservicios a desplegar. $M = \{m_1, m_2, ..., m_j\}$, donde cada microservicio tiene un tamaño dado, $size(m_n)$.
- El conjunto formado por el mapa de calor de cada microservicio. $H = \{h_{m_i,d_j} | m_i \in M \wedge d_j \in D\}$.
- El número de saltos (coste de la ruta) entre cada par de UAVs. $R = \{r_{d_i,d_j} | d_i, d_j \in D\}$.

Las variables de decisión del modelo indican si un microservicio, m_i, se despliega en un UAV específico, d_j, o no. Así, estas se pueden interpretar como un conjunto de variables binarias, $X = \{x_{m_i,d_j} | m_i \in M \wedge d_j \in D\}$, tal que:

$$x_{m_i,d_j} = \begin{cases} 1 & \text{si} \quad m_i \text{ se despliega en } d_j \\ 0 & \text{si no} \end{cases}$$

Esta forma de representar las variables de decisión está inspirada en la formulación de Adasme et al. [11], o Khoshkholghi et al. [12].

Como se ha indicado anteriormente, la función objetivo del modelo debe tratar de reducir el número de saltos que hay que realizar para servir a un cualquier usuario cualquier microservicio, $F(X) = \sum_i^D \sum_J^D \sum_K^M x_{m_k,d_j} * h_{m_k,d_i} * r_{d_i,dj}$. Por lo tanto, la formulación del modelo se puede escribir de la siguiente manera:

$$min(\sum_i^D \sum_J^D \sum_K^M x_{m_k,d_j} * h_{m_k,d_i} * r_{d_i,dj}) \tag{1}$$

$$s.t.$$

$$\sum_i^M x_{m_i,d_j}, *size(m_i) \leq capacity(d_j), \forall d_j \in D \tag{2}$$

$$\sum_i^D x_{m_j,d_i} \geq 1, \forall m_i \in M \tag{3}$$

$$x_{m_i,d_j} \leq 1, \forall x_{m_i,d_j} \in X \tag{4}$$

$$x_{m_i,d_j} \geq 0, \forall x_{m_i,d_j} \in X \tag{5}$$

Como puede observarse, existen varios grupos de restricciones. El primer grupo de ellas Eq. (2) especifica que, para cada UAV, la suma de los tamaños de los microservicios desplegados en él no puede superar su capacidad. El

Tabla I
RESUMEN DE LAS ESPECIFICACIONES DEL EQUIPO EN EL QUE SE LLEVAN A CABO LAS PRUEBAS.

Especificaciones del computador	
CPU	Intel Core i7 1270P (12 cores, 16 threads)
Memoria RAM	16 GB DDR4
Sistema Operativo	Linux (Ubuntu 22.04)
Python-MIP (versión)	1.15.0
Gurobi Optimizer (versión)	10.0.1

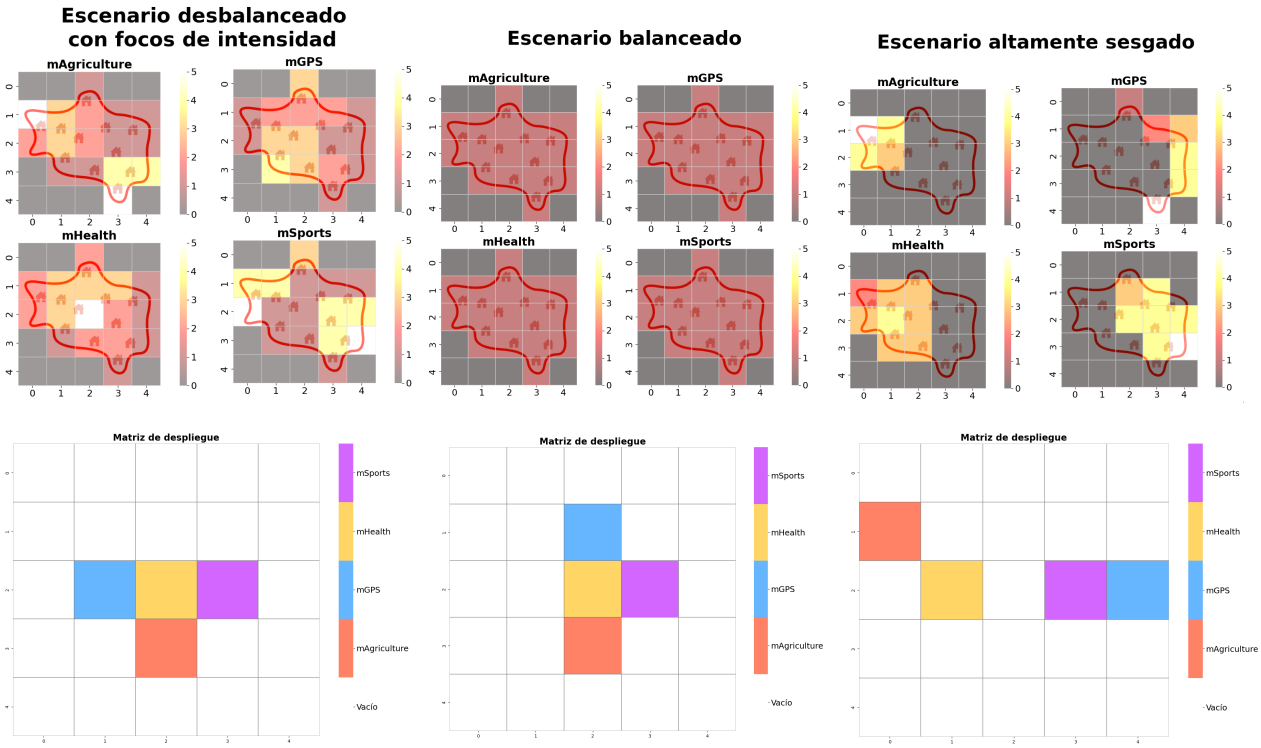


Fig. 3: Mapas de calor de los microservicios junto con las soluciones correspondientes a cada experimento.

segundo grupo Eq. (3) establece que ningún microservicio puede quedarse sin desplegar. Los dos últimos grupos de restricciones Eq. (4, 5) sirven para limitar el rango de las variables de decisión a $\{0, 1\}$. Es decir, que sean binarias.

Este modelo proporciona la solución que minimiza el número total de saltos necesarios para servir los microservicios a los usuarios o dispositivos finales. Esto, a su vez, reducirá la latencia de toda la red debido a la correlación entre el número de saltos y la latencia.

IV. RESULTADOS

En este apartado se discute la solución que ofrece el modelo para el escenario mencionado. Además, se proponen dos casos más sobre la misma población. En el primer subapartado se aborda la generación de soluciones para una entrada dada. Por otro lado, en el segundo, se emulan estas soluciones utilizando Kathará con el objetivo de proporcionar una base de referencia para futuros estudios.

A. Descripción de Soluciones

Para la generación de soluciones se ha implementado, usando el lenguaje Python, el modelo especificado por Eq. (1)-(5) utilizando la librería Python-MIP y el Optimizador Gurobi. Cabe destacar que las especificaciones del equipo informático donde se ha llevado a cabo el proceso de pruebas se exponen en la Tabla I.

Para probar el comportamiento del modelo en diferentes situaciones, se proponen 2 variaciones del escenario. La Fig. 3 muestra tres grupos de mapas de calor junto a las soluciones, uno por cada prueba realizada.

El primer mapa de calor es el mismo de la Fig. 2. Se puede observar que cada microservicio es necesario en cada celda del núcleo de población con mayor o menor intensidad en función de la luminosidad de esta. Es apreciable el hecho de que el modelo tiene una cierta tendencia a desplegar los servicios en el centro de la topología del enjambre. Esto tiene sentido porque el centro es el punto más cercano a todos los demás.

El mapa de calor del centro de la figura muestra una distribución perfectamente equilibrada. Esto significa que cada microservicio es solicitado en cada celda del núcleo de población con la misma frecuencia. Se incluye esta variación porque podría ser una forma de modelar un escenario en el que se desconoce la distribución del consumo de los microservicios que se ofrecen. En este caso, la lucha por la celda central de la red es más obvia, ya que todos los microservicios no tienen preferencia por desplegarse más cerca de un punto focal, que es lo diferencia del primer experimento.

Por último, en el extremo derecho se muestra una distribución muy sesgada. En este tercer experimento, los microservicios no son necesarios en todas las celdas del núcleo de población. Como resultado, el número de conflictos se reduce porque hay menos solapamiento entre los mapas térmicos, lo que, a su vez, permite un despliegue más espaciado. Puede decirse que la localidad de las peticiones de los microservicios es un factor positivo que ayuda a reducir el número total de saltos del escenario.

Interpretar los resultados únicamente visualizando el posicionamiento de los microservicios en los UAVs no proporciona suficiente información sobre el rendimiento del modelo. Para entender realmente si está proporcionando soluciones óptimas, se debe analizar el número de saltos necesarios para servir cada microservicio a los

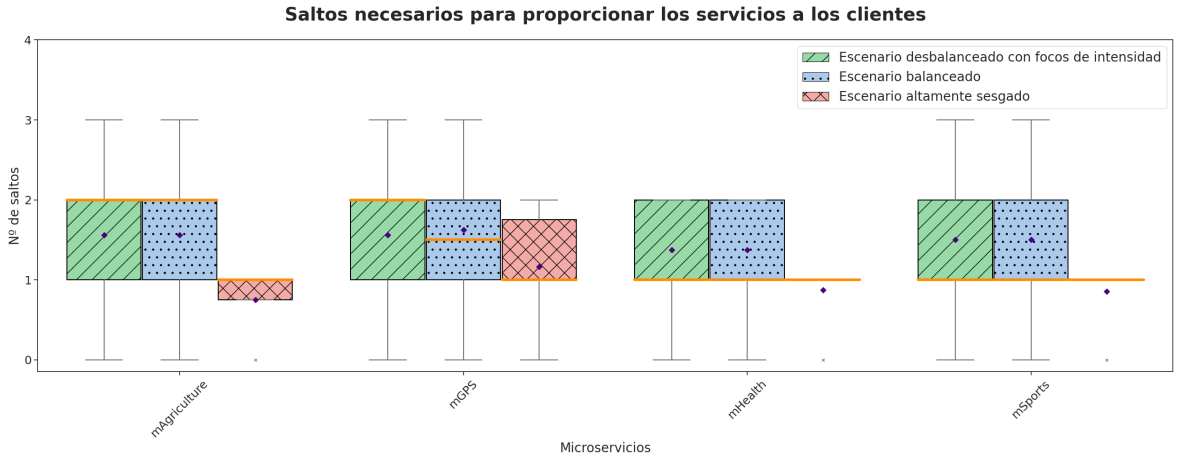


Fig. 4: Visualización del número de saltos que los usuario o dispositivos finales necesitan realizar para ser servidos en cada escenario.

usuarios finales (o dispositivos). Para ello, se muestra la Fig. 4. Mediante *box-plots* se indica la distribución del número de saltos que los usuarios deben realizar para ser servidos cada microservicio.

El peor caso, dada la topología del enjambre de UAVs estudiada, sería de cuatro saltos para servir un microservicio a un dispositivo anfitrión. Sin embargo, esto no ocurre en los escenarios propuestos. Por otro lado, el mejor caso es el de cero saltos para servir un microservicio. Esto significaría que un microservicio dado sólo se requiere en una única celda y este se despliega en ella. Para que esto ocurra, todas las peticiones de un microservicio deben concentrarse en una única celda, lo que no ocurre en los escenarios propuestos.

Se puede concluir que la localidad espacial de las peticiones a los microservicios y el solapamiento en las mismas celdas de peticiones de diferentes microservicios limitan lo buenas que pueden ser las soluciones ofrecidas por el modelo. Ambos problemas expuestos tienen solución. Si se solicita un microservicio sobre una subarea amplia del escenario (poca localidad espacial), la mejor solución es desplegar múltiples instancias del microservicio y asignar a cada usuario o dispositivo final la más cercana, reduciendo así los saltos necesarios para su envío. Por otro lado, si se solicitan dos o más servicios sobre la misma subárea de la malla (solapamiento de microservicios), la única forma de paliarlo es aumentar la capacidad de los UAV que extienden esas posiciones. Esto permitiría el despliegue de una instancia de cada microservicio en la misma celda.

Entre las dos soluciones propuestas para resolver los problemas encontrados, sólo la segunda puede aplicarse en el modelo. Esto se puede hacer aumentando la capacidad del UAV dado. A su vez, la restricción mostrada en Eq. 2 permitirá desplegar más de un microservicio en la misma celda. Sin embargo, la primera implica desplegar más de una instancia de un servicio. El modelo actual no contempla esta posibilidad. Esta situación abre la vía a seguir mejorando el modelo para permitir el despliegue de microservicios instanciados múltiples veces.

B. Emulación de Escenarios

Aunque la Fig. 4 proporciona información útil sobre el número de saltos entre los hosts y los UAV que pueden servir los microservicios que solicitan, la latencia de los escenarios propuestos sigue siendo desconocida.

Con el fin de cuantificar la latencia de la topología de red del enjambre de UAVs se realiza una emulación del escenario. Kathará [4] es el framework bajo el cual se diseña la emulación. Permite emular cualquier topología de red con la ayuda de contenedores Docker.

Se ha diseñado una herramienta para automatizar la traducción de la salida del modelo en las especificaciones que Kathará necesita para crear el entorno. Esta herramienta genera una red compuesta por:

- Una subred por cada par de UAVs adyacentes. Las máscaras de estas redes tienen una longitud fija de 30 bits
- Una subred por cada UAV para conectarlo a los dispositivos clientes. La longitud de las máscaras de las subredes dependen del número de dispositivos asignados a cada UAV, que a su vez, depende del número de microservicios.
- Un *router* por cada UAV en D. Cada *router* tiene $1+k$ interfaces, donde k es el número de UAVs adyacentes a él $k \in [1, 8]$. El protocolo de enrutamiento que emplean es RIPv2.
- n clientes conectados a cada *router* $d_i, \forall d_i \in D$, donde $n = |M|$. Cada *host* puede solicitar un único tipo de microservicio al *router* más cercano que pueda proporcionarlo. Sin embargo, los únicos *hosts* que realmente realizan peticiones son aquellos para los que el valor del mapa de calor del microservicio que solicitan, en la posición del UAV al que están asignados, es mayor que 0. Para decirlo de forma sencilla, un $host_{m_j}$ realiza peticiones si y solo si $h_{m_j,d_i} > 0$. Esto se simula con el uso de una ICMP ECHO *request* que será respondida por el servidor con un ICMP ECHO *reply*. Cabe destacar que todos los *hosts* comienzan a realizar las peticiones simultáneamente y que dejan de hacerlo pasados 60 segundos. La

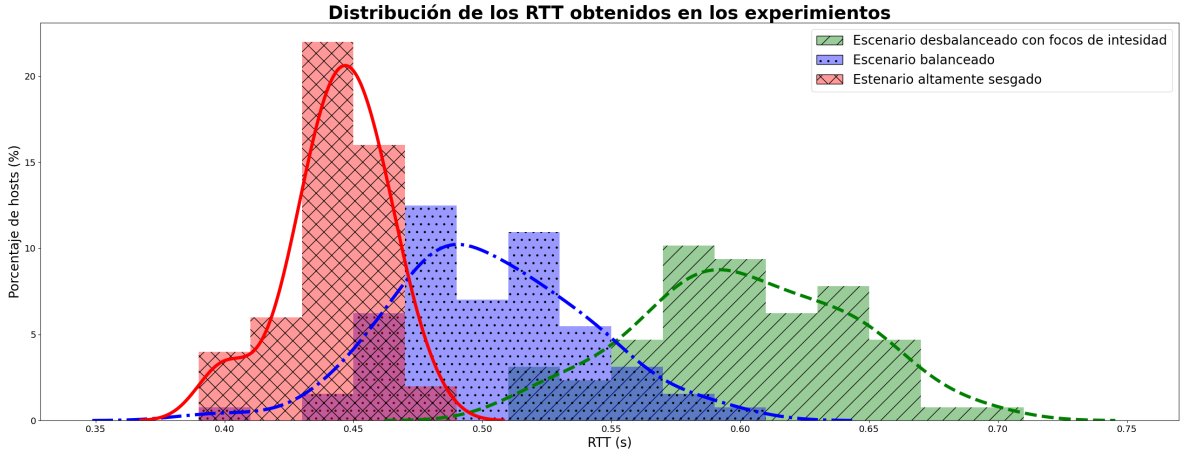


Fig. 5: Representación de la media de los RTT experimentados por los *hosts* en las emulaciones.

frecuencia con la que los *hosts* realizan las peticiones es directamente proporcional al valor del mapa de calor del microservicio que solicitan en la celda en la que están situados.

La forma en que se implementan los laboratorios de Kathará garantiza que, para una topología de red y una lista de microservicios determinadas, el número de dispositivos se mantenga constante independientemente de las variaciones en la distribución del tráfico. Así, para los tres experimentos realizados, el número de dispositivos se mantiene fijo en 16 *routers* y 64 *hosts*. Sin embargo, el volumen de tráfico sí que varía.

La imagen Docker utilizada para los *routers* y los *hosts* de las emulaciones cuenta con la librería de Python Ping3. Esto permite la ejecución, dentro del contenedor de cada host, de un script que hace *ping* al *router* más cercano que "pueda despachar su solicitud de servicio" y almacena el tiempo de ida y vuelta (RTT). Como se ha indicado previamente, tras la convergencia de RIP, cada *host* comienza a lanzar peticiones al *router* que le es asignado. Un *router* es susceptible de ser seleccionado para un *host* si puede despachar el servicio que el *host* solicita y entre todas las posibilidades, se seleciona el más cercano (en función del número de saltos). Cuando el proceso finaliza, se recogen todos los datos.

Cabe destacar el hecho de que, al tratarse de una emulación, los valores obtenidos de RTT dependen, en gran medida, de la capacidad de cómputo del dispositivo sobre el que se ejecutan. Para poder replicar los resultados obtenidos es necesario que las pruebas se ejecuten en un terminal similar al que se describe en Tabla I.

La Fig. 5 contiene un resumen de los experimentos. Muestra que el RTT medio experimentado por los usuarios oscila entre 400 ms y 700 ms. Además, a partir de la distribución del escenario altamente sesgado se puede comprobar que una mayor localidad de las peticiones conduce a una menor varianza, lo que, a su vez, se traduce en una experiencia más homogénea para todos los usuarios implicados. Otro aspecto que corroboran las observaciones es que un mayor número de peticiones conlleva, inequívocamente, una mayor latencia. Para comprobarlo, sólo es necesario observar el rango de cada una de las distribuciones RTT de los experimentos realizados. Este hecho está también respaldado por Fig. 4. En esta se puede observar como el tamaño de las cajas del tercer experimento son más pequeñas para cada microservicio. Se disponen, en orden descendente, según lo poblados que estén los mapas de calor de los microservicios de los experimentos (ver Fig. 3).

V. TRABAJOS RELACIONADOS

Hasta donde saben los autores del presente documento, no se han realizado estudios sobre el despliegue óptimo de SOAs en MEC asistidas por UAV con el objetivo de minimizar la latencia de extremo a extremo. Sin embargo, en [13], se aborda la descarga de tareas intensivas desde dispositivos móviles a servidores *fog* desplegados sobre vehículos con capacidad de computación. Los autores proponen una estrategia de programación basada en una heurística voraz con el objetivo de reducir la latencia. Sin embargo, el enfoque es diferente porque los vehículos implicados en los escenarios son terrestres.

En [14] se aborda la provisión de servicios en redes MEC implementadas mediantes UAVs, pero la principal preocupación no es la latencia. En su lugar, se centran en el despliegue de servicios minimizando el consumo de energía mediante la optimización conjunta de cuatro métricas (colocación de servicios, trayectorias de UAVs, programación de tareas y asignación de recursos). Su resultado, dada la complejidad de la formulación, es un problema de programación no lineal entera mixta no convexa. Nuestro modelo es un problema MILP, que puede resolverse con mucha más facilidad.

En [15] se propone una red MEC de baja latencia asistida por UAV con un *backhaul* de ondas milimétricas (*mmWave*). Sin embargo, se considera un único UAV en lugar de un enjambre. Además, el UAV carece de capacidades de cálculo, actuando sólo como un enlace más rápido entre los usuarios de la red y la red troncal donde se encuentran los recursos. Por otro lado, nuestro modelo considera múltiples UAV con capacidad de cálculo en lugar de uno solo.

Khoshkholghi et al. [12] abordan el problema de la colocación de SFCs en entornos *edge/cloud*. Se propone un modelo de optimización multiobjetivo para minimizar conjuntamente el coste de despliegue y la latencia extremo a extremo. A diferencia de nuestra propuesta, se centran en la infraestructura de red tradicional en lugar de en redes MEC implementadas con flotas de UAVs.

En [16] se aborda el problema del enrutamiento para peticiones de SFC, considerando conjuntamente el posicionamiento de las propias VNF y las restricciones de QoS. El ámbito de estudio son las SDNs habilitadas para el despliegue de NFVs y el objetivo es minimizar el coste de enrutamiento basado en la utilización de la CPU, el ancho de banda de los enlaces, las tablas de flujo en los nodos de conmutación y el despliegue de VNFs en nodos específicos. A pesar de ser un estudio muy relevante, no tiene en consideración el uso de UAVs en la formulación de su problema.

En [17] se toma una aproximación completamente diferente y se busca optimizar la latencia mediante la explotación de las capacidades de la tecnología de acceso multiple no ortogonal al medio de transmisión (NOMA). El trabajo trata sobre la descarga de las tareas de computación de los dispositivos de los usuarios hacia nodos fronterizos con el fin de minimizar el tiempo de computación y el consumo de batería. Sin embargo, no consideran la utilización de UAVs ni se mencionan las SOAs.

Dada la falta de investigación sobre el tema específico, nuestra contribución aborda el despliegue óptimo de SOAs en MECs inhalámbricas implementadas con enjambres de UAVs con recursos limitados.

VI. CONCLUSIÓN Y TRABAJO FUTURO

Este artículo introduce y aborda el problema de desplegar los componentes de SOAs en MECs aumentada mediante emjambres de UAVs con recursos limitados con el objetivo de minimizar la latencia. Para ejemplificar el problema se propone un escenario IoT con restricciones de QoS que tiene lugar en una zona rural. Se propone, implementa y evalúa un modelo MILP que encuentra el despliegue óptimo de microservicios dada una topología de enjambre de UAVs con unas características de tráfico específicas. Además, se emulan mediante Kathará tres soluciones a diferentes escenarios proporcionadas por el modelo. Los resultados del modelo son prometedores, ya que permiten realizar despliegues óptimos en los escenarios estudiados. Sin embargo, cuando las peticiones de servicios están muy dispersas por la red, tener un único despachador posible para ellas empeora la latencia. Además, el modelo actual no tiene en cuenta el concepto de flujo de servicio, es decir, una secuencia de microservicios que con frecuencia se solicitan conjuntamente (*workflow* o *service function chain*). En futuras investigaciones, está previsto formular nuevos modelos que tengan en cuenta los aspectos mencionadas.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por MCIN/AEI/10. 13039/501100011033 y por la Unión Europea "Next GenerationEU /PRTR", por el Ministerio de Ciencia, Innovación y Universidades (proyectos TED2021-130913B-I00, PDC2022-133465-I00), por el proyecto PID2021-124054OB-C31 y la subvención CAS21/00057 (MCI/AEI/FEDER, UE), y por la Consejería de Economía, Ciencia y Agenda Digital de la Junta de Extremadura (GR21133).

REFERENCIAS

[1] A. Fakhreddine, C. Raffelsberger, M. Sende, and C. Bettstetter, "Experiments on drone-to-drone communication with wi-fi, lte-a, and 5g," *2022 IEEE GLOBECOM Workshops, GC Wkshps 2022 - Proceedings*, pp. 904–909, 2022.

[2] J. L. Herrera, J. Galan-Jimenez, J. Berrocal, and J. M. Murillo, "Optimizing the response time in sdn-fog environments for time-strict iot applications," *IEEE Internet of Things Journal*, vol. 8, pp. 17 172–17 185, 12 2021.

[3] B. Choudhury, S. Choudhury, and A. Dutta, "A proactive context-aware service replication scheme for adhoc iot scenarios," *IEEE Transactions on Network and Service Management*, vol. 16, no. 4, pp. 1797–1811, 2019.

[4] 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.

[5] C. Tang, C. Zhu, and M. Guizani, "Coverage optimization based on airborne fog computing for internet of medical things," *IEEE Systems Journal*, 2023.

[6] M. Mahbub, "Uav assisted 5g het-net: A highly supportive technology for 5g nr network enhancement," *EAI Endorsed Transactions on Internet of Things*, vol. 6, pp. 1–19, 08 2020.

[7] J. L. Herrera, J. Galán-Jiménez, J. Berrocal, and J. M. Murillo, "Optimizing the response time in sdn-fog environments for time-strict iot applications," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 17 172–17 185, 2021.

[8] 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.

[9] F. Yakubu, S. M, M. Bagiwa, and H. Murtala, "Correlation between latency and hop count," 12 2009.

[10] X. Chen, K. Nguyen, and H. Sekiya, "An experimental study on performance of private blockchain in iot applications," *Peer-to-Peer Networking and Applications*, vol. 14, pp. 3075–3091, 9 2021.

[11] P. Adasme, A. Viveros, A. D. Firoozabadi, and I. Soto, "Mathematical models for minimizing latency in software-defined networks."

[12] M. A. Khoshkholghi, M. G. Khan, K. A. Noghani, J. Taheri, D. Bhamare, A. Kassler, Z. Xiang, S. Deng, and X. Yang, "Service function chain placement for joint cost and latency optimization," *Mobile Networks and Applications*, vol. 25, pp. 2191–2205, 12 2020.

[13] C. Tang, X. Wei, C. Zhu, Y. Wang, and W. Jia, "Mobile vehicles as fog nodes for latency optimization in smart cities," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 9364–9375, 9 2020.

[14] Y. Qu, H. Dai, H. Wang, C. Dong, F. Wu, S. Guo, and Q. Wu, "Service provisioning for uav-enabled mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3287–3305, 2021.

[15] Y. Yu, X. Bu, K. Yang, H. Yang, and Z. Han, "Uav-aided low latency mobile edge computing with mmwave backhaul," *IEEE International Conference on Communications*, vol. 2019-May, 5 2019.

[16] L. Liu, S. Guo, G. Liu, and Y. Yang, "Joint dynamical vnf placement and sfc routing in nfv-enabled sdns," *IEEE Transactions on Network and Service Management*, vol. 18, p. 4263, 2021.

[17] L. Qian, Y. Wu, J. Ouyang, Z. Shi, B. Lin, and W. Jia, "Latency optimization for cellular assisted mobile edge computing via non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5494–5507, 2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Metodología para el análisis del comportamiento de protocolos de transporte sobre canales variantes: Aplicación en escenarios NTN

Fátima Khan*, Fátima Fernández†, Luis Diez*, Ramón Agüero*
*Departamento de Ingeniería de Comunicaciones (DICOM), Universidad Cantabria
fatima.khan@unican.es,{ldiez,ramon}@tlmat.unican.es
†IoT and Digital Platforms Department, Ikerlan Technology Research Center
ffernandez@ikerlan.es

La proliferación de nuevas tecnologías inalámbricas, así como de las correspondientes topologías de acceso, hace necesaria la evaluación del rendimiento de protocolos, tanto tradicionales como los que están apareciendo de manera más reciente, sobre dichos escenarios. Aunque la comunidad científica suele utilizar emuladores de enlaces con este fin, estos no están concebidos para reflejar, de manera precisa, la gran variabilidad temporal de algunos canales inalámbricos. Por otro lado, los simuladores de red proporcionan una implementación detallada de las capas inferiores, que puede abstraerse para analizar el rendimiento de los protocolos superiores. Como solución de compromiso, en este artículo se propone un enfoque de simulación ligero, utilizando el simulador ns-3, que utiliza modelos dinámicos de capacidad de canal procedentes de la literatura. El enfoque propuesto se ha utilizado para modelar topologías de acceso basadas en LEO, sobre los que se evalúa el comportamiento del protocolo TCP bajo diferentes configuraciones.

***Palabras Clave*—ns-3, Markov Chain, modeling, RAN, LEO**

I. INTRODUCCIÓN

El continuo desarrollo y transformación de las tecnologías inalámbricas y, en particular, de las redes de acceso radio (RAN), hace necesario reexaminar el rendimiento de los protocolos que operan sobre ellas. Así, en los últimos años muchos trabajos han analizado el rendimiento de TCP sobre canales mmWave [1], y se han desarrollado además herramientas para este propósito [2]. Por otro lado, las nuevas generaciones de comunicaciones móviles promueven, como ya se ha adelantado, innovadoras arquitecturas de redes de acceso que se integran con las redes no terrestres (NTN), incluyendo aquellas basadas en drones y satélites de baja órbita, Low Earth Orbit (LEO).

En estas circunstancias, cuando los despliegues reales no son una alternativa viable, la evaluación de las nuevas tecnologías de acceso suele realizarse mediante simulación. Los modelos empleados suelen centrarse en las capas inferiores (MAC y PHY), que pueden llegar a tener una granularidad temporal de milisegundos o microsegundos. En cambio, los protocolos de capas superiores operan en una escala temporal sensiblemente mayor. En este sentido, estos protocolos de transporte suelen analizarse utilizando herramientas que normalmente no incluyen mucho detalle de capas inferiores, como los emuladores de enlace. Por ejemplo, Pantheon, un marco para el análisis de mecanismos y algoritmos de control de congestión [2], incorpora el emulador de enlace MahiMahi [3]. Por otro lado, se ha demostrado que el comportamiento de los enlaces inalámbricos puede modelarse de manera suficientemente precisa a través de cadenas de Markov, asumiendo diferentes estados de transmisión [4], [5], [6].

Teniendo en cuenta todo lo anterior, en este trabajo describimos una metodología de análisis, que aprovecha el simulador ns-3 para analizar el rendimiento de protocolos de capas altas sobre enlaces inalámbricos altamente dinámicos. La solución propuesta permite evaluar tanto enlaces simples como topologías más complejas, abarcando múltiples enlaces con diferentes características. El objetivo de este trabajo es por tanto diseñar y desarrollar una metodología de simulación eficiente y flexible, para evaluar el rendimiento de protocolos de capas superiores sobre entornos inalámbricos. Para lograr esto, se hace uso de ns-3, un popular simulador de red de código abierto, como base para la herramienta propuesta. El enfoque que se presenta en el trabajo se basa en la creación de escenarios de simulación que representan diferentes configuraciones de conectividad inalámbrica. Éstas pueden ser sencillas, como una conexión punto a punto, o más complejas, involucrando múltiples nodos y enlaces interconectados. La versión modificada de ns-3 con los modelos propuestos, así como las utilidades de automatización que se han

desarrollado se encuentran disponibles públicamente en el repositorio Github [1].

La herramienta que se ha desarrollado permite ajustar diferentes características de los enlaces, como la tasa de transmisión, el retardo y el modelo de pérdida de paquetes, para emular diferentes condiciones de red realistas. Aunque la metodología sugerida no está vinculada a una tecnología inalámbrica específica, se utilizará una red de acceso basada en satélites LEO como ejemplo ilustrativo para validar la propuesta, ya que requiere modelar tanto los propios enlaces inalámbricos como la topología RAN global. Esta metodología de análisis ha sido adoptada por Hervella *et al.* para comparar el rendimiento de TCP y QUIC [7]. Sin embargo, este trabajo se centra en describir la metodología de simulación sobre ns-3, más que en una evaluación concreta, al tiempo que se amplían las posibilidades que se incluyen.

En conjunto, el enfoque propuesto permite una simulación simplificada (de manera que se pueda adaptar a la dinámica temporal de las capas superiores), pero a la vez precisa, de las tecnologías inalámbricas, manteniendo una implementación detallada para las capas superiores. Una de las ventajas de este enfoque es que aprovecha la limitada complejidad de la simulación, lo que permite una interacción con aplicaciones en tiempo real utilizando la herramienta ns-3 TAP. Esto es especialmente útil cuando se desea evaluar cómo las aplicaciones interactúan con los protocolos de capa de transporte en entornos inalámbricos, lo que puede proporcionar información valiosa sobre el desempeño y la eficiencia de estas aplicaciones en escenarios reales. Sin embargo, es importante tener en cuenta que, a pesar de que la metodología de simulación propuesta logra reflejar apropiadamente el comportamiento de las tecnologías inalámbricas subyacentes, los modelos teóricos utilizados pueden ser menos precisos en comparación con una simulación de mayor detalle. La simulación con gran nivel de precisión implicaría por su parte tener en cuenta cada detalle y cada evento en el sistema, logrando reflejar de manera más exacta el comportamiento de los protocolos en situaciones específicas. Sin embargo, para diferentes aplicaciones y servicios, el modelado detallado que habilita una simulación *de grano fino* pueden no tener un impacto relevante.

Las contribuciones de este trabajo pueden resumirse como sigue:

1) Descripción de un enfoque de simulación ligero, que permite analizar el comportamiento de protocolos de capas altas (transporte/aplicación) sobre redes con conectividad inalámbrica. Se detalla la construcción de los escenarios de simulación, indicando qué características se pueden ajustar y qué protocolos de capa alta se utilizan para generar tráfico. Esta descripción proporciona una base sólida para futuras investigaciones y desarrollo de soluciones en el ámbito de las redes inalámbricas.
2) Evaluación de la viabilidad de la metodología propuesta, a partir de un escenario relevante que representa condiciones de red realistas, se han realizado una exhaustiva campaña de experimentos para evaluar el rendimiento de protocolos de capas alta sobre diferentes configuraciones de enlaces inalámbricos. Los resultados que se presentan demuestran la utilidad y aplicabilidad de la metodología propuesta para comprender y mejorar el rendimiento de los protocolos sobre entornos inalámbricos, incluso altamente dinámicos.
3) Disponibilidad del código en un repositorio abierto, para fomentar la colaboración y permitir que otros investigadores puedan utilizar la metodología propuesta. Esto permite a la comunidad científica acceder, revisar, utilizar, y contribuir al desarrollo de la solución de simulación ligera en ns-3 descrita en el artículo. Al hacerlo, se promueve asimismo la reproducibilidad de los resultados mostrados, así como el avance en el campo de las redes inalámbricas.

El resto del documento se estructura como sigue. En la Sección II se revisan las alternativas para evaluar protocolos de capas altas sobre diferentes redes con conectividad inalámbrica. A continuación, en la Sección III se describe la implementación de la metodología propuesta, que posteriormente se evalúa en la Sección IV, analizando el comportamiento del protocolo TCP sobre una red de acceso basada en LEO. Finalmente, en la Sección V se resumen las principales contribuciones del artículo, identificando algunos aspectos que serán abordados en el futuro.

II. ESTADO DEL ARTE

Existe una gran variedad de herramientas y entornos de simulación y emulación, con diferentes funcionalidades y características. Por un lado, los emuladores en tiempo real como Dummynet [8] o NetEm [9] permiten limitar la capacidad de transmisión, y añadir colas con diferentes políticas de gestión. Además, permiten emular diferentes retardos y eventos de pérdidas de tramas/paquetes.

Del mismo modo, las herramientas de emulación de redes como Mininet utilizan configuraciones similares para los enlaces que conectan los nodos de la red [10]. Aunque estas herramientas son capaces de proporcionar un gran grado de realismo, no están pensadas para reflejar las características de canales dinámicos, por ejemplo con capacidad variable, como es el caso de los enlaces inalámbricos.

También merece la pena mencionar el emulador de enlaces MahiMahi [3]. Además de las funcionalidades de emulación proporcionadas por las herramientas anteriores (capacidad de transmisión, retardo del enlace, colas), Mahimahi puede configurarse con trazas de capacidad registradas en configuraciones reales. Aunque esta característica se puede utilizar para emular canales dinámicos, no permite utilizar modelos teóricos, con lo que los resultados están limitados a las condiciones particulares de los escenarios en los que se obtuvieron las trazas.

En cuanto a los simuladores de red, como Omnet++ [11] o ns-3 [12], proporcionan mucha más flexibilidad en el

[1] https://github.com/tlmat-unican/Lightweight-ns-3-link-simulation

modelado de los mecanismos de comunicación subyacentes, por lo que se podrían incluir modelos teóricos. Sin embargo, se centran en replicar con precisión el comportamiento de entornos reales, lo que conduce a implementaciones con un elevado grado de complejidad en las capas inferiores necesarias, lo que dificulta la posibilidad de realizar evaluaciones sobre escenarios complicados, que contemplen un número de dispositivos elevado.

El enfoque propuesto busca mantener la complejidad de ns-3 en las capas superiores, al mismo tiempo que simplifica el modelado de las capas inferiores mediante la utilización de enlaces dinámicos, que logran un grado de realismo apropiado, aun manteniendo una complejidad limitada.

III. SIMULACIÓN A NIVEL DE ENLACE

En esta sección se detallan las modificaciones realizadas en ns-3 y el desarrollo adicional necesario para implementar la metodología propuesta, y el marco de simulación ligero en la que se basa. En resumen, se utiliza el objeto enlace punto a punto para emular canales individuales con modelos teóricos/empíricos, y se crean además utilidades para desplegar escenarios más complejos, a partir de archivos de configuración, con el objetivo de facilitar la realización sistemática y repetitiva de experimentos.

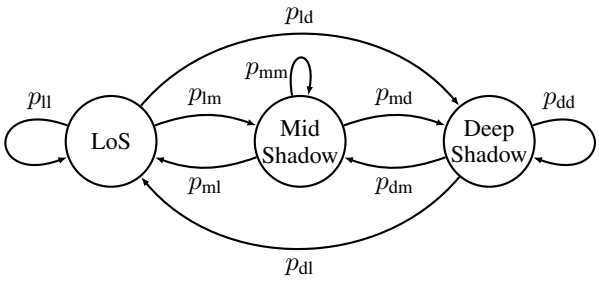
A. Adaptación enlace punto a punto

La implementación llevada a cabo se aprovecha principalmente en la funcionalidad proporcionada por el dispositivo punto a punto, implementado por la clase `PointToPointNetDevice` de ns-3. Está compuesto por un un *buffer* que recibe tráfico y, a su vez, una interfaz caracterizada por una determinada velocidad de transmisión y probabilidad de error.

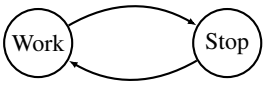
Con esas características de base, el simulador utiliza la capacidad, y la longitud de los paquetes, para programar los eventos correspondientes a la transmisión de paquetes a través de la interfaz. Además, la velocidad de transmisión se fija habitualmente al principio de la simulación, permaneciendo inalterada durante todo el experimento.

Por otro lado, la clase `PointToPointNetDevice` tiene un método público (`SetDataRate`) que permite cambiar la tasa durante la simulación. Al hacer uso del mismo se observaron algunos problemas, principalmente en casos extremos. En concreto, cuando la capacidad se establece en 0 (por ejemplo, si se pretende emular una desconexión temporal), el dispositivo no era capaz de reiniciar la transmisión de los paquetes en cola, una vez se volviera a incrementar la velocidad. Para superar esta limitación, cada vez que se produce un cambio de velocidad, y el valor anterior era 0, se comprueba la ocupación del *buffer*. Si hubiera algún paquete en espera, se iniciaría su transmisión (se llama a la función `TransmitStart`), de forma que el dispositivo recupera su funcionamiento normal.

En cualquier caso, cabe destacar que el cambio de velocidad de transmisión no modifica las transmisiones activas, por lo que los eventos de transmisión de paquetes no se reprogramaran ante una modificación de la capacidad de la interfaz.



(a) Cadena de Markov de tres estados para el canal LMS



(b) Cadena de Markov de dos estados para emular desconexiones temporales

Fig. 1: Cadenas de Markov integradas

B. Integración de los modelos de conectividad

Para reflejar el comportamiento del canal subyacente se plantea la integración de modelos basados en Cadenas de Markov, habituales para emular el comportamiento de enlaces inalámbricos. Como se ha dicho anteriormente, se utilizará un escenario basado en comunicaciones LEO para validar la metodología propuesta. Se utilizarán redes que incorporan dos tipos de enlaces: (1) entre la estación terrena y el primer satélite (o entre el último satélite y la estación terrena en recepción); y (2) enlaces entre satélites. Para el primer tipo de enlace (canal LMS), se plantea utilizar la un modelo (ver Figura 1a) basado en una cadena de Markov con 3 estados [4]. Así, se considera que los enlaces pueden pasar por diferentes condiciones, debidas a posibles desvanecimientos y obstáculos, que a su vez repercutirían en la velocidad de transmisión. Por lo tanto, teniendo en cuenta estas características ambientales, el canal transita entre los siguientes estados: line of sight (LoS) con condiciones ideales; mid-shadowing (MS), en el que las condiciones de conectividad empeoran; y deep-shadowing (DS), en el que la capacidad de transmisión se ve gravemente afectada. En cada estado, la capacidad de transmisión se considera constante.

Por su parte, para los enlaces entre satélites, se plantea la utilización de cadenas de Markov con 2 estados que modelan dos posibles situaciones: capacidad constante e interrupciones [13], como puede verse en la Fig. 1b. En este caso, la capacidad en un estado de interrupción (*off*) es cero.

Para cada enlace del escenario, se implementa una cadena de Markov independiente, que gobierna los cambios de la tasa de transmisión en una conexión punto a punto, reflejando las situaciones descritas anteriormente. De esta forma, cuando un enlace cambia a un nuevo estado, el

simulador llama a la función `SetDataRate` del dispositivo punto a punto correspondiente, y programa un evento para cambiar nuevamente de estado según la distribución de tiempos de permanencia. Este proceso se repite hasta el final de la simulación, de manera independiente para todos los enlaces, y sin tener en cuenta si están en uso actualmente, a diferencia de los emuladores de canal basados en trazas.

C. *Definición del escenario*

En esta sección se muestran el conjunto de utilidades que se han desarrollado para crear una topología personalizada basada en un escenario ns-3 genérico. El escenario se alimenta con un archivo de configuración JSON que describe completamente la topología de la red. En concreto, define el número de enlaces del escenario, así como su tipo y características. Hasta ahora, se han implementado enlaces que adoptan los modelos de las cadenas de Markov de 3 y 2 estados antes mencionados, así como enlaces ideales con configuraciones estáticas. En el caso de aquellos basados en cadenas de Markov, se definen a su vez varios subtipos, que representan configuraciones para diferentes bandas de frecuencia, utilizando datos de la bibliografía [4].

Por simplicidad, las configuraciones detalladas de los modelos (matrices de transición, tiempos medios de permanencia en los estados, entre otros) se implementan en código C++, y el fichero de configuración sólo indica el modelo que hay que cargar para cada enlace. En el futuro, se pretende generalizar la implementación, de modo que las propias cadenas de estados puedan definirse también mediante ficheros de configuración.

En todos los casos, el JSON también establece el valor de los parámetros heredados del enlace punto a punto, y de los dispositivos que no dependen del modelo adoptado: tasa de error, retardo, MTU y tamaño del *buffer*. Además, la implementación permite inyectar tráfico de fondo TCP en cada enlace de forma independiente. Este flujo de tráfico se genera mediante una aplicación `ON-OFF`, y el fichero de configuración define sus principales atributos: tasa de tráfico, tamaño del paquete, número de bytes, y tiempos en los estados ON y OFF.

Por último, esta configuración proporciona la versatilidad necesaria para instanciar tanto un cliente como un servidor con el propósito de generar el tráfico que será objeto de análisis. Este enfoque es aplicable tanto en el entorno nativo de ns-3 como mediante la aplicación de técnicas de virtualización. La virtualización implica la creación de un enlace transparente de tipo `csma` caracterizado por una capacidad infinita y la ausencia de retardo. Este flujo atraviesa la topología definida y se generan registros que permiten monitorizar la transmisión y recepción de paquetes, la ventana de congestión, la ocupación de los *buffer* de todos los enlaces, así como la variación de la capacidad del canal.

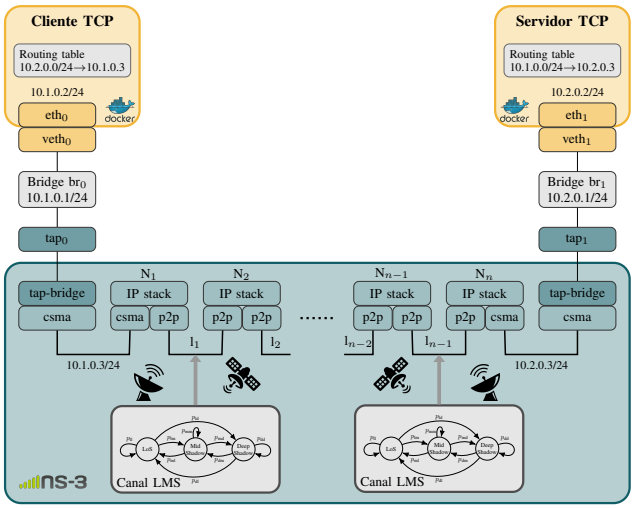


Fig. 2: Diagrama del banco de pruebas para llevar a cabo la evaluación que integra NS-3, contenedores Docker y el modelo LMS.

IV. VALIDACIÓN Y CASO DE USO DE LA APLICACIÓN

En esta sección, se valida la metodología propuesta, así como el correcto funcionamiento de las utilidades implementadas sobre un escenario de comunicaciones LEO, y que se muestra en la Figura 2. Como puede observarse, el escenario emula dos enlaces de satélite móvil terrestre (LMS), que conectan estaciones terrestres y satélites. Además, se pueden añadir varios enlaces intersatélite (ISL), con diferentes características. En los nodos que emulan las estaciones terrenas, se envía tráfico TCP, utilizando CUBIC como algoritmo de control de congestión [14], con una tasa constante que iguala la tasa media de los enlaces de acceso.

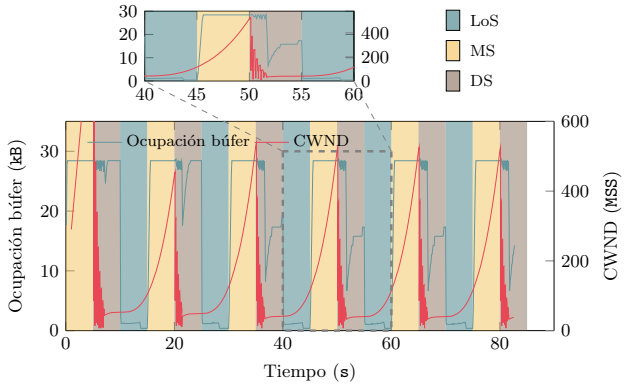
En primer lugar, se muestran los resultados obtenidos sobre un único enlace sencillo, controlado para validar el correcto comportamiento de la metodología propuesta. Posteriormente, se amplía la evaluación con una configuración más realista, abarcando tanto canales LMS como enlaces ISL, con tráfico de fondo y desconexiones. La Tabla I resume los parámetros de configuración utilizados en los experimentos llevados a cabo. Como se puede apreciar la velocidad de transmisión máxima, corresponde a la situación de línea de visión directa, cuyo valor es de 80 `Mbps` [4]. Además, las capacidades en situaciones de mid-shadowing y deep-shadowing se ajustan al 50% y al 20%, respectivamente, de esa capacidad máxima.

A. *Modelo sintético*

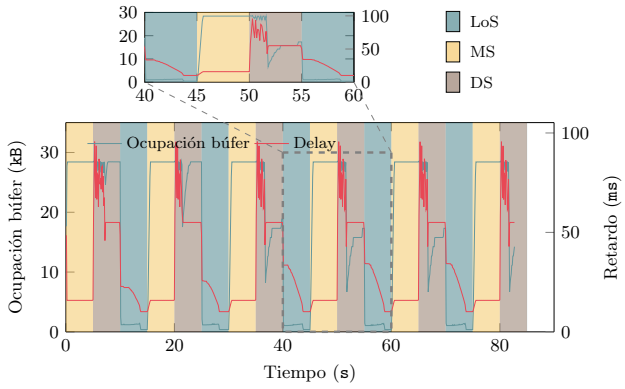
Este escenario contempla un único enlace modelado como una cadena de Markov de 3 estados y una capacidad de transmisión en cada estado definida en la Tabla I. La configuración imita el canal LMS definido en [7], pero con tiempos de permanencia constantes de 5 segundos, para hacerlo más predecible y poder analizar mejor el comportamiento de TCP y su mecanismo de control de congestión. Los valores de capacidad de transmisión se

Tabla I: Configuración de los escenarios

Enlace LMS	
Modelo	Cadena de 3 estados Sintético y Realista [7]
Capacidad media del enlace	$\approx$ 45 Mbps
Capacidad de los estados	[80, 40, 16] Mbps
Capacidad de la cola	0.5· BDP
Retardo base	10 ms
Enlace ISL	
Modelo	Cadena de 2 estados Tasa constante
Capacidad de los estados	2 estados: [0, 80] Mbps Tasa constante: 50 Mbps
Tasa del tráfico de fondo	[5, 20] Mbps
Tiempo medio de interrupción	[0, 3] s
Capacidad de la cola	Sin límite (∞)
Retardo base	10 ms
Aplicación	
Tasa de los datos de aplicación	40 Mbps
Tamaño del fichero	300 MB



(a) Búfer vs. Ventana de congestión

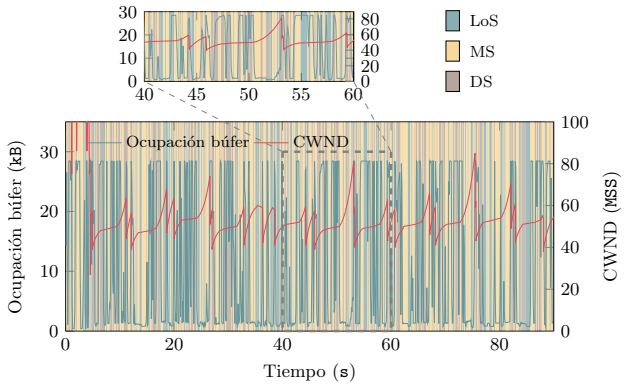


(b) Búfer vs. Retardo

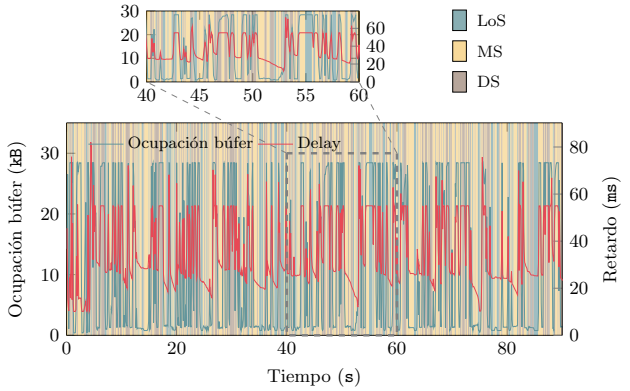
Fig. 3: Creación de un único canal de 3 estados con un tiempo constante de permanencia de 5 segundos.

definen para emular las condiciones de los canales LoS, MS y DS, tal y como se menciona en la Sección A.

La Figura 3 muestra la evolución temporal de la conexión, utilizando diferentes colores de fondo para cada estado. Primero, la Figura 3a muestra la evolución de la ocupación del *buffer* (eje izquierdo) y de la ventana de congestión (cwnd, eje derecho). Como puede observarse,



(a) Búfer vs. Ventana de congestión



(b) Búfer vs. Retardo

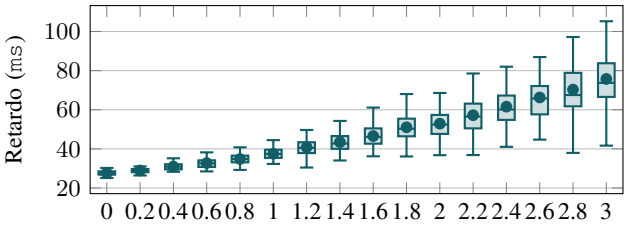
Fig. 4: Realización de un único canal LMS como se define en [7]

el *buffer* se satura durante los periodos MS y DS, y se vacía durante las condiciones LoS. En cuanto a la ventana de congestión, crece de forma constante durante los estados MS y LoS, y se reduce de manera brusca al principio de las fases DS, cuando se producen pérdidas por saturación del *buffer*, el cual se ajustó previamente al 50% del producto entre el ancho de banda y el retardo medio de propagación (BDP). Este ajuste se realiza con el propósito de estudiar de manera detallada este fenómeno.

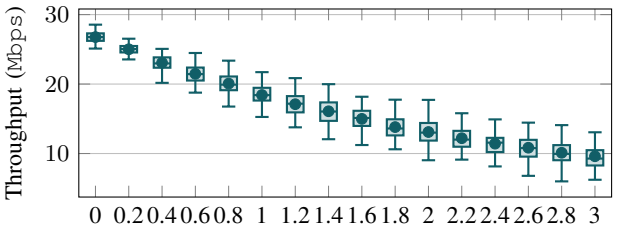
La Figura 3b ilustra, para la misma simulación, la evolución del *buffer* y del retardo, definido como la diferencia entre la transmisión de un paquete y su correcta recepción en el destino. En este caso, se puede observar que el retardo aumenta notablemente en la transición de los estados MS a DS. Además, se ve que, cuando el canal está en la situación de DS, se reduce bruscamente, como consecuencia del agotamiento de la ventana de congestión. Finalmente, alcanza su valor mínimo durante las condiciones LoS. Este comportamiento es coherente con el funcionamiento de TCP Cubic, aunque se obtendrían patrones diferentes para distintos tamaños de *buffer*.

B. Análisis basado en comunicaciones LEO

Se amplían ahora los resultados, utilizando un modelo de canal realista para el enlace LMS. El modelo es de nuevo una cadena de Markov de 3 estados, con tiempos de permanencia distribuidos exponencialmente, y transiciones



(a) Distribución del retardo medio.



(b) Distribución del rendimiento medio.

Fig. 5: Impacto de las desconexiones en el enlace ISL para las comunicaciones extremo a extremo. Los resultados se obtienen a partir de 100 simulaciones para cada valor de tiempo medio de desconexión.

con probabilidades definidas en [7]. En primer lugar, se analiza el comportamiento de un único enlace LMS, para abordar posteriormente las comunicaciones extremo a extremo, abarcando dos enlaces LMS, ascendente y descendente, y enlaces entre satélites con tráfico de fondo e interrupciones.

La Figura 4 muestra los resultados que presentaron para el canal LMS 'sintético', en este caso para una configuración realista. De nuevo, los colores de fondo indican el estado del canal. Como puede observarse, existe una variación muy rápida de las condiciones del canal entre cualquier par de estados, lo que lo vuelve muy impredecible.

En la Figura 4a se muestra la evolución del *buffer* y la ventana de congestión a lo largo del tiempo. En este caso, la ventana de congestión apenas aumenta, debido a las continuas variaciones de capacidad, que dificultan la adaptación de la tasa de transmisión. Esto se evidencia en el detalle que se hace de un intervalo de la conexión, donde se observa que el *buffer* está vacío incluso en situación DS, debido al pequeño tamaño de la ventana de congestión. En este sentido, la Figura 4 muestra que el retardo del tráfico sigue la tendencia de ocupación del *buffer*. Dado que el control de congestión no es capaz de aprovechar toda la capacidad del canal, se envía menos tráfico, y el retardo es relativamente inferior al observado en la Figura 3b para el canal LMS sintético. De nuevo, los resultados son coherentes con el funcionamiento del mecanismo de control de congestión empleado por TCP, Cubic.

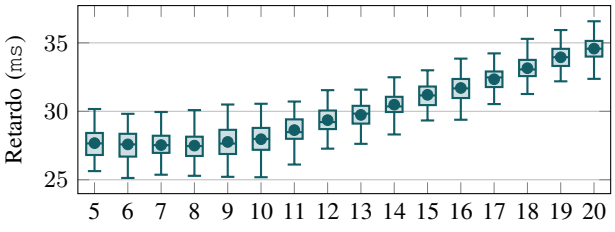
A continuación, se analiza el comportamiento extremo a extremo en un escenario con dos enlaces LMS (ascendente y descendente) y un ISL que los conecta. El enlace ISL se modela como una cadena de Markov de 2 estados con tiempos de permanencia distribuidos exponencialmente y definidos como sigue (ver Tabla I): periodo activo con una tasa de transmisión de 80 `Mbps` (mayor que la capacidad media de los enlaces LMS realistas) y un tiempo de permanencia medio de 5 `s`; y estado inactivo, que refleja las situaciones de interrupción del enlace, y cuyo tiempo de permanencia medio aumenta de 0 a 3 segundos. Para cada valor del tiempo medio de interrupción, se ejecutan 100 simulaciones independientes, en cada una de las cuales se envía un fichero de 300 `MB`. Cabe señalar que el experimento se detiene cuando la aplicación recibe todos los bytes.

La Figura 5 muestra la distribución del retardo medio extremo a extremo y el rendimiento observado, a medida que se va incrementando el tiempo medio de interrupción en el enlace ISL. Como se podía prever, en la Figura 5a se puede observar que el retardo medio extremo a extremo aumenta de forma constante, y también lo hace su dispersión. A su vez, la Figura 5b muestra que el rendimiento disminuye como consecuencia tanto de la reducción de la capacidad debida a los tiempos de interrupción, como de la reacción del control de congestión ante dichas desconexiones temporales.

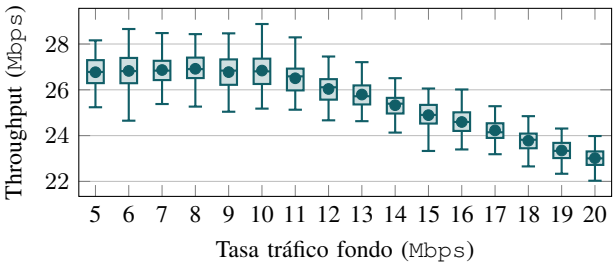
Por último, se ha configurado el enlace ISL con tráfico de fondo (TCP). Para simplificar la evaluación de los resultados obtenidos, en esta configuración no se tienen en cuenta las interrupciones. Por lo tanto, el enlace ISL se configura con una velocidad constante (funcionamiento normal del enlace punto a punto) de 50 `Mbps`. Se realiza un barrido del tráfico de fondo en el enlace ISL de 5 a 20 `Mbps`, realizando 100 ejecuciones independientes para cada configuración. De nuevo, cada ejecución conlleva la transmisión de 300 `MB` y sólo se detiene cuando se reciben todos los bytes. La Figura 6 muestra las distribuciones del retardo medio de extremo a extremo y el rendimiento a medida que se incrementa el tráfico de fondo.

En la Figura 6a se observa que el retardo extremo a extremo se mantiene bastante estable hasta que la tasa de tráfico de fondo alcanza los 10 `Mbps`. A partir de ese punto, el retardo aumenta constantemente, mostrando una variabilidad que no cambia de manera apreciable. Mientras tanto, en la Figura 6b se muestra un comportamiento diferente para el rendimiento, que se mantiene bastante estable, en torno a los 27 `Mbps`, hasta que la tasa de tráfico de fondo alcanza los 10 `Mbps`, disminuyendo a partir de ese punto. Cabría esperar que el caudal del flujo extremo a extremo con bajo tráfico de fondo, alcanzara un valor más cercano al promedio de los enlaces LMS (40 `Mbps`). Sin embargo, es importante señalar que la falta de adaptación de la ventana de congestión tiene un gran impacto en el rendimiento, como se observó en el análisis mencionado anteriormente, y esto es aún más acusado en presencia de dos enlaces LMS (enlace ascendente y enlace descendente).

Como se ha podido ver, la metodología propuesta permite acometer la evaluación de protocolos de capas



(a) Distribución del retardo medio.



(b) Distribución del rendimiento medio.

Fig. 6: Impacto del tráfico de fondo en el enlace ISL sobre las comunicaciones de extremo a extremo. Los resultados se obtienen a partir de 100 simulaciones para cada valor de tráfico de fondo.

altas sobre enlaces inalámbricos únicos y topologías más complejas, sin incurrir en un modelado muy detallado, que implicaría unos tiempos de simulación mucho más elevados. Así, aunque estas evaluaciones pueden realizarse utilizando modelos de simulación más precisos de las capas bajas (PHY, MAC) la complejidad subyacente aumentaría sustancialmente, y en diferentes circunstancias su impacto en los protocolos de capas altas podría no ser relevante. Además, la metodología ligera propuesta permite la interacción con tráfico real, utilizando ns-3 TAP. Por el contrario, las simulaciones con modelos muy precisos de la conectividad subyacente podrían no adaptarse a una ejecución en tiempo real (dependiendo de diferentes factores: complejidad del modelo, el hardware y el escenario), lo que dificultaría la interacción con aplicaciones e implementaciones de protocolos de transporte reales.

V. CONCLUSIÓN

La aparición de nuevas tecnologías inalámbricas y topologías de red de acceso hace que sea necesario evaluar el comportamiento los protocolos de capas superiores, tanto los más tradicionales, como nuevas soluciones que han ido apareciendo de manera más reciente. Aunque los emuladores de enlace se pueden utilizar para realizar este tipo de análisis (por ejemplo, los mecanismos de control de congestión), no están diseñados para reflejar la dinámica de las tecnologías de acceso y, en ocasiones, se limitan a escenarios más específicos. Por otro lado, la implementación detallada de procedimientos de capa baja (PHY y MAC) en los simuladores de red podría no tener un impacto significativo en el rendimiento de las soluciones de capas superiores, incrementando de manera considerable el tiempo de simulación, limitando así la interacción con las aplicaciones reales. Como complemento a los enfoques de análisis existentes, en este trabajo se describe un esquema de simulación ligero que combina el simulador ns-3 con modelos teóricos de los canales subyacentes.

El enfoque descrito, que se ha puesto a disposición de la comunidad científica a través de un repositorio público, se basa en una modificación de la implementación original de la clase *point-to-point-device* en ns-3, y varias utilidades que simplifican la definición de escenarios con diferentes topologías y configuraciones, incluyendo la posibilidad de emular enlaces con desconexiones y compartidos.

Utilizando este enfoque, se ha utilizado una red LEO, utilizando modelos de capacidad basados en cadenas de Markov para reflejar el comportamiento de los dos tipos de enlace que la conforman. A continuación, se ha validado la metodología propuesta, analizando el rendimiento del protocolo TCP en diferentes configuraciones. El correcto funcionamiento de la implementación se ha validado inicialmente sobre un canal sintético. Posteriormente, se ha ampliado el estudio, utilizando modelos más realistas de enlaces LEO, propuestos en la literatura y adaptados para ser integrados en la herramienta que se ha desarrollado. En este sentido, el enfoque propuesto permite el despliegue de diferentes topologías y escenarios, que van desde enlaces únicos hasta redes extremo a extremo, que abarcan múltiples enlaces con tráfico de fondo y desconexiones.

Hay dos líneas de trabajo diferentes que se pretender abordar en el futuro. Por un lado, se mejorará la operación de la implementación propuesta, habilitando la definición de modelos de canal mediante archivos de configuración. Esto permitirá personalizar la configuración de cada uno de los enlaces, incluso los que sean del mismo tipo, con tiempos de permanencia, probabilidades de transición y capacidades diferenciadas. Por otro lado, se aprovechará la metodología de análisis propuesta para estudiar el comportamiento de otros protocolos de transporte, en particular QUIC, así como para proponer técnicas que mejoren su rendimiento.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, MINECO-FEDER, a través del proyecto SITED: *Semantically-enabled Interoperable Trustworthy Enriched Data-spaces (PID2021-125725OB-I00).* El trabajo de Fátima Fernández ha contado con la financiación del Gobierno Vasco a través del programa Elkartek en el marco del proyecto EGIA (KK-2022/00119) y del Programa de Doctorado Industrial de la Universidad de Cantabria (Convocatoria 2020).

REFERENCIAS

[1] M. Polese, R. Jana, and M. Zorzi, "Tcp and mp-tcp in 5g mmwave networks," *IEEE Internet Computing*, vol. 21, no. 5, pp. 12–19, 2017.

[2] F. Y. Yan, J. Ma, G. D. Hill, D. Raghavan, R. S. Wahby, P. Levis, and K. Winstein, "Pantheon: the training ground for Internet congestion-control research," in *2018 {USENIX} Annual Technical Conference ({USENIX} {ATC} 18)*, 2018, pp. 731–743.

[3] R. Netravali, A. Sivaraman, K. Winstein, S. Das, A. Goyal, and H. Balakrishnan, "Mahimahi: A lightweight toolkit for reproducible web measurement," *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, p. 129–130, aug 2014. [Online]. Available: https://doi.org/10.1145/2740070.2631455

[4] F. P. Fontan, M. Vazquez-Castro, C. E. Cabado, J. P. Garcia, and E. Kubista, "Statistical modeling of the lms channel," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 6, pp. 1549–1567, 2001.

[5] A. Chen, C. Chang, and Y. Yao, "Performance evaluation of arq operations with obp and inter-satellite links: delay performance," in *IEEE 54th Vehicular Technology Conference Proceedings (VTC Fall 2001)*, vol. 4, 2001, pp. 2346–2350 vol.4.

[6] R. Hermenier, C. Kissling, and A. Donner, "A delay model for satellite constellation networks with inter-satellite links," in *2009 International Workshop on Satellite and Space Communications*, 2009, pp. 3–7.

[7] C. Hervella, L. Diez, F. Fernández, N. J. Hernández Marcano, R. Hylsberg Jacobsen, and R. Agüero, "Realistic assessment of transport protocols performance over leo-based communications," ser. PE-WASUN '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 91–98. [Online]. Available: https://doi.org/10.1145/3551663.3558680

[8] L. Rizzo, "Dummynet: a simple approach to the evaluation of network protocols," *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 1, pp. 31–41, 1997.

[9] S. Hemminger *et al.*, "Network emulation with NetEm," in *Linux conf au*, vol. 844. Citeseer, 2005.

[10] N. Handigol, B. Heller, V. Jeyakumar, B. Lantz, and N. McKeown, "Reproducible network experiments using container-based emulation," in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, 2012, pp. 253–264.

[11] A. Varga, "OMNeT++," in *Modeling and tools for network simulation*. Springer, 2010, pp. 35–59.

[12] G. F. Riley and T. R. Henderson, "The ns-3 network simulator." in *Modeling and Tools for Network Simulation*, K. Wehrle, M. Günes, and J. Gross, Eds. Springer, 2010, pp. 15–34. [Online]. Available: http://dblp.uni-trier.de/db/books/collections/Wehrle2010.html#RileyH10

[13] Y. Zhu, M. Sheng, J. Li, and R. Liu, "Performance analysis of intermittent satellite links with time-limited queuing model," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2282–2285, 2018.

[14] S. Ha, I. Rhee, and L. Xu, "Cubic: a new tcp-friendly high-speed tcp variant," *ACM SIGOPS operating systems review*, vol. 42, no. 5, pp. 64–74, 2008.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Estrategias de scheduling sobre QUIC en entornos NTN

Fátima Khan*, Fátima Fernández†, Luis Diez*, Ramón Agüero*
*Departamento de Ingeniería de Comunicaciones (DICOM), Universidad Cantabria
fatima.khan@unican.es,{ldiez,ramon}@tlmat.unican.es
†IoT and Digital Platforms Department, Ikerlan Technology Research Center
ffernandez@ikerlan.es

En este trabajo se realiza un análisis del rendimiento de los protocolos de transporte TCP y QUIC, sobre redes satelitales realistas, promoviendo una metodología innovadora que combina implementación real (técnicas de virtualización) y simulación para llevar a cabo experimentos sistemáticos y repetitivos. Con el objetivo de asegurar un análisis exhaustivo, se evalúan diferentes configuraciones, que incluyen cambios en la banda de operación y el tamaño del *buffer*. Además, se determina el impacto de la capacidad de gestionar múltiples flujos (multi-streaming) que integra QUIC. Se observa que QUIC ofrece retardos inferiores a los de TCP, aunque puede experimentar un mayor *jitter* en determinadas configuraciones. Los resultados obtenidos evidencian que el uso de múltiples flujos en QUIC no consigue una mejora significativa con el planificador por defecto, Round-Robin. Por lo tanto, se proponen estrategias de scheduling más adecuadas, capaces de obtener mejores resultados en situaciones de tráfico desequilibrado. Gracias a estas políticas, se concluye que el comportamiento de los protocolos de transporte en redes no terrestres no siempre es óptimo, y que QUIC puede aportar ventajas claras en comparación con TCP.

***Palabras Clave*—QUIC, LEO, Scheduling, LMS, TCP**

I. INTRODUCCIÓN

Se espera que la presencia de las Non-Terrestrial Networks (NTN) en las próximas generaciones de comunicaciones celulares, redes 5G y Beyond 5G (B5G), sea notablemente más relevante que en sistemas anteriores. Así, estas redes se consideran elementos fundamentales de dichos despliegues, ya que permitirán proporcionar servicios de conectividad en zonas remotas (satélite) así como desplegar rápidamente recursos de comunicación en ubicaciones específicas, utilizando Unmanned Aerial Vehicle (UAV). En definitiva, el impacto de la integración de redes NTN en arquitecturas celulares más tradicionales conlleva beneficios en términos de fiabilidad, escalabilidad, cobertura y continuidad del servicio.

Sin embargo, estos beneficios también plantean varios desafíos. Por ejemplo, es bien conocido que los protocolos de transporte tradicionales, especialmente Transmission Control Protocol (TCP), presentan un rendimiento deficiente en los enlaces inalámbricos. Se espera que esto se acentúe con la llegada de la tecnología 5G y los sistemas B5G, incluyendo las NTN [1]. Además, los protocolos de transporte tradicionales no se comportan de manera óptima con los patrones de tráfico característicos de los nuevos servicios.

Para hacer frente a las deficiencias mencionadas, la comunidad científica está trabajando en el diseño y desarrollo de nuevos protocolos de transporte que puedan superar algunas de las limitaciones que presenta TCP. Una de las alternativas más relevantes es el protocolo Quick UDP Internet Connections (QUIC), promovido originalmente por Google, y que ha sido estandarizado recientemente [2]. Entre sus ventajas, se puede destacar que habilita la gestión de múltiples flujos (multi-streaming), lo que evita el bloqueo Head-of-line (HOL), producido en TCP por una pérdida que, a pesar de afectar a un flujo concreto, afecta al resto de los que estén activos.

Con el objetivo de mejorar la eficiencia del sistema, surgen algoritmos de *scheduling* (planificación) para gestionar los recursos, aumentando el rendimiento y reduciendo el retardo. El mecanismo de planificación decide el flujo del que se extrae información para construir un paquete QUIC. El funcionamiento por defecto del protocolo QUIC es la estrategia Round-Robin (RR), en la que el reparto de la información en los paquetes es equitativo para cada flujo, utilizando siempre el mismo orden, por lo que los flujos se gestionan sin ningún tipo de prioridad.

En definitiva, en este trabajo se evalúa la combinación de QUIC y NTN y, en particular, se estudian las prestaciones de este protocolo de transporte sobre comunicaciones Low Earth Orbit (LEO). Se hace uso de una metodología novedosa, que permite imitar el comportamiento de los enlaces inalámbricos que caracterizan

la conectividad subyacente. Gracias a ella se extraen conclusiones sobre el rendimiento de protocolos de transporte sobre estas redes. En primer lugar, se comparan las prestaciones de QUIC con las de TCP y, a continuación, se estudia el comportamiento de diferentes estrategias de scheduling, con el fin de conocer las posibles ventajas que podrían aportar.

Las principales contribuciones de este trabajo son las que se enumeran a continuación:

1) Evaluación del rendimiento de la capacidad de transmisión multi-streaming de QUIC sobre canales inalámbricos muy variables (NTN).
2) Propuesta de políticas de scheduling basadas en la teoría de control dinámico de colas, que tienen como objetivo ajustar el retardo de flujos de tráfico heterogéneos.
3) Uso de una metodología que combina implementaciones reales de protocolos (técnicas de virtualización) con plataformas de simulación, imitando con precisión las características de la conectividad subyacente, lo que permite realizar experimentos sistemáticos.
4) Evaluación del rendimiento de los esquemas propuestos en comparación con alternativas más tradicionales. Los resultados muestran que los planificadores propuestos armonizan el retardo sufrido por flujos de tráfico heterogéneos, sin perjudicar el rendimiento, incluso en escenarios saturados.

El resto del documento está estructurado como se indica: en la Sección II se describen los trabajos relacionados, identificando las principales diferencias con el que se presenta en el artículo. La Sección III describe el modelo del sistema que se propone, que permite integrar esquemas de scheduling novedosos, basados en el retardo. La Sección IV introduce la metodología utilizada para evaluar el rendimiento del protocolo QUIC y las estrategias de scheduling propuestas, así como el escenario sobre el que se realizarán los experimentos, mientras que los resultados se describen en la Sección V. Finalmente, la Sección VI concluye el artículo, e introduce líneas de investigación futuras que surgen a partir de este trabajo.

II. ESTADO DEL ARTE

En los últimos años ha aumentado notablemente el interés por las comunicaciones NTN, en general, y LEO, en particular. Una de las principales causas en su posible uso como tecnología habilitadora para redes 5G y B5G [3], [4]. Liu *et al.* introducen una Satellite Access Network (SAN), prestando mayor atención a las nuevas líneas de investigación que surgen a partir de este paradigma de comunicación [5]. Más allá de las propuestas que se han realizado para las arquitecturas subyacentes, existen trabajos que analizan la convergencia de redes terrestres y no terrestres. En este sentido, Leyva-Mayorga *et al.* proporcionan una visión general de la integración de conectividad basada en satélites LEO en la red de acceso 5G/B5G, centrándose en caracterizar los enlaces físicos de la SAN resultante [6].

Los organismos de estandarización también han empezado a considerar las comunicaciones por satélite como elemento fundamental de las redes B5G. En [7] se analiza la integración de NTN y tecnologías 5G. El trabajo introduce la estrategia de estandarización del 3rd Generation Partnership Project (3GPP), y presenta los principales retos que surgen: (1) elección de arquitectura; (2) técnicas de gestión de red; y (3) modificaciones necesarias en los terminales de los usuarios. Del mismo modo, Darwish *et al.* estudian la defición de la red de acceso basada en LEO por parte de los organismos de estandarización, prestando especial atención a las normas 3GPP NR [8]. También hay trabajos que proponen diferentes soluciones y técnicas especialmente concebidas para redes de acceso basadas en LEO. Así, los autores de [9] proponen una solución de scheduling en enlaces ascendentes, para servicios Massive Machine-Type Communications (mMTC)-Narrow Band IoT (NB-IoT) soportados por constelaciones LEO. No requiere modificaciones en los dispositivos NB-IoT, y reduce el desplazamiento Doppler diferencial.

Los trabajos mencionados anteriormente sientan las bases del uso de satélites LEO para la provisión de servicios en 5G/B5G. En este contexto, un segundo grupo de trabajos se centra en las capacidades de SAN en general, y LEO en particular, para habilitar servicios de computación en el *edge*/nube asistidos por satélite.

Xie *et al.* ofrece una visión general de Satellite-Terrestrial integrated Edge Computing Networks (STECN), que combina las redes satelitales-terrestres con la computación en el *edge* para mejorar la Quality of Service (QoS) [10]. El documento analiza los principios de diseño, las funcionalidades clave, así como algunos retos de este enfoque. En la misma línea, los autores de [11] proponen la combinación de Mobile Edge Computing (MEC) con LEO, promoviendo la denominada LEO Satellite Edge Computing (LSEC), e investigan la asignación de recursos para la descarga de computación en la red LSEC.

Otros trabajos asumen una integración total de los satélites y las redes terrestres existentes. Este es el caso de [12], que considera una topología con estaciones base y pequeñas células basadas en LEO, así como enlaces de backhaul terrestres y satelitales. Sobre dicha topología, se formula un problema de optimización para minimizar el consumo energético de toda la red, manteniendo los requisitos de QoS. Aunque estos trabajos comparten el marco de aplicación con el que se presenta en este documento, el objetivo es diferente, por lo que la contribución que aquí se presenta es complementaria.

En ese sentido, este trabajo se centra en la evaluación del rendimiento de protocolos de transporte tradicionales (TCP) y emergentes (QUIC) sobre estas redes. Para contextualizar esta línea de investigación, se indican a continuación algunos trabajos que han estudiado el comportamiento de QUIC, comparándolo con TCP.

Shreedhar *et al.* evalúan el rendimiento obtenido con QUIC utilizando tráfico web, y servicios de almacenamiento en la nube y vídeo en entornos no controla-

dos [13]. Sin embargo, no prestan mucha atención a la red de acceso subyacente. Del mismo modo, Qian *et al.* analizan el rendimiento de QUIC en redes de acceso LTE Advanced (LTE-A) [14]. Los autores estudian el comportamiento de diferentes algoritmos de control de congestión, y amplían la operación habitual de QUIC, aprovechando su funcionalidad multi-camino, a través de un prototipo que permite comunicaciones duales WiFi y LTE-A. Más cerca del escenario de computación en el *edge*/nube, Dizdarević y Jukan evalúan el rendimiento de QUIC para el continuo IoT-*edge*-nube [15]. La evaluación se basa en dos escenarios diferentes: (1) IoT-edge, y (2) IoT-edge-cloud. El análisis, que utiliza dispositivos Raspberry Pi, se centra en la arquitectura de la nube, mientras que se presta poca atención a la red de acceso.

Como se puede observar, no existen trabajos que hayan afrontado la evaluación de protocolos de transporte en general, y de QUIC en particular, sobre SAN y, más concretamente, sobre escenarios que integran redes terrestres y no terrestres.

En el pasado, Tsonuda *et al.* analizan el rendimiento de Stream Control Transmission Protocol (SCTP) sobre redes LEO utilizando multi-stream adaptativo [16]. Además, la característica multi-homing de SCTP también se utilizó sobre redes LEO para mejorar los traspasos [17]. Respecto a QUIC, Yang *et al.* abordan la interacción entre redes LEO y QUIC, pero desde una perspectiva diferente [18]. Se centran en Multi-Path QUIC (MPQUIC), y proponen un modelo de rendimiento para LEO (traspaso, interrupción, etc.).

Martin y Khademi analizan la idoneidad del control de congestión Bottleneck Bandwidth and Round-trip propagation time (BBR), utilizando QUIC, sobre redes Geoestacionary Equatorial Orbit (GEO) SATCOM [19]. Así, estudian el comportamiento de la combinación de QUIC y BBR sobre enlaces satelitales en órbita geoestacionaria.

El estudio que parece más cercano al que aquí se presenta, es el trabajo de Yang, Li y Wu [20]. Modelan el canal inalámbrico con Matlab Satellite Tool Kit (STK) y utilizan las trazas correspondientes para analizar el rendimiento de QUIC sobre constelaciones LEO. También emplean la emulación sobre dispositivos reales, pero utilizan enlaces cableados de alta capacidad, cuyas características (retardo y tasa de pérdida) se modifican estadísticamente en base a los resultados previos. La diferencia con el enfoque de evaluación que aquí se propone es significativa, ya que: (1) no tienen en cuenta la longitud del *buffer* en los dispositivos, por lo que no consideran el retardo correspondiente; (2) en este trabajo se asume que las pérdidas son consecuencia directa de las fluctuaciones en la capacidad del canal, causando saturación de los *buffer* en los diferentes enlaces; (3) los autores de [20] usan archivos pequeños (páginas web de unas decenas de `kB`), mientras que en este trabajo se contemplan flujos de mayor duración, a partir del envío de archivos grandes; (4) los experimentos realizados en este trabajo tienen en cuenta además la capacidad del canal, permitiendo evaluar así el impacto de la carga de tráfico (tasa de aplicación).

Por último, cabe mencionar que ninguno de los trabajos analizados aborda realmente el estudio de las técnicas de scheduling para orquestar la funcionalidad multi-stream de QUIC. Existen pocos trabajos que se hayan centrado en el scheduling, pero lo han hecho siempre en el ámbito de las comunicaciones múlti-camino, no como medio para gestionar los múltiples *streams* que una conexión QUIC puede gestionar [21], [22].

En definitiva, este trabajo cubre dos aspectos que no han sido tratados hasta ahora por la comunidad científica: (1) la evaluación de protocolos de capa de transporte, en particular QUIC, sobre nuevas topologías de red de acceso, abarcando constelaciones de satélites; y (2) el análisis de técnicas de scheduling que aproveche la capacidad que proporciona QUIC de gestionar varios flujos (streams) de manera simultánea.

III. MODELO DEL SISTEMA

En esta sección se procede a la explicación de las estrategias propuestas así como las soluciones de referencia. Con este objetivo se modifica el funcionamiento del planificador incluido por defecto en la implementación de QUIC que se utiliza en los experimentos, RR. Esta estrategia selecciona secuencialmente uno de los flujos activos en cada oportunidad de transmisión. Inicialmente se han implementado Fair Queuing (FQ) y Weighted Fair Queuing (WFQ), con fines comparativos. Con FQ, se distribuyen equitativamente la capacidad compartida entre los flujos activos y, siempre que se satisface la necesidad de uno, el excedente se vuelve a repartir equitativamente entre el resto. En WFQ el comportamiento es similar, pero en este caso, la capacidad se distribuye en función de ciertos pesos, que se utilizan para priorizar determinados flujos, asignándoles más recursos.

Estos tres planificadores tienen algunas limitaciones. Por un lado, RR y FQ gestionan todos los flujos por igual, y no permiten, por tanto, establecer prioridades. Por otro lado, WFQ es capaz de asignar cierta priorización, pero los pesos deben establecerse a priori, por lo que no podría adaptarse a las variaciones del canal o del tráfico.

Para superar esta limitación, se propone un planificador Back Pressure (BP), basado en la teoría de Lyapunov [23]. Esta política pretende garantizar la estabilidad de las colas de cada flujo para cualquier condición de tráfico y canal, siempre que se respete la capacidad promedio. Cabe destacar que, a diferencia de WFQ, la solución propuesta no requiere ninguna configuración previa.

La estrategia de scheduling entre los flujos activos se modela como un sistema de colas. $Q_k(t)$ corresponde a la ocupación del *buffer* de aplicación para el flujo k-ésimo, en cualquier momento t, mientras que la decisión del scheduling para ese flujo se define como $\alpha_k(t)$. La capacidad de transmisión de la conexión en cualquier momento t se modela como una variable aleatoria $\omega(t)$, cuyo valor depende de las condiciones del canal y del mecanismo de control de la congestión. Bajo estas premisas, cada cola se actualiza siguiendo la siguiente expresión:

$$Q_k(t+1) = \max[Q_k(t) - b_k(t), 0] + a_k(t) \qquad (1)$$

donde $a_k(t)$ y $b_k(t)$ son las variables de llegada y salida, respectivamente. Cabe señalar que $b_k(t)$ corresponde en realidad a cada variable de decisión $\alpha_k(t)$, mientras que $a_k(t)$ es una variable aleatoria sobre la que no se tiene control alguno. En cada instante de tiempo, se toma una decisión de scheduling a partir de un conjunto $\mathcal{A}$ que estabilice las colas de los flujos de aplicación, garantizando que no se supere la capacidad de transmisión, $\sum_k \alpha_k(t) \leq \omega(t)\ \forall t$. Como se ha expuesto anteriormente, este problema se resuelve mediante el algoritmo BP, basado en la teoría de Lyapunov. Por lo tanto, en cada instante de tiempo se toma la decisión que optimiza el siguiente problema:

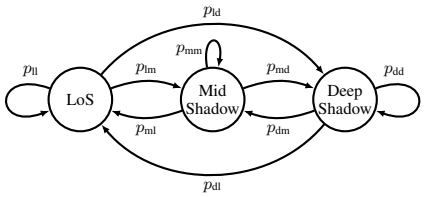
$$\begin{aligned} \max_{\alpha(t)} \quad & \sum_{k=1}^{N} Q_k(t) \cdot b_k(t) \\ \textbf{s.t.} \quad & \alpha \in A \end{aligned} \tag{2}$$

Finalmente, su comportamiento se traduce en seleccionar el flujo con mayor ocupación en el *buffer* de aplicación para ser transmitido.

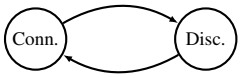
IV. ESCENARIO DE APLICACIÓN

En esta sección se evalúa el rendimiento de los planificadores propuestos sobre comunicaciones NTN, en particular una red de satélites LEO. Se pueden distinguir dos configuraciones: un único enlace Land Mobile-Satellite (LMS), entre una estación terrestre y un satélite; y una comunicación extremo a extremo entre dos estaciones terrestres, incluyendo dos enlaces LMS y varios enlaces Inter-Satellite Link (ISL). Estos últimos podrían sufrir desconexiones temporales, debido al continuo movimiento de los satélites en constelaciones LEO. Para modelar los dos tipos de enlace, se utilizan las cadenas de Markov mostradas en la Fig. 1: para el enlace LMS, se parte del trabajo de Fontán *et al.*, que considera tres situaciones/estados diferentes: (l) Line of Sight (LoS), con condiciones ideales de propagación; (m) mid-shadowing, donde las condiciones se deterioran; y (d) deep-shadowing, en el que la conectividad se ve gravemente perjudicada [24]. Para el ISL, se utiliza una cadena de dos estados, para capturar las desconexiones temporales. Otros trabajos también han utilizado cadenas de Markov para modelar enlaces por satélite [25], [26], [27].

La implementación que se usa de QUIC en este trabajo es `quic-go` (versión 0.15.1), que se modifica para integrar las políticas de scheduling propuestas. La plataforma que se usa para evaluar su rendimiento se representa en la Fig. 2. Combina una implementación real del protocolo QUIC (usando virtualización y contenedores Docker) con la simulación de la conectividad subyacente, explotando el simulador *ns-3*. En ese sentido, se cambia el funcionamiento por defecto del enlace *point-to-point* de ns-3, de forma que sus características puedan ser modificadas durante el experimento, de acuerdo con la dinámica establecida por las cadenas de Markov antes mencionadas. De esta forma, se captura el dinamismo de los enlaces LMS y ISL.



(a) Cadena de Markov de tres estados para el canal LMS



(b) Cadena de Markov de dos estados para emular desconexiones temporales en enlaces ISL

Fig. 1: Modelos de cadena de Markov para los enlaces LMS e ISL, utilizados para la evaluación del rendimiento.

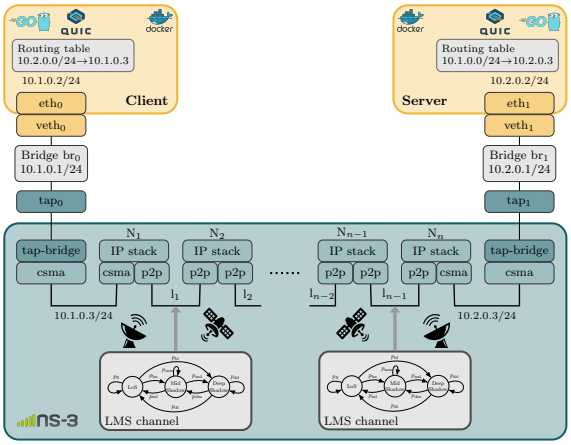


Fig. 2: Esquema de la plataforma utilizado para la evaluación, que integra contenedores docker, ns-3, y los modelos LMS e ISL

V. RESULTADOS

En esta sección se discuten los resultados de la evaluación de TCP y QUIC, incluyendo las diversas estrategias de scheduling, sobre escenarios LEO. Los principales parámetros de la configuración del escenario que se contempla se resumen en la Tabla I. Como puede observarse, se usan dos bandas diferentes, *Ka* y *S*, cada una de ellas caracterizada por su matriz de probabilidades de transición, el tiempo medio de permanencia en cada estado, y las tasas de transmisión correspondientes. La velocidad de transmisión máxima, que corresponde a la situación de línea de visión directa, es de 4 y 80 `Mbps` para las bandas *S* y *Ka*, respectivamente. Además, las capacidades del mid- y deep-shadowing se fijan al 50% y 20% de dicha capacidad máxima.

Se compara en primer lugar el rendimiento de QUIC y TCP sobre enlaces LMS, utilizando sus configuraciones por defecto. A continuación, la evaluación se centra en QUIC, analizando el comportamiento de los schedulers descritos en la Sección III sobre enlaces LMS, considerando flujos de tráfico asimétricos. Finalmente, se amplía el escenario, y se analiza el impacto de las desconexiones en enlaces ISL sobre una conectividad

Tabla I: Configuración del escenario

Banda *Ka* Parámetros LMS [24, Table XVII]	
Tasa LMS	$[80, 40, 16]$ Mbps
Tasa del enlace	45.33 Mbps
Matriz de transición LMS	$\mathcal{P} = \begin{pmatrix} 0 & 0.93156 & 0.068437 \\ 0.34526 & 0 & 0.65474 \\ 0.070012 & 0.92999 & 0 \end{pmatrix}$
Tiempo de estancia de LMS	$[0.2530, 0.7299, 0.1666]$ s
δ	100 ms
Banda *S* Parámetros LMS [24, Table XVII]	
Tasa de LMS	$[4, 2, 0.8]$ Mbps
Tasa del enlace	2.32 Mbps
Matriz de transición LMS	$\mathcal{P} = \begin{pmatrix} 0 & 0.94076 & 0.059243 \\ 0.77084 & 0 & 0.22916 \\ 0.49418 & 0.50582 & 0 \end{pmatrix}$
Tiempo de estancia LMS	$[0.5485, 0.4992, 0.3529]$ s
δ	100 ms
Aplicación y búfer	
Tamaño del búfer	$[7, 15, \infty]$ Paquetes
Tamaño de paquete	1000 Bytes

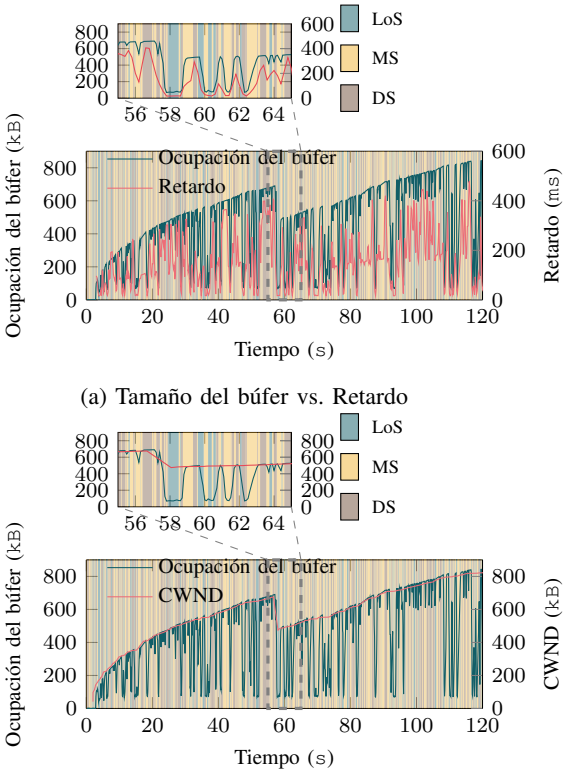
extremo a extremo.

A. *Rendimiento de TCP y QUIC en un único enlace LMS*

Antes de analizar el rendimiento medio de TCP y QUIC en enlaces LMS, la Figura 3 ilustra la variabilidad mostrada por un canal LMS concreto y su impacto en la ocupación del *buffer*, la ventana de congestión (cwnd) y el retardo. Aquí, el *buffer* se refiere al del dispositivo punto a punto (Fig. 2), y su ocupación es un indicador de la adaptación del mecanismo de control de congestión a la variabilidad del canal. Los resultados se obtienen a partir de una única transmisión QUIC, con un flujo y tasa de tráfico de aplicación constante, sobre la banda *Ka*, y con capacidad de *buffer* infinita. Los colores de fondo de la Figura 3 representan el estado del canal, y las líneas la evolución instantánea de los diferentes parámetros que se monitorizan.

Como puede observarse en la Figura 3a, el tamaño del *buffer* aumenta durante la conexión, en la que se transmite un fichero de gran tamaño, con variaciones lentas, correladas con la evolución de la ventana de congestión, y variaciones rápidas que dependen del estado concreto del canal. Al mismo tiempo, se observa que el retardo de los paquetes sigue la misma tendencia que la ocupación del *buffer*, mostrando grandes variaciones (lo que daría lugar a un *jitter* elevado) cuando el canal pasa de un estado a otro. Por otro lado, la Figura 3b, muestra la evolución de la ventana de congestión en el mismo experimento. Como puede observarse, este parámetro (cwnd) crece mientras el tamaño del *buffer* aumenta, hasta que se produce un evento de tiempo de expiración, en el que la ventana de congestión disminuye su valor, al igual que la ocupación del *buffer*. En definitiva, la Figura 3 muestra que el enlace LMS dificulta la adaptación de los protocolos de transporte a la capacidad de transmisión real, lo que tiene un fuerte impacto en el rendimiento de la aplicación, particularmente en el retardo y el jitter.

A continuación, se compara el rendimiento obtenido por



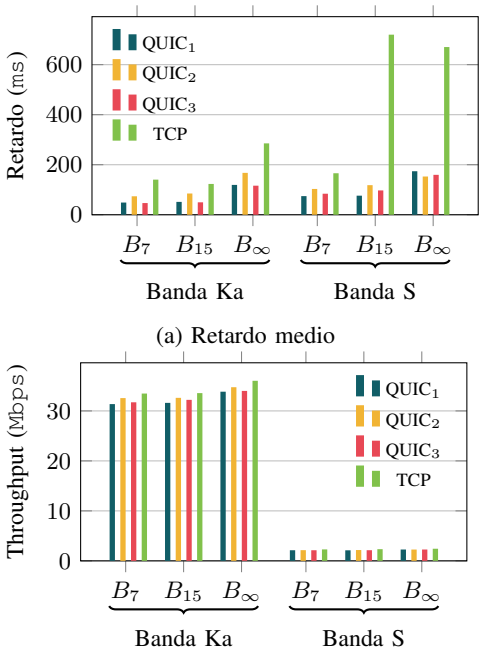
(a) Tamaño del búfer vs. Retardo

(b) Tamaño del búfer vs. Ventana de congestión

Fig. 3: Ejemplo de evolución del tamaño del *buffer*, ventana de congestión y retardo sobre un canal LMS con *buffer* de capacidad infinita.

QUIC y TCP sobre enlaces *LMS* con diferentes configuraciones. Cabe destacar que ambos protocolos utilizan CUBIC como mecanismo de control de congestión, por lo que es previsible que las variaciones en la conectividad subyacente tenga un impacto similar sobre ambos protocolos [28]. La Figura 4 muestra el retardo por paquetes y el rendimiento de la aplicación obtenidos en ambas bandas, *Ka* y *S*, para diferentes tamaños de *buffer*. Los resultados se promedian sobre 30 ejecuciones independientes, y cada experimento comprende transmisiones activas durante 60 *s*. En todos los casos, la aplicación genera tráfico a la capacidad media del enlace. Al utilizar QUIC, se representan los resultados obtenidos para distintos números de flujos, utilizando el scheduler por defecto, RR.

Como se podía esperar, se observa que el tamaño del *buffer* tiene un fuerte impacto sobre el retardo. Sin embargo, los resultados muestran que el rendimiento de TCP se ve más perjudicado con valores muy elevados, al utilizar un *buffer* infinito. Dado que ambos protocolos utilizan el mismo control de congestión, esta diferencia se debe a las mejoras en la detección de pérdidas que se implementan en QUIC [29]. Por otro lado, la tasa media es bastante estable, independientemente del protocolo utilizado o del tamaño del *buffer*. En cuanto a la capacidad multi-stream de QUIC, los resultados demuestran que tiene poco impacto en el comportamiento del protocolo. A la vista de los resultados obtenidos se podría concluir que QUIC aporta



(a) Retardo medio

(b) Rendimiento medio

Fig. 4: Comparación de rendimiento entre TCP y QUIC sobre un único enlace LMS. Los resultados medios se obtienen a partir de 30 ejecuciones en las que se genera tráfico durante 1 minuto a una tasa igual a la capacidad media del canal.

Tabla II: Configuración del tráfico de aplicación

Configuración común	
Capacidad media del canal (*Ka*)	45 *Mbps*
Tamaño del paquete	1000 *Bytes*
Distribución	Poisson
Tiempo de generación del tráfico	60 s
Escenario 1 (50%)	
Tasa transmisión de datos	$Stream_1$: 15 *Mbps*; $Stream_2$: 7.5 *Mbps*
Escenario 2 (80%)	
Tasa transmisión de datos	$Stream_1$: 27 *Mbps*; $Stream_2$: 15 *Mbps*

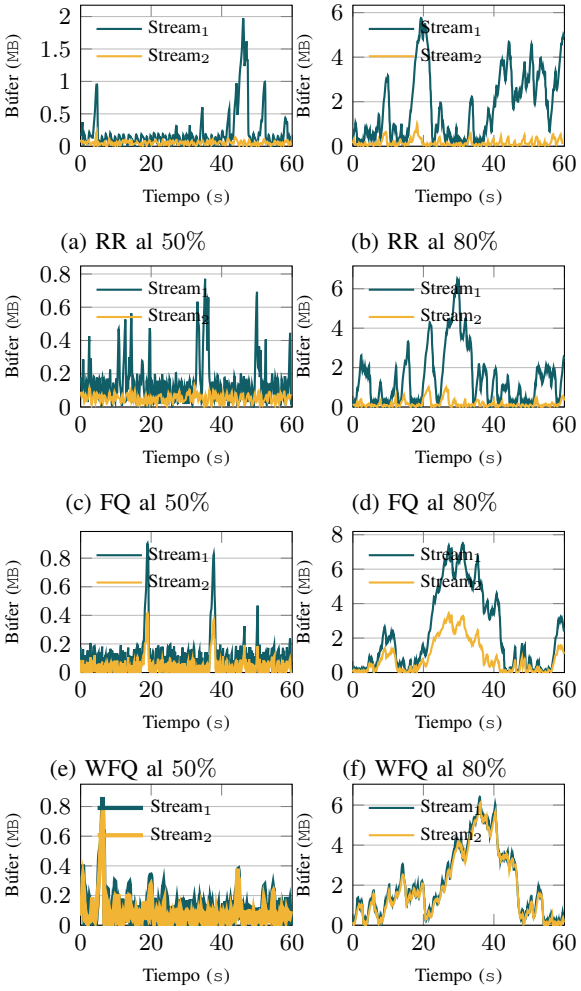


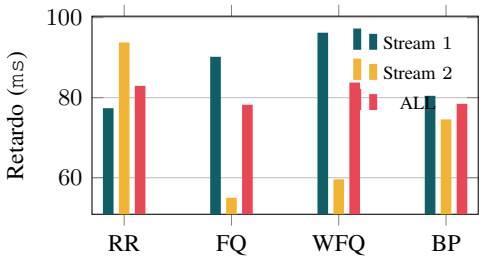
Fig. 5: Evolución del *buffer* de los dos flujos a lo largo del tiempo, utilizando diferentes algoritmos de scheduling y tasas de envío, sobre un enlace LMS.

una solución de transporte más adecuada para canales muy variables, tales como los enlaces LMS.

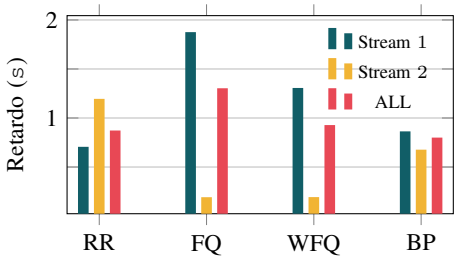
B. Rendimiento de las estrategias de scheduling sobre un canal LMS

En primer lugar, la Figura 5 muestra la evolución del *buffer* de aplicación utilizando los distintos flujos (filas) y tasas de tráfico (columnas). Se observa que RR y WFQ (primera y segunda filas) no son capaces de adaptarse al desequilibrio de tráfico, dando lugar a ocupaciones del *buffer* de aplicación bastante diferentes. Este comportamiento es más acusado a medida que la tasa de tráfico de aplicación se acerca a la capacidad media del canal. Por otro lado, WFQ proporciona una distribución de recursos (capacidad del canal) más justa entre los dos flujos, dando lugar a una ocupación de *buffer* más equilibrada a medida que se aumenta el tráfico. Sin embargo, no aprovecha de forma óptima a las situaciones en la que las tasas de ambos flujos no se corresponden de manera exacta con su valor medio. Este comportamiento subóptimo se corrige con el tiempo cuando el canal no está saturado, aunque en circunstancias cercanas a la saturación no se observa este efecto, y WFQ es incapaz de proporcionar una distribución justa de los recursos compartidos. Por otro lado, la política propuesta, basada en el algoritmo BP, es capaz de armonizar la ocupación del *buffer* en ambos flujos, independientemente de la variación del canal, incluso cuando el tráfico alcanza la capacidad del mismo, debido a que comprueba la ocupación del *buffer* al construir el paquete QUIC.

A continuación, se analiza el impacto de los schedulers en el comportamiento percibido. Se presta atención al retardo, ya que todos los experimentos anteriores muestran que la diferencia en términos de rendimiento es insignificante. La Figura 6 muestra el retardo medio por paquete (en la aplicación), para cargas de tráfico del 50% y 80% de la capacidad media del canal. Los resultados se obtienen a partir de 30 experimentos independientes, en los que se llevan a cabo conexiones de 60 segundos de duración. Para

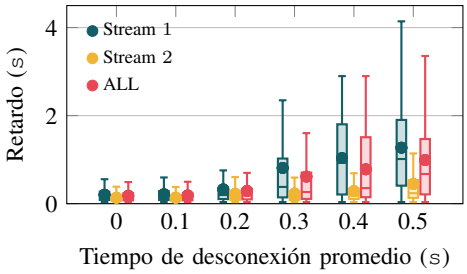


(a) Tasa de tráfico al 50%

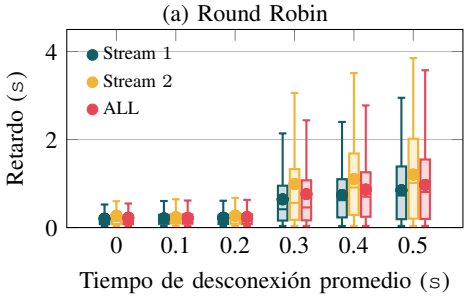


(b) Tasa de tráfico al 80%

Fig. 6: Retardo medio experimentado en un enlace LMS de banda Ka, cuando se utilizan los distintos planificadores



(a) Round Robin



(b) Back Pressure

Fig. 7: Distribución de retardos con RR y BP sobre una red E2E con interrupciones e en el canal ISL.

ilustrar mejor el comportamiento se muestran los retardos medios experimentados por cada flujo de forma independiente, así como el promedio de la conexión global. Como puede observarse, los schedulers estáticos (RR, FQ y WFQ) dan lugar a retardos sensiblemente desbalanceados entre ambos flujos, independientemente de la carga de tráfico. Curiosamente, la implementación por defecto RR produce una distribución más justa del retardo, debido al mejor uso de la capacidad total de transmisión. Por otro lado, el algoritmo BP propuesto en este trabajo es capaz de armonizar el retardo de ambos flujos, prácticamente consiguiendo el mismo valor. Además, la solución propuesta produce una ligera reducción global del retardo.

C. Estrategias de planificación sobre escenarios End to End (E2E)

A continuación, se analizan las políticas de programación con una configuración extremo a extremo, abarcando dos enlaces *Ka* que emulan las condiciones de conectividad entre las estaciones terrestres y los satélites, conectados por un único enlace. Como se mencionó en la Sección IV, el ISL se modela con una cadena de Markov de dos estados, lo que permite tener en cuenta situaciones de interrupción en la comunicación, habituales en comunicaciones basadas en redes LEO.

Utilizando la misma configuración que antes, se consideran 2 flujos QUIC desequilibrados, donde la tasa del primero es el doble que la del segundo. Dado que se contemplan situaciones de interrupción con distinta duración, lo que tiene como consecuencia una disminución de la capacidad global de la red, la tasa de tráfico agregada se fija en 50% de la capacidad media LMS. El enlace ISL se modela con tiempos de permanencia distribuidos exponencialmente: (1) periodo activo con una tasa de transmisión de 80 `Mbps` y un tiempo medio de permanencia de $5s$; y (2) estado inactivo que emula las eventuales desconexiones del enlace y cuyo tiempo medio de permanencia se incrementa de 0 a 0.5 segundos. Para cada valor del tiempo medio de interrupción, se ejecuta una simulación independiente, con una duración de 120 s, lo que garantiza un número suficiente de transiciones de los estados en los diferentes canales, obteniendo resultados estadísticamente significativos. La ejecución sólo se detiene cuando se reciben todos los bytes en la aplicación receptora.

La Figura 7 muestra la distribución del retardo medio extremo a extremo de ambos flujos para las cuatro estrategias de scheduling estudiadas, a medida que se aumenta el tiempo medio de interrupción en el enlace ISL. Se puede observar que el retardo medio de extremo a extremo aumenta de forma constante, así como su dispersión. Esta variabilidad es ligeramente mayor cuando se utiliza el algoritmo propuesto en este trabajo, BP, que, por otro lado, consigue garantizar un comportamiento más equilibrado para los dos *streams*, ya que consigue armonizar los *buffer* para todas las configuraciones.

VI. CONCLUSIÓN

Este trabajo evalúa el rendimiento de los protocolos de transporte (TCP y QUIC) sobre redes LEO. Como se ha podido ver, se trata de uno de los primeros trabajos en afrontar un análisis con este objetivo. Para ello se ha propuesto una metodología novedosa que combina técnicas de virtualización, implementaciones reales de protocolos, y entornos de simulación de red, para realizar experimentos sistemáticos con un modelado realista de la conectividad subyacente. Los resultados demuestran que QUIC mejora el comportamiento de TCP sobre canales con alta variabilidad.

Además, se han caracterizado minuciosamente distintas estrategias de scheduling que explotan la capacidad multi-streaming de QUIC. Se ha visto que el algoritmo utilizado por defecto en la implementación `quic-go` es subóptimo con flujos con tráfico desequilibrado. La estrategia WFQ puede mantener la velocidad de transmisión sin afectar significativamente al retardo medio, pero no es capaz de adaptarse a situaciones de tráfico desequilibrado. Por otro lado, el esquema propuesto, basado en el algoritmo *Back-pressure*, estabiliza dinámicamente las longitudes de cola de los flujos activos, utilizando únicamente el estado del buffer.

La metodología propuesta permite aumentar la complejidad de la conectividad subyacente introduciendo tiempos de desconexión en enlaces ISL (como se ha descrito en los últimos experimentos comentados en el artículo) o tráfico de fondo. Como trabajo futuro, se explotará esta funcionalidad para ampliar el análisis e incluir otras características, como los mecanismos de control de la congestión. Además, se estudiará la utilización de nuevos algoritmos de scheduling que tengan en cuenta explícitamente el retardo del tráfico.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, MINECO-FEDER, a través del proyecto SITED: *Semantically-enabled Interoperable Trustworthy Enriched Data-spaces (PID2021-125725OB-I00)*. El trabajo de Fátima Fernández ha contado con la financiación del Gobierno Vasco a través del programa Elkartek en el marco del proyecto EGIA (KK-2022/00119) y del Programa de Doctorado Industrial de la Universidad de Cantabria (Convocatoria 2020).

REFERENCIAS

[1] Poorzare, Reza and Calveras, Anna, "Open Trends On TCP Performance Over Urban 5G MmWave Networks," in *Proceedings of the 17th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, ser. PE-WASUN '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 85–92. [Online]. Available: https://doi.org/10.1145/3416011.3424749

[2] J. Iyengar and M. Thomson, "QUIC: A UDP-Based Multiplexed and Secure Transport," RFC 9000, may 2021. [Online]. Available: https://rfc-editor.org/rfc/rfc9000.txt

[3] I. F. Akyildiz, A. Kak, and S. Nie, "6g and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133 995–134 030, 2020.

[4] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6g: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021.

[5] S. Liu, Z. Gao, Y. Wu, D. W. Kwan Ng, X. Gao, K.-K. Wong, S. Chatzinotas, and B. Ottersten, "Leo satellite constellations for 5g and beyond: How will they reshape vertical domains?" *IEEE Communications Magazine*, vol. 59, no. 7, pp. 30–36, 2021.

[6] I. Leyva-Mayorga, B. Soret, M. Röper, D. Wübben, B. Matthiesen, A. Dekorsy, and P. Popovski, "Leo small-satellite constellations for 5g and beyond-5g communications," *IEEE Access*, vol. 8, pp. 184 955–184 964, 2020.

[7] M. Hosseinian, J. P. Choi, S.-H. Chang, and J. Lee, "Review of 5g ntn standards development and technical challenges for satellite integration with the 5g network," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 8, pp. 22–31, 2021.

[8] T. Darwish, G. K. Kurt, H. Yanikomeroglu, M. Bellemare, and G. Lamontagne, "Leo satellites in 5g and beyond networks: A review from a standardization perspective," *IEEE Access*, vol. 10, pp. 35 040–35 060, 2022.

[9] O. Kodheli, S. Andrenacci, N. Maturo, S. Chatzinotas, and F. Zimmer, "An uplink ue group-based scheduling technique for 5g mmtc systems over leo satellite," *IEEE Access*, vol. 7, pp. 67 413–67 427, 2019.

[10] R. Xie, Q. Tang, Q. Wang, X. Liu, F. R. Yu, and T. Huang, "Satellite-terrestrial integrated edge computing networks: Architecture, challenges, and open issues," *IEEE Network*, vol. 34, no. 3, pp. 224–231, 2020.

[11] K. Wei, Q. Tang, J. Guo, M. Zeng, Z. Fei, and Q. Cui, "Resource scheduling and offloading strategy based on leo satellite edge computing," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, 2021, pp. 1–6.

[12] Z. Tang, H. Zhou, T. Ma, K. Yu, and X. S. Shen, "Leveraging leo assisted cloud-edge collaboration for energy efficient computation offloading," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[13] T. Shreedhar, R. Panda, S. Podanev, and V. Bajpai, "Evaluating quic performance over web, cloud storage, and video workloads," *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 1366–1381, 2022.

[14] P. Qian, N. Wang, and R. Tafazolli, "Achieving robust mobile web content delivery performance based on multiple coordinated quic connections," *IEEE Access*, vol. 6, pp. 11 313–11 328, 2018.

[15] J. Dizdarević and A. Jukan, "Experimental benchmarking of http/quic protocol in iot cloud/edge continuum," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.

[16] H. Tsunoda, N. Kato, A. Jamalipour, and Y. Nemoto, "Performance evaluation of sctp wth adaptive multistreamiing over leo satellite networks," in *2007 International Workshop on Satellite and Space Communications*, 2007, pp. 150–154.

[17] Z. Zhang, Q. Guo, and Z. Gao, "A prediction based sctp handover scheme for ip/leo satellite network," in *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, 2010, pp. 1–4.

[18] W. Yang, S. Shu, L. Cai, and J. Pan, "Mm-quic: Mobility-aware multipath quic for satellite networks," in *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*, 2021, pp. 608–615.

[19] A. Martin and N. Khademi, "On the suitability of bbr congestion control for quic over geo satcom networks," in *Proceedings of the Workshop on Applied Networking Research*, ser. ANRW '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3547115.3547194

[20] S. Yang, H. Li, and Q. Wu, "Performance analysis of quic protocol in integrated satellites and terrestrial networks," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 2018, pp. 1425–1430.

[21] X. Shi, L. Wang, F. Zhang, B. Zhou, and Z. Liu, "Pstream: Priority-based stream scheduling for heterogeneous paths in multipath-quic," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, 2020, pp. 1–8.

[22] T. Viernickel, A. Froemmgen, A. Rizk, B. Koldehofe, and R. Steinmetz, "Multipath quic: A deployable multipath transport protocol," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.

[23] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan and Claypool Publishers, 2010.

[24] F. P. Fontan, M. Vazquez-Castro, C. E. Cabado, J. P. Garcia, and E. Kubista, "Statistical modeling of the lms channel," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 6, pp. 1549–1567, 2001.

[25] A. Chen, C. Chang, and Y. Yao, "Performance evaluation of arq operations with obp and inter-satellite links: delay performance," in *IEEE 54th Vehicular Technology Conference Proceedings (VTC Fall 2001)*, vol. 4, 2001, pp. 2346–2350 vol.4.

[26] R. Hermenier, C. Kissling, and A. Donner, "A delay model for satellite constellation networks with inter-satellite links," in *2009 International Workshop on Satellite and Space Communications*, 2009, pp. 3–7.

[27] Y. Zhu, M. Sheng, J. Li, and R. Liu, "Performance analysis of

intermittent satellite links with time-limited queuing model," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2282–2285, 2018.

[28] S. Ha, I. Rhee, and L. Xu, "Cubic: a new tcp-friendly high-speed tcp variant," *ACM SIGOPS operating systems review*, vol. 42, no. 5, pp. 64–74, 2008.

[29] J. Iyengar and I. Swett, "QUIC Loss Detection and Congestion Control," RFC 9002, May 2021. [Online]. Available: https://www.rfc-editor.org/info/rfc9002

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Procesamiento de Tráfico en el Kernel de Linux con Machine Learning

Jorge Gallego-Madrid, Irene Bru-Santa, Ramon Sanchez-Iborra, Antonio Skarmeta
Departamento de Ingeniería de la Información y de las Comunicaciones,
Universidad de Murcia
Facultad de Informática, Campus de Espinardo, 30100, Murcia, España.
{jorgegm, ireneleonor.brus, ramonsanchez, skarmeta}@um.es

Resumen

El procesado de tráfico mediante mecanismos basados en *Machine Learning* (ML) está cobrando gran auge durante los últimos años dada la complejidad de los nuevos sistemas de comunicaciones y la necesidad de tener un control avanzado de los flujos que éstos transportan. Así, nuevas técnicas y herramientas están surgiendo para permitir la integración de algoritmos inteligentes en el camino que siguen los datos extremo a extremo. Una de ellas es *extended Berkeley Packet Filter* (eBPF), que permite la ejecución de programas dentro del kernel de Linux sin tener que recompilar éste, enriqueciendo su funcionalidad. En este trabajo aprovechamos las opciones que brinda eBPF para integrar dentro del kernel de Linux un algoritmo complejo de análisis de tráfico basado en ML. Esta estrategia de desarrollo permite reducir notablemente latencias de procesamiento con respecto a soluciones similares implementadas fuera del kernel, además de permitir el despliegue de funciones de red avanzadas de forma ágil en dispositivos no especializados en tareas de gestión de red.

***Palabras Clave*—eBPF, red neuronal, XDP.**

I. INTRODUCCIÓN

La llegada de las redes de próxima generación (*Next-Generation Networks*, NGN) exigirá mejoras significativas en las infraestructuras de comunicación en términos de velocidad, flexibilidad, inteligencia y latencia. Esto permitirá el desarrollo de nuevos casos de uso en todos los verticales identificados para las redes más allá del 5G (*Beyond 5G*, B5G) [1]. Para hacer frente a esta explosión, actualmente están surgiendo múltiples soluciones para la gestión de redes avanzadas y el tráfico que éstas transportan, ya que la complejidad para coordinar los diferentes segmentos de red que componen las arquitecturas B5G desagregadas está en constante aumento [2]. Además, la transición a esta nueva generación de redes plantea un gran desafío debido a la inversión necesaria en nuevo *hardware*, *software* e infraestructura.

En este sentido, se necesitarán tecnologías novedosas para proporcionar flexibilidad e inteligencia a las infraestructuras NGN mediante la adopción de un enfoque de "softwarización" de la red. Una tecnología de vanguardia en este sentido es el *extended Berkeley Packet Filter* (eBPF), que permite un procesamiento de tráfico ultra rápido en equipamiento no especializado en tareas de red, al permitir la ejecución segura de código dentro del kernel de Linux [3]. Esto resulta de gran relevancia, dadas las capacidades que se habilitan para este tipo de equipos, concretamente en escenarios que consideran el continuo *fog/edge/cloud*. Con la ayuda de técnicas de aprendizaje automático (*Machine Learning*, ML), eBPF es un habilitador dentro del ámbito de las NGNs para realizar tareas inteligentes de gestión de red y monitoreo de tráfico en cualquier punto de la infraestructura, lo cual es un pilar fundamental para alcanzar el alto rendimiento y baja latencia demandados por los nuevos servicios y aplicaciones soportados por las infraestructuras B5G e, incluso, permitir el despliegue de funciones de red orientadas a la ciberseguridad. Además, es bien sabido que ML es una de las tecnologías clave para proporcionar una toma de decisiones inteligente en la gestión y orquestación de la red [4]. Permite que los dispositivos de red se adapten automáticamente a las condiciones cambiantes del sistema, ya que los mecanismos de control basados en ML pueden detectar anomalías, predecir el comportamiento de la red o anticipar fallos y cuellos de botella. De esta forma, esta proactividad se puede utilizar para automatizar tareas de gestión de red y optimizar las redes en tiempo real sin intervención humana, siguiendo el paradigma *Zero Touch Network Service Management* (ZSM) [5].

Las sinergias entre eBPF y ML han sido exploradas en la literatura en los últimos años. Normalmente, se utilizan juntos de la siguiente manera: eBPF se encarga de recopilar datos de los flujos de tráfico a nivel del espacio de kernel de Linux y los modelos ML, en el plano usuario, analizan esa información y hacen predic-

ciones o decisiones. eBPF permite una baja sobrecarga de procesamiento al inspeccionar el tráfico y el estado de los recursos *hardware*, lo que lo hace muy adecuado como herramienta que no introduce alta latencia ni penalizaciones computacionales. En cuanto a los algoritmos ML, pueden procesar eficazmente los datos que se les proporciona para detectar patrones en el tráfico y predecir futuras condiciones de red o, incluso, detectar ataques como se mencionó anteriormente. En este contexto, presentamos como principal contribución de este trabajo (aún en desarrollo), una solución que integra modelos avanzados de inspección de tráfico dentro del kernel de Linux, aprovechando las capacidades proporcionadas por eBPF para combinar el procesamiento rápido de paquetes y la toma de decisiones inteligente basada en ML en el mismo nivel, es decir, en espacio de kernel y no de usuario de Linux. Esta estrategia permite ahorrar recursos en el dispositivo que realiza dicho cálculo y reduce notablemente las latencias de procesamiento que, como se ha mencionado anteriormente, es algo muy relevante en el ecosistema B5G e, incluso, de Internet de las Cosas (*Internet of Things*, IoT). Según el conocimiento de los autores, éste es el primer trabajo que presenta una implementación funcional de un modelo de red neuronal dentro del kernel de Linux con el fin de habilitar el procesamiento inteligente de tráfico.

El resto del artículo se organiza de la siguiente manera. La Sección II presenta trabajos previos que explotan tanto eBPF como algoritmos basado en ML. La Sección III discute la arquitectura del sistema, su diseño y los detalles de implementación de la solución propuesta. Finalmente, la Sección IV concluye el artículo y presenta líneas de investigación futuras.

II. ESTADO DEL ARTE

eBPF es una tecnología que permite que se ejecuten programas aislados en un contexto privilegiado dentro del espacio de kernel de Linux. De este modo, puede ampliar de forma segura las capacidades del kernel sin perder eficiencia ni requerir cambios en el código fuente del mismo. Estos programas pueden añadir funcionalidades adicionales al sistema operativo en tiempo de ejecución como si estuvieran compilados de forma nativa. Los programas eBPF pueden dirigirse a un amplio conjunto de casos de uso, pero se desarrollan principalmente para seguridad, observabilidad y funciones de red [6]. Para los últimos escenarios, suele combinarse con el *eXpress Data Path* (XDP) de Linux, lo que resulta en una herramienta potente para implementar funciones de red flexibles, eficientes y portátiles, por ejemplo en forma de función de red virtualizada (*Virtualised Network Function*, VNF).

Se espera que las NGN manejen una gran variedad de aplicaciones, tecnologías y dispositivos. Esta heterogeneidad inherente dificulta hacer frente eficazmente a los cambios en las condiciones de la red. En consecuencia, es necesario integrar funciones de gestión y control inteligentes dentro de la arquitectura de red. eBPF está ganando impulso recientemente debido a su flexibilidad y portabilidad. Por esa razón, diferentes trabajos han propuesto el uso de programas eBPF como herramientas de aplicación de decisiones tomadas por mecanismos implementados con ML que operan a nivel de aplicación. En esta línea, los autores de [7] presentan un entorno basado en ML para seleccionar y desplegar dinámicamente algoritmos de control de congestión. La solución se basa en dos módulos eBPF, uno para recopilar información sobre los flujos TCP analizados y reenviarla a un programa en espacio de usuario, y otro que implementa un algoritmo de control de congestión, que puede ser reconfigurado en tiempo de ejecución por dicho programa. Los experimentos realizados, tanto en redes emuladas como en producción, muestran su efectividad sobre soluciones más básicas. El trabajo en [8] desarrolla un modelo basado en eBPF y *Long Short-Term Memory* (LSTM) para la predicción de las tareas de red del equipo. La solución utiliza un programa eBPF para monitorizar las peticiones y respuestas HTTP en la pila de red del kernel. Estos datos se reenvían al modelo LSTM para predecir la situación de red futura. En comparación con métodos similares, la propuesta muestra una mayor precisión para realizar predicciones en tiempo real. En [9], se presenta un sistema de monitorización a nivel del kernel. Este está compuesto por un programa eBPF no intrusivo que recopila tráfico de capa de aplicación. La información recopilada se analiza mediante métodos ML y se realiza un diagnóstico del rendimiento de la red, lo que permite localizar los cuellos de botella de la red. Los autores de [10] proponen una solución para identificar y clasificar microservicios. La identificación se realiza utilizando un módulo eBPF para rastrear las llamadas al sistema. Con estos datos, una combinación de aprendizaje Bayesiano y *autoencoders* LSTM es capaz de identificar muchos microservicios del mundo real con un 99 % de precisión, con sólo un 1-2 % de uso adicional de CPU.

Como se puede ver a través de los trabajos existentes, eBPF ha surgido como una solución muy adecuada para realizar tareas de monitoreo y gestión de red dentro del kernel de Linux. Cuando se combina con algoritmos de ML, proporciona una herramienta potente para detectar automáticamente las imperfecciones en la red y actuar en consecuencia en tiempo real. Sin embargo, todavía no existe una integración de modelos ML complejos dentro del kernel de Linux utilizando las capacidades de eBPF. Solo se ha encontrado un trabajo preliminar en el que se implementa un modelo simple de árbol de decisión utilizando una serie de instrucciones *if/else* concatenadas [11]. Sin embargo, es común encontrar soluciones que utilizan programas eBPF para recopilar datos de la red y luego reenviarlos al espacio de usuario, donde los algoritmos ML se alimentan y generan las decisiones [12]. Esto implica un sobrecoste adicional, ya que el procesamiento basado en ML se saca del espacio del kernel añadiendo latencias adicionales en este proceso. Por lo tanto, según nuestro conocimiento, esta es la primera integración dentro del kernel de Linux de un algoritmo ML complejo, a saber, una red neuronal MLP, integrada en un programa eBPF para permitir un procesamiento de tráfico ultra rápido e inteligente.

III. DESARROLLO

A. GENERACIÓN DEL MODELO DE ML

Como ya se ha discutido, el objetivo final de esta propuesta es implementar una solución avanzada de procesado de tráfico, basada en ML, dentro del kernel de Linux, gracias a las posibilidades que brinda eBPF. Si bien se podrían haber considerado diferentes tipos de algoritmos de ML para este propósito, hemos optado por el uso de modelos *Multi Layer Perceptron* (MLP) debido a la flexibilidad y precisión que ofrecen las redes neuronales para analizar tráfico de datos [13]. Se mencionó anteriormente que una investigación previa ha implementado con éxito soluciones basadas en árboles de decisión en eBPF [11]. Sin embargo, en este trabajo vamos un paso más allá al implementar un algoritmo de ML mucho más complejo como MLP. En el contexto de la monitorización y gestión de red, las redes neuronales MLP han ganado atención significativa debido a su capacidad para manejar relaciones no lineales entre los vectores de entrada y las etiquetas de salida, lo cual es una característica común de los datos de tráfico de red. Además, estos modelos pueden reconocer patrones complejos en los datos, lo que los hace adecuados para detectar comportamientos anómalos sutiles que pueden no ser evidentes a primera vista. Las redes neuronales MLP pueden lograr alta precisión y velocidad, lo que las hace ideales para su uso en sistemas con requerimientos de baja latencia.

Para el desarrollo de la red neuronal, se ha utilizado una implementación de MLP basada en la librería *Scikit-Learn*[1] de Python. Hemos utilizado la función de activación ReLU (*Rectified Linear Unit*), ya que, debido a su simplicidad y eficiencia, se trata de una función de activación versátil y adecuada para varios tipos de casos de uso en redes neuronales. Particularmente, esta función es útil para tareas de clasificación binaria, como en la detección de intrusiones o de tráfico malicioso, o en la identificación de problemas de conectividad y rendimiento. Además, es una elección adecuada para el entorno eBPF, ya que es eficaz en la detección de patrones pero computacionalmente eficiente para aplicaciones que requieren un procesamiento en tiempo real. De esta forma, se evitan cálculos costosos, como las funciones exponenciales presentes en otras funciones de activación. Asimismo, esta función no requiere operaciones de punto flotante en su implementación, lo que permite su uso dentro del kernel de Linux a través de eBPF, aspecto a considerar tal y como se detalla en la próxima sección.

Finalmente, para portar el modelo MLP obtenido de *Scikit-Learn* a un programa eBPF, se ha realizado un paso intermedio, basándonos en el paradigma TinyML [14]. En concreto, se ha utilizado la librería *emlearn*[2], que transforma el modelo de Python en código ANSI C, lo que facilita notablemente la integración de la red neuronal dentro de un programa eBPF. Esta librería proporciona utilidades de conversión para generar matrices de puntos flotantes que contienen los pesos y sesgos de cada capa, así como una estructura completa que representa el modelo de red neuronal entrenado. Es necesario mencionar que no se ha nombrado el dataset empleado para el entrenamiento del modelo, ya que la presentación que hacemos de la solución propuesta en este trabajo es agnóstica del caso de uso adoptado.

B. IMPLEMENTACIÓN EBPF

La integración del modelo MLP en código C dentro de un programa eBPF no es algo trivial, dado lo estricto que es el verificador eBPF antes de permitir la carga en el kernel de un programa [6]. Como primer paso, se ha hecho uso de las cabeceras que proporciona la librería *emlearn*, que contiene las funciones que implementan la lógica de la red neuronal. En concreto, por cada una de las capas de la red neuronal, se hace una combinación lineal entre las salidas de la capa anterior con los pesos y sesgos de la actual, seguido de la aplicación de una función de activación no lineal, ReLU en este caso, tal y como se especificó anteriormente. A pesar de la "simplicidad" del código C generado por *emlearn*, no se puede integrar el modelo MLP directamente en eBPF debido a diversas restricciones propias de este entorno de trabajo [6]. Para solucionar esto, se han tenido que introducir modificaciones en la librería *emlearn* para adaptar el código generado y que sea aceptado por el verificador eBPF.

En primer lugar, los programas eBPF no pueden realizar operaciones de punto flotante. Esto es una gran limitación, ya que los pesos y sesgos de las redes neuronales son números fraccionarios. Para solventar este problema, se ha implementado una representación de punto fijo, que usa 16 bits para la parte entera, 15 bits para la parte fraccionaria, y 1 bit para el signo. Además, se han usado funciones para convertir todos los valores de punto flotante a esta representación de 32 bits. La implementación de esta solución requiere, por tanto, modificar el código de la red neuronal. En concreto, la multiplicación de los enteros de punto fijo tiene que soportar enteros de 64 bits para alcanzar la precisión necesaria. En la misma línea, también se ha adaptado la función de activación para aceptar este tipo de enteros y verificar si son positivos.

Por otro lado, el verificador de eBPF no permite el uso de bucles abiertos en el código. Por ello, se han ajustado todos los bucles para que el número de iteraciones sea conocido en tiempo de compilación. Esto se ha conseguido cambiando las variables que representan el número de capas y las salidas de la red neuronal y las longitudes de los *buffers* auxiliares por MACROS con un valor global fijo. Además, para poder aplicar la combinación lineal de pesos y sesgos en cada capa se han usado *arrays* estáticos constantes que contienen el número de entradas y salidas de cada capa. Gracias a todas estas modificaciones, la implementación ha sido diseñada para permitir el uso de redes neuronales con diferente número de capas y neuronas, lo que otorga a la solución gran poder de generalización. Sin embargo, hay que puntualizar que eBPF presenta una serie de limitaciones adicionales

[1] https://scikit-learn.org/
[2] https://github.com/emlearn/emlearn

para proteger al propio kernel, por ejemplo en el número máximo de instrucciones que puede contener un programa, la cantidad de llamadas anidadas, etc., lo cual restringe el número de capas y neuronas que se pueden usar.

Además de la red neuronal capaz de procesar tráfico, es necesario implementar un mecanismo de parseo de cabeceras para analizar el tráfico de red. Este proceso en eBPF requiere de la comprobación continua de los límites de acceso a memoria, para asegurar que siempre se hacen en espacio correctamente reservado. En este sentido, existen librerías que pueden ayudar a este recorrido de las cabeceras más conocidas, como Ethernet, IP o TCP; sin embargo, el parseo de cabeceras "menos comunes" como 6LoWPAN, GTP, etc. debe realizarse a mano.

Por último, si se desea realizar un análisis de flujos de tráfico de forma agregada en vez de paquete a paquete, después de parsear las cabeceras del paquete recibido, los datos recogidos tienen que ser almacenados de forma persistente entre ejecuciones del programa eBPF, ya que éstas son activadas por eventos (la recepción de un paquete en este caso) y no se mantiene un estado entre ejecuciones. Para ello, se pueden utilizar mapas eBPF, que permiten almacenar la información para poder aplicar el modelo MLP implementado cuando se considere, por ejemplo, cuando transcurra una ventana de tiempo de muestreo determinada. Así, al final de cada ventana de tiempo, el modelo MLP analiza los datos recogidos y devuelve las conclusiones o decisiones para las que esté entrenado.

IV. CONCLUSIONES

Este trabajo ha presentado los detalles de implementación de un mecanismo de procesado de tráfico de datos basado en ML dentro del kernel de Linux. Para ello, se ha empleado la tecnología eBPF, la cual brindas opciones avanzadas para poder introducir funciones complejas de red en dispositivos no especializados en tareas de gestión de tráfico. Pese a las limitaciones que presenta eBPF, en cuanto al número de instrucciones por programa, no compatibilidad con operaciones en aritmética de punto flotante, etc., se ha conseguido una implementación completa de un modelo complejo de redes neuronales (MLP) que es capaz de analizar el tráfico recibido por un equipo para la toma autónoma de decisiones, por ejemplo, para aplicaciones relacionadas con la Calidad del Servicio (*Quality of Service*, QoS), ciberseguridad, etc. Esta propuesta abre un camino interesante en cuanto a la implementación de funciones de red eficientes y portables que pueden ser ejecutadas en distintos puntos del continuo *fog-edge-cloud*. En este sentido, como trabajo futuro, se prevé la aplicación de los modelos desarrollados en un caso de uso específico, considerando distintos tipos de dispositivos dentro del ecosistema B5G e IoT.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación (MCIN) y la Agencia Española de Investigación (AEI), con el proyecto ONOFRE-3 (PID2020-112675RB-C44 financiado por MCIN/AEI/10.13039/ 501100011033); por la Fundación Séneca—Agencia de Ciencia y Tecnología de la Región de Murcia—con la beca FPI 21429/FPI/20, y cofinanciado por Odin Solutions S.L., España; por la Comisión Fulbright en España con la beca Fulbright 00003/FLB/21; por la Comisión Europea con los proyectos 5GASP (101016448) y NANCY (101096456); y por el Ministerio de Economia y Transformación Digital de España, con el proyecto CERBERUS-ZEUS (TSI-063000-2021-36).

REFERENCIAS

[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds," *IEEE Communications Surveys Tutorials*, vol. 25, no. 2, pp. 905–974, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10054381/

[2] C. Yeh, G. D. Jo, Y.-J. Ko, and H. K. Chung, "Perspectives on 6G wireless communications," *ICT Express*, vol. 9, no. 1, pp. 82–91, feb 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S240595952100182X

[3] 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

[4] C. Hardegen, B. Pfulb, S. Rieger, and A. Gepperth, "Predicting Network Flow Characteristics Using Deep Learning and Real-World Network Traffic," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2662–2676, dec 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9201025/

[5] J. Gallego-Madrid, R. Sanchez-Iborra, P. M. Ruiz, and A. F. Skarmeta, "Machine learning-based zero-touch network and service management: a survey," *Digital Communications and Networks*, vol. 8, no. 2, pp. 105–123, apr 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2352864821000614

[6] S. Miano, M. Bertrone, F. Risso, M. Tumolo, and M. V. Bernal, "Creating complex network services with ebpf: Experience and lessons learned," in *2018 IEEE 19th International Conference on High Performance Switching and Routing (HPSR)*, 2018, pp. 1–8.

[7] J. Zhou, X. Qiu, Z. Li, G. Tyson, Q. Li, J. Duan, and Y. Wang, "Antelope: A framework for dynamic selection of congestion control algorithms," in *2021 IEEE 29th International Conference on Network Protocols (ICNP)*, 2021, pp. 1–11.

[8] X. Zhang, Z. Liu, and J. Bai, "Linux network situation prediction model based on ebpf and lstm," in *2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, 2021, pp. 551–556.

[9] C. Liu, Z. Cai, B. Wang, Z. Tang, and J. Liu, "A protocol-independent container network observability analysis system based on ebpf," in *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)*, 2020, pp. 697–702.

[10] H. Chang, M. Kodialam, T. Lakshman, and S. Mukherjee, "Microservice fingerprinting and classification using machine learning," in *2019 IEEE 27th International Conference on Network Protocols (ICNP)*, 2019, pp. 1–11.

[11] J. F. Maximilian Bachl and T. Zseby, "A flow-based ids using machine learning in ebpf," *Arxiv*, 2022. [Online]. Available: https://arxiv.org/abs/2102.09980

[12] S.-Y. Wang and J.-C. Chang, "Design and implementation of an intrusion detection system by using extended bpf in the linux kernel," *Journal of Network and Computer Applications*, vol. 198, 2022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120624023&doi=10.1016\%2fj.jnca.2021.103283\&partnerID=40

[13] M. Abbasi, A. Shahraki, and A. Taherkordi, "Deep Learning for Network Traffic Monitoring and Analysis (NTMA): A Survey," *Computer Communications*, vol. 170, pp. 19–41, mar 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0140366421000426

[14] R. Sanchez-Iborra and A. F. Skarmeta, "TinyML-enabled frugal smart objects: Challenges and opportunities," *IEEE Circuits and Systems Magazine*, vol. 20, no. 3, pp. 4–18, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9166461/

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Mejora del balanceo de carga en redes SDN utilizando Deep Reinforcement Learning

José A. Gómez de la Hiz, 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

La flexibilidad y programabilidad de las redes definidas por software (*Software-Defined Networks*) ha permitido a la comunidad investigadora proponer nuevas técnicas de Ingeniería de Tráfico para mejorar su rendimiento. Aunque la instalación de soluciones heurísticas u óptimas en el controlador de la red permiten obtener buenos resultados, estas se basan en datos históricos que pueden no estar actualizados ante variaciones de tráfico real. Por ello, en este trabajo se propone una solución basada en *Multi-Agent Deep-Reinforcement Learning* que permite reducir *Maximum Link Utilization* de las redes definidas por software. Dicha solución permite disponer de múltiples agentes instalados en los *switches* de la red, de modo que cada agente es capaz de reaccionar y adaptarse ante los cambios localizados en su entorno. Los resultados experimentales obtenidos son prometedores. El algoritmo consigue reducir la *Maximum Link Utilization* de la red en comparación con la aplicación del algoritmo de *Dijkstra*.

***Palabras Clave*—Software Defined Network, Ingeniería de Tráfico, Machine Learning, Deep Reinforcement Learning**

I. INTRODUCCIÓN

En la actualidad, casi dos tercios de la población tiene acceso a Internet, unos veintinueve mil trescientos millones de dispositivos están conectados a la red [1]. El incremento en el número de usuarios, número de dispositivos y tráfico en Internet, hace que se deba prestar especial atención a las técnicas de Ingeniería de Tráfico, de modo que se permitan gestionar grandes volúmenes de datos de manera eficiente.

La irrupción del paradigma de las redes *Software-Defined networks* (SDN), gracias a su flexibilidad y su programabilidad, hace que la migración de redes basadas en *Intenet Protocol* (IP) a SDN sea un hecho. Las redes SDN se diferencian principalmente de las redes IP en el desacople del plano de datos y el plano de control, que pasa a estar supervisado por un elemento centralizado externo denominado controlador, que gestiona los planos de control de varios *switches* SDN [2].

El algoritmo propuesto en este trabajo se denomina *Reinforce Forward* (RF), que consiste en una solución basada en *Multi-Agent Deep-Reinforcement Learning* (MADRL) [3]. El objetivo principal de este trabajo consiste en mejorar el balanceo de carga de las redes SDN mediante la propuesta de una solución distribuida, basada en Deep-Reinforcement Learning (DRL). De entrada, lo más lógico sería pensar en utilizar algún algoritmo instalado en el controlador, que tenga una visión global de la red y que tome decisiones a partir de esta, pero esto supondría una comunicación con el controlador por cada flujo a encaminar, lo que supondría una sobrecarga del controlador producida por las peticiones de todos los *switches*, y una pérdida de tiempo por cada una de estas comunicaciones. Ante esta situación, surge la idea de que estas decisiones las tomen los propios *switches*. De esta forma se evitaría sobrecargar al controlador SDN de peticiones, reduciendo el tiempo requerido en las comunicaciones *switch*-controlador.

Se propone un algoritmo basado en MADRL, el cual despliega un *agente* en cada *switch* SDN, cuya función será elegir qué regla instalar en cada momento para encaminar los flujos teniendo una visión local de la red. Es decir, cada agente solo tendrá conocimiento de la carga de sus enlaces. Presentándose así un algoritmo basado en DRL multi-agente que permita mejorar la Ingeniería de Tráfico (IT) de las redes SDN al balancear la carga de la red a través de un encaminamiento dinámico y en tiempo real, teniendo únicamente una visión parcial de la red [4].

II. ALGORITMO *REINFORCE FORWARD*

En el ámbito del *Machine Learning* (ML) se pueden diferenciar tres campos: los métodos supervisados [5], los no supervisados [6], que se centran en clasificar y/o agrupar un conjunto de datos respectivamente, y *Reinforcement Learning* (RL) [7], que se basa en la declaración

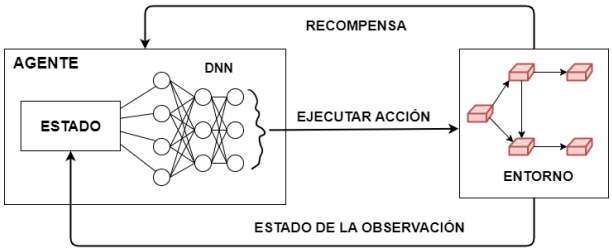


Fig. 1. Arquitectura y funcionamiento del algoritmo DRL.

de un agente que debe tomar decisiones con el objetivo de encontrar soluciones a un problema tras un entrenamiento basado en recompensas.

DRL [8] se basa en RL fusionado con Redes Neuronales Profundas (*Deep Neural Network,* DNN), lo que permite al agente aprender a actuar de la forma que le permita maximizar las recompensas obtenidas. Este algoritmo parte de cinco conceptos fundamentales. Uno de ellos es el *agente*, componente central del algoritmo, que se encarga de tomar decisiones tras observar su *entorno* y aprender al obtener una *recompensa*. El *entorno*, por otra parte, es todo aquello que rodea, observa y sobre lo que actúa el *agente*. Las *recompensas* son valores numéricos predefinidos en el algoritmo que indican al *agente* la calidad de la decisión que ha tomado mediante la diferencia que existe entre estas recompensas. Las decisiones que toma el *agente* también están predefinidas y se conocen como *acciones*. El último concepto a destacar sería el de *estado*. El *estado* es la representación del entorno en un instante de tiempo concreto, es decir, el elemento de entrada al agente (observación).

En este punto, podríamos definir el funcionamiento del algoritmo DRL de la siguiente manera: un agente aprende a tomar decisiones sobre qué acción tomar ante diferentes estados de un entorno, con el objetivo de maximizar las recompensas. Para ello (ver Figura 1), el agente observa el entorno, ejecuta la acción más recomendada tras introducir la observación (estado) por su DNN, provocando esta acción un cambio en el entorno que será evaluado y a la cual se le otorgará una recompensa que será la que ajustará los pesos de la DNN del agente.

El algoritmo RF, como ya se ha comentado, se basa en DRL. Pero, ¿tiene sentido implementar un *agente* DRL en un controlador?, ¿se trata de la mejor decisión? Instalar un *agente* DRL en el controlador SDN puede derivar en buenos resultados. Sin embargo, al hacerlo de una forma centralizada podría ser más conveniente utilizar heurísticas tradicionales o soluciones óptimas basadas en programación lineal que se traducen en mejores (o más rápidas) soluciones. Por ello se propone un algoritmo basado en MADRL, en el que se instala un *agente* en cada nodo SDN de la red, pasando cada uno de estos nodos de tener una visión global de la red, como el controlador SDN, a tener una visión limitada de ella como veremos en la definición del estado de estos agentes posteriormente.

A continuación se detalla el diseño de los elementos principales de los *agentes* DRL utilizados en este algoritmo (ver Figura 1):

- *Entorno*. El entorno en este algoritmo se compone de los enlaces que conectan los nodos SDN de la red. En concreto, representa la carga de cada uno de los enlaces. En el algoritmo se representa como una matriz de $2 * N$ elementos, donde N hace referencia al número de nodos. Sobre esta matriz, los agentes realizarán sus observaciones, procesarán los datos y lo enviarán a la DNN para obtener una predicción de cada actuación posible.
- *Estado*. Se trata de una representación del entorno en un instante de tiempo. Esta es enviada hacia la DNN para que estime cuál es la mejor forma de actuar en ese momento. En el algoritmo RF, el estado se declara como un *array* de dos dimensiones de tamaño N, siendo N el número de nodos de la red a la que se quiere aplicar el algoritmo. La primera dimensión contiene el porcentaje de carga de cada uno de sus enlaces salientes y un valor igual a 999 en las casillas de los nodos con los que ese agente no tiene conectividad. En la segunda dimensión se encuentra el tamaño del flujo a encaminar (en la primera posición) y el valor 999 en las demás posiciones. Por tanto, solamente se pueden variar los valores distintos de 999.
- *Acciones*. El encaminamiento en las redes SDN se realiza mediante la instalación de reglas de flujo en los *switches* SDN, por lo que si se procura mejorar el balanceo de carga en estas redes, las acciones se pueden definir como diferentes reglas de flujo, entre las cuales el agente deberá elegir que regla instalar en el *switch* en el que se está ejecutando, en función del estado de su entorno.
  Inicialmente, en este algoritmo se definen cuatro acciones distintas:
  1) Encaminar según el camino más corto (*Dijkstra*).
  2) Encaminar según el enlace menos cargado.
  3) Encaminar equitativamente según los dos enlaces menos cargados.
  4) Encaminar equitativamente según los tres enlaces menos cargados.
- *Recompensas*. Para que el algoritmo cumpla con el objetivo, deben definirse correctamente las recompensas que retroalimentan a los *agentes*. Pueden implementarse varias distribuciones de recompensas que produzcan buenos resultados. Cada una de ellas se han definido siguiendo una serie de reglas con prioridad, las cuales suman o restan más puntuación mientras más alta sea la prioridad de esta regla. A continuación se definen las reglas que definen los *rewards* de este algoritmo junto con los valores con los que se obtienen los mejores resultados:
  1) Encaminar el flujo por un enlace válido. Si el *agente* consigue encaminar el flujo de paquetes por un enlace que pertenezca a uno de los

caminos que pueden llegar al nodo destino y además el flujo de paquetes tiene un tamaño menor a la capacidad residual de ese enlace, entonces, puede decirse que ha encaminado el flujo por un enlace válido. De ser así se pasa a valorar la siguiente regla. De lo contrario, se le otorga una recompensa muy negativa, con un valor igual a -4000, para indicarle al *agente* que en ese estado esa acción no es la más adecuada.

2) Disminuir la *Maximum Link Utilization* (MLU) de los enlaces salientes al nodo en el que está instalado el *agente*. Esto permite hacer ver al *agente* que para maximizar la recompensa obtenida debe minimizar el porcentaje de carga máxima de entre todos sus enlaces salientes. Para indicarle esto, se le ofrece una recompensa con un valor directamente proporcional al enlace saliente del nodo más cargado. En el algoritmo RF, esto se refleja estableciendo 5 valores distintos de recompensa, en función de si el porcentaje de carga del enlace más cargado supera el 35%, 50%, 65%, 80% o 95%. En el caso de que no se supere ninguno de estos porcentajes, se establece un valor de recompensa igual a 1.000. Sin embargo, en el caso de que se supere alguno de estos umbrales, el valor de recompensa inicial es igual a 500, al cual se le va restando 100 puntos de recompensa por cada uno de los umbrales que se superen.
3) Encaminar el flujo por el enlace que pueda llegar al destino con el menor número de saltos posibles. Esto es vital para no sobrecargar un mayor número de enlaces, que supondría un aumento de la media de carga global de la red, como se puede observar en las Figuras 2 y 3. Ambas se apoyan en una topología de 3 nodos en la que se necesita encaminar 2 flujos desde el nodo A al nodo B, varían las métricas en función de la acción escogida por el *agente*. En la Figura 2 se puede observar el resultado de encaminar los flujos siguiendo el algoritmo de *Dijkstra*, donde se obtiene una MLU igual al 90%. Encaminar por el enlace menos cargado supone disminuir la MLU un 30% (Figura 3). Sin embargo, la carga media de todos los enlaces asciende un 10%, porcentaje que puede aumentar considerablemente conforme ascienden las dimensiones de la red. Esto supone instalar las reglas de flujo distintas a *Dijkstra* el menor número de ocasiones posibles, solo cuando pueda llegar a mejorar las dos reglas de los *rewards* anteriores.

Definidos los elementos principales de MADRL, se describe la comunicación entre los diferentes *agentes*. Esta comunicación se produce a través de la representación del entorno, una matriz de $N * N$ (matriz de cargas) siendo N el número de nodos de la red. A cada *agente*, cuando tenga que actuar, se le envía una observación procesada

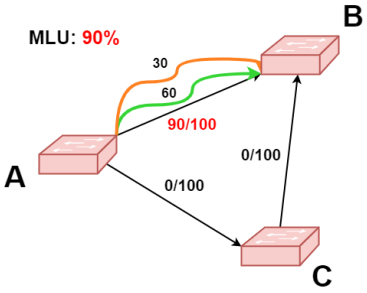


Fig. 2. Solución según camino más corto.

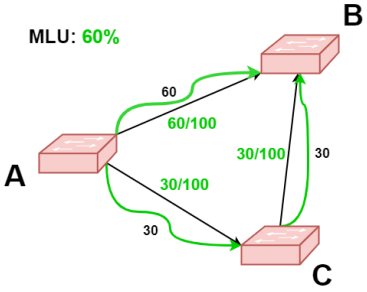


Fig. 3. Solución incluyendo la acción de enlace menos cargado.

de esta matriz y tras elegir de qué forma encaminar el flujo, modificará la matriz en función de esta decisión. El procesamiento de la matriz de cargas se basa en filtrar únicamente la fila de la matriz que contenga la carga del propio nodo y añadirle la segunda dimensión con el tamaño del flujo a encaminar.

III. Resultados experimentales

Previamente a la ejecución de las pruebas, el algoritmo RF necesita de un entrenamiento previo para que los *agentes* regulen los pesos de sus DNN y tomen decisiones acordes a lo establecido a través de las recompensas. Este entrenamiento se hace de forma independiente para cada uno de los *agentes*, realizando cada uno de ellos, unas 150.000 iteraciones con representaciones de estados reales creados de forma pseudoaleatoria, con una duración de 1 hora y 20 minutos por cada uno de los agentes. Estos tiempos varían en función de la topología a la que se quiere aplicar el algoritmo y la máquina en la que se ejecute.

El algoritmo se ha implementado utilizando el lenguaje de programación *Python*, haciendo uso de multitud de librerías, de entre las cuales cabe destacar *TensorFlow* [9], que facilita la implementación de la DNN y el uso de la GPU de la máquina para la ejecución del algoritmo, y *Keras-Rl* [10], basada en *TensorFlow*. Además, se ha utilizado *Keras* [11], que ofrece varios algoritmos de RL.

La DNN utilizada se compone de 7 capas, de 48, 56, 48, 38, 28, 18 y 4 (número de acciones) neuronas respectivamente. Todas ellas tienen como función de activación *Linear*.

Las pruebas se han realizado sobre una topología real obtenida de la librería SDNLib, *Abilene* (Figura 4), que se compone de 12 nodos y 30 enlaces unidireccionales de diferentes anchos de banda, según los datos recuperados

Fig. 4. Topología Abilene. Fuente: SndLib

de [12]. Se ejecuta RF sobre esta topología por medio de 5 pruebas con matrices de tráfico distintas y de carga creciente, siendo la matriz de tráfico con identificador 4 aquella que, realizando un encaminamiento basado en Dijkstra, obtiene una MLU cercana al 100%.

Se implementan dos soluciones para cada matriz de tráfico, una a través del algoritmo RF y otra a través del algoritmo de encaminamiento por *Dijkstra*. En la Figura 5 se puede observar cómo el algoritmo RF, configurado como se comentó en la Sección II, produce una mejora sobre el encaminamiento basado en *Dijkstra*, en términos de reducción de MLU. Sin embargo, la solución obtenida impacta en la carga media de los enlaces de la red, tal y como se puede observar en la Figura 6.

IV. CONCLUSIONES

Este trabajo tiene como objetivo el diseño e implementación de un algoritmo que mejorase la IT de las redes SDN al disminuir el porcentaje de carga máximo de los enlaces de la red o MLU. Mediante la propuesta de un algoritmo (RF) basado en MADRL, que instalando un *agente* DRL en cada *switch* SDN, conseguía tomar decisiones de enrutamiento en tiempo real, en función del estado de los enlaces locales al *switch*. Una vez realizadas las pruebas experimentales sobre Abilene, se comprueba que el algoritmo consigue disminuir entre un 2% y un 6% la MLU de la red. Sin embargo, la carga media de los enlaces se ve ligeramente incrementada. Como trabajo en proceso, se espera mejorar la definición del modelo DRL para que obtenga buenos resultados sobre redes más grandes.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por MCIN/AEI/10. 13039/501100011033 y por la Unión Europea ”Next GenerationEU /PRTR”, por el Ministerio de Ciencia, Innovación y Universidades (proyectos TED2021-130913B-I00, PDC2022-133465-I00), por el proyecto PID2021-124054OB-C31 y la subvención CAS21/00057 (MCI/AEI/FEDER, UE), y por la Consejería de Economía, Ciencia y Agenda Digital de la Junta de Extremadura (GR21133).

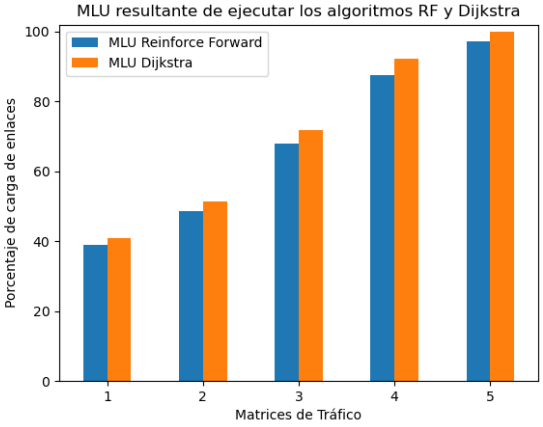


Fig. 5. Valores de MLU obtenidos por RF y por *Dijkstra*.

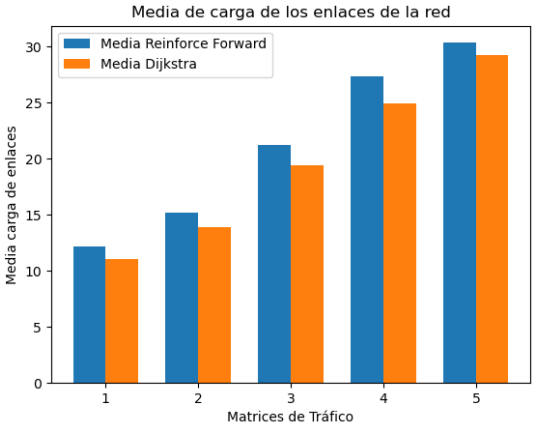


Fig. 6. Carga media de los enlaces tras aplicar RF y *Dijkstra*.

REFERENCIAS

[1] Cisco, “Cisco annual internet report - cisco annual internet report (2018–2023) white paper - cisco,” [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.

[2] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, “A survey on software-defined networking,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 27–51, 2015.

[3] S. Gronauer and K. Diepold, “Multi-agent deep reinforcement learning: a survey,” *Artificial Intelligence Review*, pp. 1–49, 2022.

[4] I. F. Akyildiz, A. Lee, P. Wang, M. Luo, and W. Chou, “A roadmap for traffic engineering in sdn-openflow networks,” *Computer Networks*, vol. 71, pp. 1–30, 2014.

[5] S. B. Kotsiantis, “Supervised machine learning: A review of classification techniques,” in *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies*. NLD: IOS Press, 2007, p. 3–24.

[6] I. C. Pérez Verona and L. Arco García, “Una revisión sobre aprendizaje no supervisado de más de distancia,” *Revista Cubana de Ciencias Informáticas*, vol. 10, pp. 43 – 67, 12 2016.

[7] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *J. Artif. Int. Res.*, vol. 4, no. 1, p. 237–285, may 1996.

[8] Y. Li, “Deep reinforcement learning: An overview,” *arXiv preprint arXiv:1701.07274*, 2017.

[9] “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[10] M. Plappert, “keras-rl,” https://github.com/keras-rl/keras-rl, 2016.

[11] F. Chollet *et al.* (2015) Keras. [Online]. Available: https://github.com/fchollet/keras

[12] S. Orlowski, M. Pioro, A. Tomaszewski, and R. Wessäly, “Sndlib 1.0 - survivable network design library,” *Networks*, vol. 55, no. 3, pp. 276–286, 2009.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Preliminary approaches towards the integration of TSN communications into the NFV architectural framework

Jorge Sasiain, Asier Atutxa, David Franco, Jasone Astorga, Eduardo Jacob

Department of Communications Engineering, University of the Basque Country (UPV/EHU). 48013 Bilbao, Spain.

jorge.sasiain@ehu.eus, asier.atutxa@ehu.eus, david.franco@ehu.eus, jasone.astorga@ehu.eus, eduardo.jacob@ehu.eus

This paper presents a preliminary architecture for the integration of Time-Sensitive Networking (TSN) communications into the Network Functions Virtualization (NFV) architectural framework. Synergies between functional blocks and constructs of NFV, and components of TSN networks, are investigated in order to arrive at an integrated architecture. Additionally, mechanisms and configuration procedures to enable TSN-compliant, real-time, and virtualized end stations under the NFV framework are explored.

***Keywords*—Time-Sensitive Networking, Network Functions Virtualization, Software-Defined Networking**

I. INTRODUCTION AND MOTIVATION

5G technology is undoubtedly having a significant impact on the industrial manufacturing and automation vertical, being, in fact, a key driver of the fourth industrial revolution (Industry 4.0). Industry 4.0 embraces a series of technologies and paradigms primarily focused around ubiquitous communication scenarios, which enables innovative use cases. In order to materialize all these applications, there is a clear trend towards the virtualization and cloudification of their functionalities [1], encompassing entities and technologies such as Programmable Logic Controllers (PLCs), robotics, IIoT gateways, Cyber-Physical Systems (CPSs), automotive systems, and digital twins. Virtualization increases flexibility and agility, facilitates offloading of complex processing to an edge cloud, and enables coexistence of mixed-criticality applications.

Industrial scenarios often involve closed-loop control systems with stringent requirements of determinism, latency, and reliability that standard Ethernet cannot satisfy. IEEE Time-Sensitive Networking (TSN) standards have been developed to bridge this gap forgoing traditional proprietary fieldbus protocols that lack interoperability. However, challenges arise when trying to incorporate virtualization technologies into TSN, at both performance and configuration management level. The literature primarily focuses on the former aspect, proposing mechanisms for real-time containers and Virtual Machines (VMs) such as real-time task scheduling policies [2], preemptable kernel and co-kernel approaches [3], and specific hypervisor architectures [4]. However, works addressing the orchestration aspect are scarce and focus mainly on specific modules of container platforms like Kubernetes [5].

To the best of our knowledge, initiatives towards a full-fledged integration of time-sensitive communications with virtualization and orchestration architectures is missing. Thus, this paper sketches an initial approach for integrating TSN and ETSI Network Functions Virtualization (NFV), the latter considered a building block of 5G networks [6].

II. BACKGROUND

IEEE TSN comprises a toolbox of standards which provide functionalities spanning time synchronization, bounded low latency, high reliability, and resource management. A TSN *stream* is an unidirectional flow of data between *end stations*, i.e. from a *Talker* to one or more *Listeners*, that traverses TSN *bridges*, and that belongs to a traffic class identified by the VLAN Priority Code Point (PCP). The 802.1Qcc standard specifies three types of configuration modes for a TSN *domain*. In the fully centralized one, the Centralized Network Configuration (CNC) entity retrieves bridge capabilities and configures them, while the Centralized User Configuration (CUC) entity retrieves application stream requirements and discovers and configures end station capabilities. CUC sends stream requirements to the CNC via the User/Network Interface (UNI). The separation of the network into TSN domains is an administrative decision (e.g. by production lines, machine units, hierarchical network segments, etc.). The primary TSN standard to achieve bounded low latency for periodic traffic is 802.1Qbv, which provides temporal isolation by computing a network-wide schedule where each traffic class is assigned a dedicated time slot enforced by egress device ports. 802.1Qbv is supported by the capabilities provided by the 802.1AS standard to synchronize the clocks of all TSN devices. Multiple TSN domains may share a common working clock domain.

The ETSI NFV architectural framework is concerned with the abstraction, or decoupling, of network functions

from physical hardware. The NFV Infrastructure (NFVI) contains and abstracts physical resources for use by Virtual Network Functions (VNFs), which can be composed of multiple sub-components —VMs or containers— to encapsulate a specific networking functionality. VNFs can be chained together to form Network Services (NS). Virtual Links (VLs) interconnect VNF sub-components and VNFs. The Management and Orchestration (MANO) block is subdivided into the Virtualized Infrastructure Manager (VIM), responsible for the management of NFVI resources, the VNF Manager (VNFM), responsible for the lifecycle management of VNFs, and the NFV Orchestrator (NFVO), which offers end-to-end NS orchestration and can interact with OSS systems. Additionally, a WAN Infrastructure Manager (WIM) can provide multi-site connectivity management between NFVI Points of Presence (NFVI-PoPs). NFV and Software-Defined Networking (SDN) are widely regarded as complementary technologies, and NFV standards explicitly address the interworking with SDN.

The remainder of this paper assumes that the fully centralized TSN configuration model is used and that, if communications traverse multiple TSN domains, they share the time synchronization domain. Additionally, the article focuses on enabling support for the 802.1Qbv standard. The proposal is contextualized in the industrial automation and manufacturing vertical under the Industry 4.0 umbrella. In this scenario, time-sensitive applications could be allocated in heterogeneous platforms, such as multi-processor SoCs or FPGAs with embedded Operating Systems that connect to field devices in production lines or robots, as well as to traditional edge cloud resources that provide additional computational capabilities.

III. ARCHITECTURE-LEVEL INTEGRATION

This section proposes an architecture-level integration between NFV functional blocks and TSN components to enable time-sensitive communications involving applications deployed under the NFV framework using hypervisor-based or container-based virtualization. Two integration modes are considered to address two different communication use cases: (a) integration to enable intra-VIM TSN communications; and (b) integration to enable inter-VIM TSN communications. Use case (a) involves a single TSN domain —mapped 1-to-1 to a NFVI-PoP—, for communications between field devices, local PLCs, local computing, etc., while use case (b) allows to span multiple TSN domains for communications between production lines, to a centralized edge cloud, etc. A high level view of the proposed architecture, which is described in the remainder of this section, is shown in Figure 1.

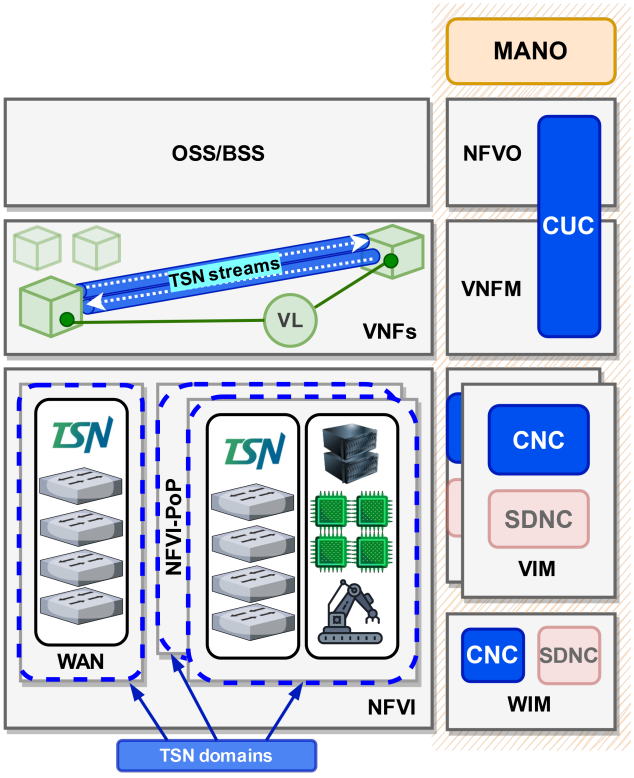


Fig. 1. Proposed architecture for integrating TSN support into the NFV architectural framework. An orange background denotes the NFV MANO functional blocks (right). TSN components and entities are depicted in blue. A simple NS is overlaid in green (left).

A. Intra-VIM Communication Use Case

The intra-VIM communication use case involves the communication between VNFs and their underlying resources located within a single NFVI-PoP. The NFVI incorporates TSN bridges apart from the computing resources. As the VIM is the functional block ultimately responsible for the interaction with the NFVI, the CNC is placed at VIM level, where three implementation possibilities are considered: (a) CNC functionality running in parallel to VIM functionality in a coordinated manner; (b) CNC functionality being implemented as part of the VIM; and (c) CNC functionality being implemented as part of, in parallel to, or on top of an SDN Controller (SDNC) (i.e. as an SDN application) that is located in the VIM. In (c), the CNC can inherit functionalities such as topology discovery from the SDNC.

The user-facing CUC functionality is provided by the NFVO functional block of NFV MANO. The NFVO receives NS lifecycle requests that, in this case, contain additional configuration that targets the TSN resources. Information that concerns the TSN bridges in the NFVI is sent to the CNC located at VIM level. Hence, information exchanged through the UNI takes place between the NFVO and the VIM, which translates to the Or-Vi reference point in the NFV architectural framework. On the other hand, CUC functionality related to configuring end stations and retrieving their capabilities is implemented in the VNFM via the Ve-Vnfm interface towards the VNFs.

B. Inter-VIM Communication Use Case

The inter-VIM communication use case involves the communication between VNFs and their underlying resources located in separate NFVI-PoPs. On the one hand, each NFVI-PoP requires the support discussed in the intra-VIM case for the portions of the TSN streams that involve resources in the extent of that NFVI-PoP.

Table I
PROPOSED EQUIVALENCE (BLUE) AND FUNCTIONAL MAPPING (GREEN) BETWEEN TSN AND NFV ENTITIES AND CONSTRUCTS

TSN	Domain	End station	Stream	CUC	CNC	UNI
NFV	NFVI-PoP or WAN segment	VM, container, or single-component VNF	One direction of VLs or of an NS	NFVO and VNFM	VIM or WIM	Or-Vi or Or-Wi

On the other hand, additional TSN resources are also found in the WAN or transport network between NFVI-PoPs, which is managed by the WIM. Two WIM types are considered: (a) standalone WIM; and (b) WIM being itself an SDNC or an SDN application running on top of a SDNC. In case (a), CNC functionality can be either embedded in the WIM, or be implemented in parallel or on top of it in a coordinated manner. In case (b), CNC functionality can be implemented as part of, in parallel to, or on top of the SDNC. Like before, the NFVO/VNFM pair is in charge of executing CUC functionality, and the UNI exists between the NFVO and the CNC, which this time is at WIM level. This corresponds to the Or-Wi reference point in the NFV architectural framework.

So far this scenario has assumed that there is a number of TSN domains equal to the number of NFVI-PoPs plus WAN segments, yet there is only a single location housing CUC functionality —the NFVO/VNFM— for all domains. On the one hand, NFV standards do not constrain the management scope of a VIM to one or to all NFVI-PoPs in the NFV architecture. Hence, the considered approach is to have one VIM manage each NFVI-PoP in order to match the scope of each TSN domain. On the other hand, the TSN 802.1Qcc standard specifies each TSN domain to be managed by its own CNC and CUC. However, as the NFVO —and thus CUC— is expected to receive service requests from a centralized OSS, stream requirements do not need to be provided directly by end stations.

Depending on the architecture of the industrial network at hand, it might be possible that no central management segment exists in it to deploy a NFVO/VNFM pair with direct management access to all involved TSN domains. In this case, where each separate MANO controls a subset of all TSN domains, it becomes effectively a concatenation of intra-VIM communications. TSN inter-domain coordination and configuration mechanisms would be required to realize end-to-end streams, but this would be considered outside of the scope of NFV. On the contrary, if the NFVO/VNFM pair has access to management connectivity for all involved TSN domains, a centralized CUC in the NFVO is used, relying on NFV MANO procedures to derive the target VIM —and thus the target CNC— of each service component and involved TSN stream.

C. Communication with External Systems

It has to be taken into account that streams may not only span end stations that reside inside the NFVI, for example in communications in which industrial sensors, actuators, and/or legacy control systems participate. If the functional lifecycle of such external components can be orchestrated by NFV MANO procedures, they can become part of the NFVI and their behavior modelled as a Physical Network Function (PNF) under MANO's visibility. This would result in a scenario equivalent to the intra-VIM communication one. On the contrary, if the external components cannot be exposed to orchestration by MANO, NFV is not responsible for any resources located outside the NFVI with regards to TSN streams that cross these boundaries.

IV. PLATFORM AND VNF CONFIGURATION

In addition to the coordinated orchestration of TSN and NFV resources, the other main challenge is that those resources have to satisfy the performance requirements of time-sensitive communications in spite of the virtualization layers. These concerns can be grouped into: minimization of latency and jitter caused by virtualization overhead and resource sharing; time synchronization for VMs/containers; and time-triggered transmission scheduling for VM/container streams. Other requirements are network topology and bridge capabilities discovery, the determination of bridge and propagation delays, and 802.1Qbv schedule synthesis, which are inherent to TSN networks and thus outside of this paper's scope. This section proposes how to manage and carry out the configuration required for TSN support in the NFV environment.

A. Mechanisms for Time-sensitive Communications with Virtualization

The toolbox of mechanisms that can contribute to providing real-time guarantees and TSN compliance to end stations —with and without virtualization— includes:

- TSN features provided by the Linux ecosystem [7]. This includes Linux PTP for 802.1AS time synchronization in the NIC and system clocks, and Traffic Control Queuing Disciplines (Qdiscs) to implement traffic shaping functionality like 802.1Qbv.
- High-priority, real-time scheduling policies such as SCHED_DEADLINE in Linux, which incorporates awareness of the deadline of the tasks.
- Real-time operating systems (RTOS).
- Real-time co-kernels that run in parallel to the main kernel and are responsible for time-critical workloads.
- The PREEMPT_RT patch for the Linux kernel, which increases preemptability of kernel code.
- Hypervisors with real-time capabilities.
- Hardware resources isolation for VMs/containers, e.g. CPU pinning and SR-IOV passthrough.

The activation of these mechanisms involves hardware and/or software requirements. Notably, hypervisor-based virtualization without resource pinning requires the hypervisor to take charge of VM transmissions according to a global schedule, e.g. dispatching vCPUs appropriately [8].

B. Stream Requirements Retrieval Procedure

Whenever a NS lifecycle operation is triggered, the CUC at NFVO level shall provide TSN configuration information to the CNC at VIM or WIM level via the UNI. Multiple CNCs in separate VIMs and/or WIMs may be targeted depending on the placement of the NS components. This information is derived from the user-defined NFV service descriptors, which shall be augmented to accommodate TSN stream requirements. Based on 802.1Qcc specifications, the required information per new stream during an NS instantiation procedure includes:

a. Identification of the end stations and network interfaces participating in the stream.
b. Data frame specification (MAC and IP addresses of participating end stations, and VLAN tag).
c. 802.1Qbv traffic specification, such as period, maximum frame size, frames transmitted in a period, and maximum allowed latency.

The information in c. cannot be derived through native NFV descriptors and procedures. Therefore, additional information shall be provided within the corresponding VL constructs in a VNF or NS descriptor, as appropriate. It has to be taken into account that TSN streams are unidirectional, but VLs are bidirectional. For example, an initial design approach may assume a single Listener per TSN stream and consider a constraint that a NS must map exactly to two TSN streams in the two directions of a communication between two end stations.

C. Configuration Procedure of End Stations

After receiving and processing stream requirements, the CNC configures the TSN bridges and sends the result back to the CUC, which concludes by forwarding configuration directives to the end stations. Although several real-time capabilities such as the use of a real-time operating system, kernel, or hypervisor, are platform-wide, specific per-stream VM/container configuration also has to be carried out by the CUC. VNF configuration is conducted by the VNFM via the Ve-Vnfm reference point, so this CUC functionality is allocated to the VNFM functional block. This configuration includes:

- Installation and execution of any required software, such as daemons for 802.1AS time synchronization.
- VLAN configuration of the network interfaces.
- Setting of internal Linux packet priorities (socket option SO_PRIORITY), and mapping to VLAN PCPs.
- Setting of process scheduling policies and priorities.
- Configuration of TSN traffic classes via TSN Qdiscs, such as the Time Aware Priority (TAPRIO) and the Earliest TxTime First (ETF) Qdiscs for 802.1Qbv.

Adhering to the timings specified in the stream requirements is the responsibility of the application running in each VM/container. Because the configuration described above is applied to each VM/container, no specific action is required in NFVI resources during an NS termination procedure (only by the CNC). On the other hand, NS or VNF update actions may involve reconstruction of streams and thus reconfiguration of VMs/containers.

V. CONCLUSIONS AND OUTLINE

This paper proposes guidelines towards a preliminary approach for supporting TSN in NFV. However, there is considerable room for delving further into integration details and exploring alternative methods and technologies.

1) Container-based NFV architectures: following the cloud-native paradigm shift, NFV Rel. 4 has introduced explicit support for containerized applications. Platforms like Kubernetes are gaining traction as orchestration tools. An integration of TSN into Kubernetes could be explored, leveraging its Container Network Interface framework, and Kubernetes Operators to extend its orchestration scope.

2) Kernel-bypass packet processing technologies: Many research works point to the use of kernel-bypass and userspace technologies such as XDP and DPDK to improve packet processing performance in the context of TSN, especially when virtualization is involved. Support for the acceleration of VNF packet processing through these technologies could bring performance benefits.

3) Integration of additional TSN standards: in addition to 802.1Qbv, the integration with other latency standards like 802.1Qbu and 802.1Qch, and reliability standards like 802.1CB and 802.1Qci, could be considered.

4) Service Function Chaining: NFV supports other constructs based on forwarding graphs, where the communication paths between VNFs are fixed, often leveraging SDN. Their mapping to TSN streams could be investigated.

5) Wireless TSN: a significant effort is currently being carried out by researchers and SDOs alike to enable wireless TSN communications, e.g. in 3GPP Release 16 onwards. 5G-TSN integration is a research line of its own; still, significant complementary synergies could be found.

ACKNOWLEDGEMENTS

This work was supported in part by the Spanish Ministry of Science and Innovation through the national project (PID2019-108713RB-C54) titled "Towards zeRo toUch nEtwork and services for beyond 5G" (TRUE-5G), and in part by the Basque Government through the project Social Network for Machines (SONETO) (KK-2023/00038) of the ELKARTEK Program.

REFERENCES

[1] R. Sabella, A. Thuelig, M. Carrozza, and M. Ippolito, "Industrial automation enabled by robotics, machine intelligence and 5G," *Ericsson Technology Review*, no. 2, 2018.
[2] L. Abeni, A. Balsini, and T. Cucinotta, "Container-based real-time scheduling in the Linux kernel," *ACM SIGBED Review*, vol. 16, no. 3, pp. 33–38, 2019.
[3] T. Tasci, J. Melcher, and A. Verl, "A container-based architecture for real-time control applications," in *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE, 2018, pp. 1–9.
[4] L. Leonardi, L. L. Bello, and G. Patti, "Towards time-sensitive networking in heterogeneous platforms with virtualization," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1. IEEE, 2020, pp. 1155–1158.
[5] A. Garbugli, L. Rosa, A. Bujari, and L. Foschini, "KuberneTSN: a deterministic overlay network for time-sensitive containerized environments," *arXiv preprint arXiv:2302.08398*, 2023.
[6] ETSI NFV ISG, "Network operator perspectives on NFV priorities for 5G," 2017.
[7] "TSN documentation project for Linux*." [Online]. Available: https://tsn.readthedocs.io/
[8] J. Ruh, W. Steiner, and G. Fohler, "Clock synchronization in virtualized distributed real-time systems using IEEE 802.1 AS and ACRN," *IEEE Access*, vol. 9, pp. 126 075–126 094, 2021.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Despliegue automatizado de red inalámbrica virtualizada

María Canales Compés, Julián Fernández Navajas, Carmen Arasanz Jordán, José Ruiz Mas, Ángela Hernández Solana, José Ramón Gállego Martínez
Departamento de Ingeniería Electrónica y Comunicaciones
Universidad de Zaragoza
Edificio Ada Byron, C/ María de Luna,1. 50018, Zaragoza.
mcanales@unizar.es, navajas@unizar.es, 779838@unizar.es, jruiz@unizar.es, anhersol@unizar.es, jrgalleg@unizar.es

El incremento del uso de smartphones y dispositivos de vídeo y la ubicuidad de los usuarios han incrementado el tráfico de datos en redes inalámbricas demandando mayor número de puntos de acceso y una coordinación más efectiva que garantice una conectividad de calidad, requiriendo arquitecturas de red robustas, flexibles y escalables. Siguiendo la corriente de desarrollo de las redes definidas por software y la virtualización de funciones de red, se propone automatizar el despliegue de una red inalámbrica coordinada, propuesta previamente en el contexto del grupo investigador. La inteligencia centralizada, que desacopla el plano de datos y el de control, permite un diseño, dimensionado y optimización independientes facilitando a su vez la integración en un escenario más amplio de conectividad global. Este trabajo pone el punto de partida, dotando de mayor flexibilidad y eficiencia a la arquitectura con un despliegue y orquestación de sus funciones virtuales mediante Kubernetes.

***Palabras Clave*- SDN, SDWN, Wi-Fi, Kubernetes, Docker.**

I. Introducción

Actualmente, las redes inalámbricas se encuentran en un periodo de gran expansión, debido a su bajo coste, su facilidad en el despliegue y a la libertad de movimiento que proporcionan. Wi-Fi es la tecnología preferida para escenarios interiores por su fácil acceso y las facilidades que ofrece para la implementación de redes locales privadas en términos de modelo comercial, lo que ha favorecido la consolidación del estándar IEEE 802.11 y su predominio en el sector. Así, su integración en el contexto actual de desarrollo de redes 5G y 6G centra un gran interés en el estudio y evolución hacia soluciones con una mayor coordinación y flexibilidad de despliegue. Tal como lo promueven los estándares 5G, SDN (Software Defined Network) y NFV (Network Function Virtualization) son dos de los habilitadores tecnológicos más prometedores en este escenario. SDN facilita un nuevo nivel de coordinación en las redes Wi-Fi (Software-Defined Wireless Networks, SDWN). SDWN permite monitorizar el entorno inalámbrico y gestionar en consecuencia los puntos de acceso programables, facilitando la implementación de algoritmos inteligentes, gracias a la administración por parte de controladores SDN centrales. En cuanto a NFV, la virtualización de las funciones, mediante máquinas virtuales o contenedores, incrementa la flexibilidad de la red facilitando la personalización de la infraestructura y la adecuación al dinamismo de los usuarios y las demandas de servicio. En este contexto, la inclusión de plataformas en la nube, como Kubernetes, para orquestar la infraestructura virtualizada y su correspondiente gestión, requiere un análisis detallado de las particularidades de la red inalámbrica: dependencia de la tecnología de radio, limitaciones de hardware, intermitencia de conectividad y recursos radio específicos. Estas particularidades pueden requerir una infraestructura virtual específica y controladores adaptados.

Bajo este paraguas, se desarrolló una solución SDWN propia [1] sobre la que este trabajo pretende estudiar la virtualización de las funcionalidades implementadas y flexibilizar la arquitectura permitiendo un despliegue automatizado de la red mediante orquestación. A continuación, se describe el progreso en la investigación mostrando el estudio de referencia y la prueba de concepto preliminar con la que se establecen las bases de una infraestructura de red completa, de mayor envergadura, adaptable al diseño de nuevos algoritmos de coordinación e inteligencia de red y orquestada en un contexto de conectividad heterogéneo aplicable en múltiples casos de uso.

II. Punto de partida y motivación

A. *Escenario preliminar: SDWN no virtualizada*

En el contexto de un proyecto de investigación ya consolidado [1], se propuso una nueva arquitectura de red inalámbrica definida por software (SDWN) que integra mecanismos de coordinación para mejorar las capacidades de un conjunto de puntos de acceso (AP) Wi-Fi administrados centralmente [2], [3]. Como muestra la Fig. 1, dicha arquitectura incluye herramientas de monitorización y otras

funcionalidades, proporcionando inteligencia pese a utilizar puntos de acceso comerciales de bajo coste, gracias a la utilización de un controlador central que, con toda la información disponible, es capaz de tomar decisiones inteligentes sobre la asignación de los recursos de la red. Por ejemplo, la implementación de soluciones de traspaso (*handover*) proactivo, permite menores latencias e integrar funcionalidades como el balanceo de carga que, en última instancia, facilitan una mayor escalabilidad y eficiencia en la utilización de los recursos. Mediante la incorporación de la abstracción *Light Virtual Access Point* (LVAP), propuesta por primera vez en [4], el controlador asigna a cada usuario el equivalente a un "AP virtual" (dirección MAC y SSID del AP para el cliente, MAC real de la STA y dirección IP exclusiva) que "acompaña" a este en su movilidad entre AP físicos de manera transparente para el terminal de usuario (no se requieren modificaciones en las estaciones, que ejecutan el estándar 802.11). Un AP físico, por lo tanto, alojará tantos LVAP como clientes tenga conectados, desvinculando las funciones de control del envío físico de los datos. Este nuevo plano de control requiere de la implementación de funcionalidades específicas, asociadas a la gestión inalámbrica, más allá del estándar de redes SDN (como OpenFlow), por lo que la arquitectura incluye, no solo la implementación específica del software del AP, sino la integración de un controlador propio y su protocolo de comunicación (Odin [3]).

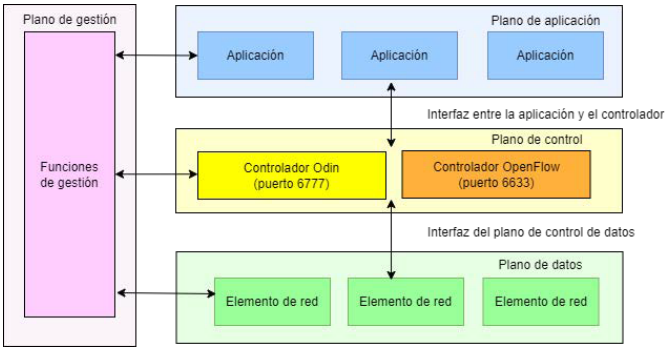


Fig. 1. Red definida por software para Wi-Fi (SDWN)

B. *Virtualización mediante máquinas virtuales*

Tomando como base la arquitectura de red descrita, se realiza un primer despliegue sobre máquinas virtuales en el entorno de emulación GNS3. Se facilita así una interfaz visual que integra en el escenario las máquinas virtuales emuladas (ejecutadas en distintos servidores), incluyendo su gestión remota a través de los terminales telnet y su interconexión a través de la red. De este modo, se puede verificar el funcionamiento correcto de la arquitectura propuesta analizando la viabilidad de la virtualización.

El análisis de la red SDWN [1] identifica como uno de los puntos críticos la virtualización de los puntos de acceso, dada la necesidad de acceder a las interfaces del nodo físico donde estos puedan ubicarse, así como el manejo directo de las tarjetas inalámbricas. Así, la red SDWN implementada consta de nodos AP y controlador emulados con máquinas virtuales desplegadas manualmente en servidores ubicados en el laboratorio de trabajo. Se consigue así un aislamiento completo de los procesos del nodo físico en que residen y un acceso exclusivo de las interfaces de red de interés, a costa de un elevado consumo de recursos.

C. *Virtualización mediante contenedores*

El escenario previo supone un primer análisis de la posibilidad del despliegue virtualizado en un entorno controlado de laboratorio. El objetivo, no obstante, es avanzar en el estudio de un despliegue automático en entornos reales. En este sentido, siguiendo la tendencia actual de desarrollo tecnológico en el 5G, así como el estado del arte en el contexto de computación [5], [6], se ha valorado la utilización de contenedores Docker [7] y Kubernetes como orquestador [6] de la infraestructura de red virtualizada, específicamente, del plano de control de SDWN (AP programables y controlador central). Para validar las ventajas teóricas de los contenedores frente a máquinas virtuales, en términos de eficiencia, se ha comparado la ejecución de un proceso patrón análogo a los AP de la red SDWN, pero simplificado con el propósito de analizar exclusivamente la viabilidad de la virtualización (consistente en un proceso de cifrado, de alta carga computacional, y una inyección continuada de tráfico en Wi-Fi). Como han demostrado las prestaciones observadas, un contenedor consume un porcentaje de CPU similar a la ejecución de su proceso directamente en el nodo físico (43 %) frente a la sobrecarga de la máquina virtual (102,12%, en un equipo de 4 núcleos). Si bien los contenedores no están exentos de posibles desventajas (menor aislamiento del nodo físico y una gestión algo más compleja de las interfaces de red), presentan una mayor eficiencia y una clara ventaja en el contexto de automatización del despliegue, dada su ligereza y menor latencia de activación, por lo que el estudio se ha centrado en verificar las especificaciones técnicas necesarias para la adecuación del contexto inalámbrico dentro de Kubernetes, admitiendo su viabilidad.

III. Despliegue automatizado en Kubernetes

Tomando como referencia los resultados previos, y centrándonos inicialmente en el plano de control de la red SDWN, se han virtualizado las funciones de AP y controlador mediante contenedores Docker. Asimismo, para automatizar y orquestar el despliegue, se ha analizado la conveniencia de utilizar la herramienta Kubernetes [7], ampliamente utilizada para la orquestación de contenedores, sin emular el escenario mediante GNS3. Concretamente, se ha utilizado la distribución ligera de Kubernetes k3s [8] especialmente indicada para entornos IoT y Edge, lo que, además, simplifica la configuración y utilización de la herramienta, siendo por tanto adecuado para el entorno de trabajo del presente estudio, como prueba de concepto no orientada a producción.

A. *Creación de imágenes Docker y configuración de Pod*

Una vez demostrada la idoneidad de los contenedores, se han identificado las necesidades específicas para implementar los AP y el controlador, así como el modo de ejecución adecuado dentro de Kubernetes. Partiendo del código correspondiente se han creado las imágenes para contenedores Docker, ubicándolas en el repositorio público Docker Hub, para ser utilizadas desde Kubernetes.

Kubernetes despliega los contenedores como parte de su unidad mínima de cómputo, denominada Pod. Si bien en un Pod pueden ejecutarse varios contenedores, en este trabajo se

considera un único contenedor por Pod, siendo por tanto análogo la ejecución de un contenedor con la ejecución de un Pod. El despliegue ubica los Pod dentro del Clúster de Kubernetes, es decir, del conjunto de nodos que forman parte de la infraestructura de cómputo. La decisión a la hora de ubicar los Pod la toma Kubernetes, aunque existe cierta capacidad por parte del usuario final de seleccionar los nodos adecuados para ello. En el caso del escenario de trabajo, la SDWN, resulta indispensable ubicar los AP de acuerdo a la cobertura que se desea planificar, abriendo la posibilidad a una activación/desactivación dinámica de los mismos, de acuerdo a la demanda. En el caso de la aplicación del controlador, su ubicación, aun pudiendo realizarse de acuerdo a restricciones de latencia u otros parámetros, inicialmente puede establecerse sin condiciones. Por otra parte, a la hora de poner en ejecución los Pod, hay que considerar que los AP, dado que requieren acceso directo a la tarjeta inalámbrica, deben configurarse de modo que tengan acceso a la red del nodo físico, así como privilegios para la manipulación de la misma.

Con estas premisas, se definen los Pod necesarios a desplegar. La Fig. 2 muestra un ejemplo de Pod para un AP que incluye, como especificaciones distintivas, el acceso a la red del nodo físico (hostNetwork: true), el acceso privilegiado (securityContext: privileged: true) y la ubicación en un nodo particular, de acuerdo a una etiqueta característica del mismo (nodeSelector: disktype: master).

```
apiVersion: v1
kind: Pod
metadata:
  name: ap
spec:
  hostNetwork: true
  containers:
    - image: <nombre_imagen_docker_AP>
      imagePullPolicy: Always
      securityContext:
        privileged: true
      name: ap
      command: ["sh", "/home/ap"]
  nodeSelector:
      disktype: master
```

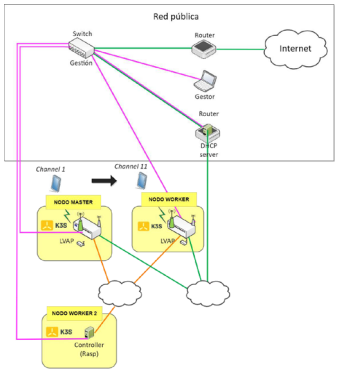
Fig. 2. Pod para el despliegue de un AP

B. *Estudio de conectividad de red en Kubernetes*

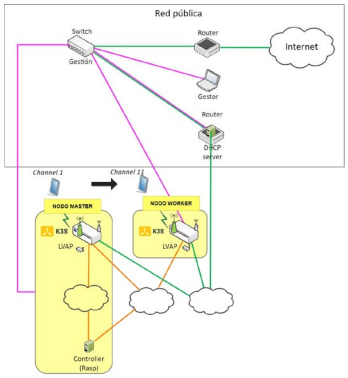
Un aspecto clave a la hora de desplegar nuestras funciones virtualizadas mediante Kubernetes es garantizar la conectividad esperada entre las mismas. Kubernetes implementa su propio modelo de red que asegura la comunicación directa entre los diversos Pod de un Clúster, y entre los nodos y los Pod que residen en ellos, de manera directa, sin traducción de direcciones (NAT). Aunque dicho modelo permite un funcionamiento adecuado en el despliegue estándar de aplicaciones, donde el requisito es garantizar su accesibilidad, replicabilidad o balanceo de carga de manera transparente (por ejemplo, un servicio Web con demandas variables de servicio), cuando lo que se despliega son funciones de red en sí mismas, de cuya utilización depende la conectividad real entre los usuarios (terminales móviles conectados a la SDWN o los propios AP y el controlador), es necesario estudiar la viabilidad de la propuesta o identificar las adaptaciones necesarias.

Respetando el modelo de red de Kubernetes, existen diversas implementaciones (plugin CNI – *Container Network Interface*) que establecen la conectividad y alcanzabilidad de modo diferente, si bien en su mayoría crean redes *overlay* sobre la infraestructura física de conexión entre los nodos del Clúster que hace transparente para los Pod la comunicación sobre la misma y a su vez gestionan y actualizan automáticamente las tablas de rutas y direccionamiento físico relacionado. Asimismo, el direccionamiento de las interfaces así gestionadas queda al cargo de Kubernetes, que asigna las direcciones IP dentro de unos CIDR específicos propios. En el contexto de este trabajo, donde se ha utilizado la distribución de Kubernetes k3s, el CNI configurado por defecto, Flannel [9], [10], establece una red *overlay* basada en VxLAN entre los nodos del Clúster y gestiona, mediante el *daemon* flanneld, el mantenimiento de tablas de rutas y ARP asociadas a las IP asignadas a los diversos Pod dentro de los nodos.

En este escenario se plantea la necesidad de estudiar la conectividad entre los Pod de los AP y el controlador (red de control de la SDWN), con las particularidades de implementación mencionadas previamente (como el acceso privilegiado a la red del host de los AP), para verificar la posibilidad real de su despliegue.



(a) Clúster de 3 nodos: un Pod (AP o controlador) por nodo

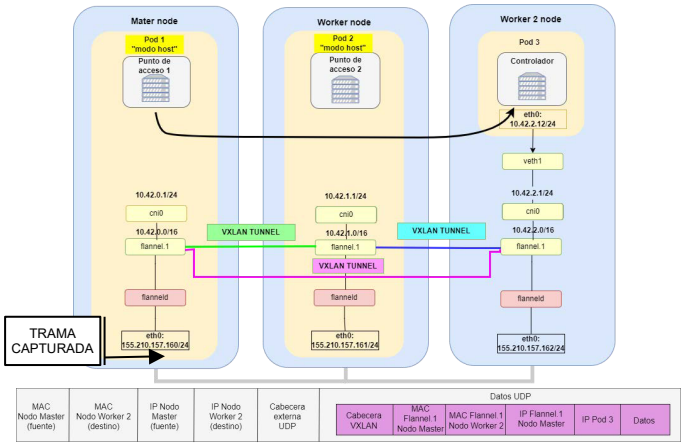


(b) Clúster de 2 nodos. AP y controlador co-localizados

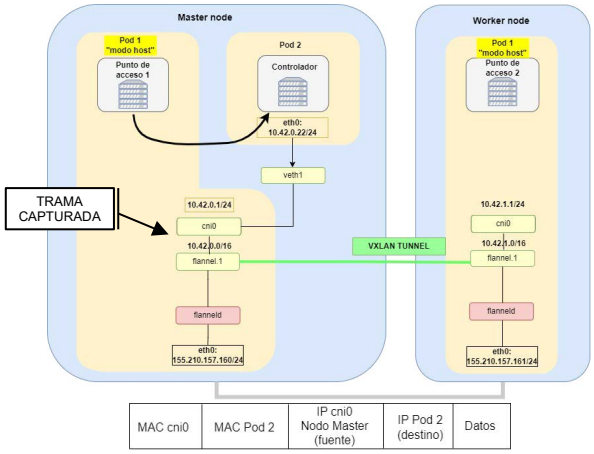
Fig. 3. Escenario de red. Ubicación de APs y controlador en Clúster Kubernetes.

Asumimos un Clúster de Kubernetes con un nodo central (máster) que incluye también capacidad de cómputo (worker). Tomando como referencia un escenario de red mínimo (2 AP y un controlador), se plantean dos posibles alternativas: ubicación de un Pod por nodo físico (por lo tanto, un clúster de tres nodos) o ubicación independiente entre los AP y co-localización del controlador con uno de ellos (clúster de dos nodos), como compara la Fig. 3. En ambos casos, tomando como referencia el modelo de red Flannel, la comunicación debe establecerse de acuerdo a lo mostrado en la Fig. 4.

El plano de control de la red SDWN debe permitir una comunicación transparente entre los AP y el controlador ubicados en sus correspondientes Pod. El controlador recibe una dirección disponible del CIDR de Flannel en Kubernetes en su interfaz por defecto (eth0), conectada al bridge correspondiente del nodo (cni0) y con acceso a cualquier otro Pod del Clúster. Para ello, se establecen comunicaciones directas en el mismo nodo, mediante el bridge cni0, o entre nodos a través de los túneles VxLAN desplegados entre las interfaces flannel.1. Los AP, por su parte, se alojan en Pod con acceso al host y, por lo tanto, sin interfaz eth0. No obstante, dado su acceso a las interfaces propias del nodo, su conectividad, tanto a Pod del nodo como a externos, queda garantizada mediante la interfaz bridge cni0 (también con dirección IP del CIDR de Flannel). Es decir, cualquier Pod, controlador aislado o AP con acceso a la red del nodo, puede comunicarse con los Pod del Clúster mediante las interfaces eth0 internas o cni0 del nodo, gracias a la red Flannel.



(a) Pod localizados en dos nodos. Red Flannel.



(b) Pod co-localizados en un nodo. Bridge cni0.

Fig. 4. Conectividad entre un Pod de AP (host) y otro de controlador en k3s.

C. *Prueba de concepto*

Una vez establecidas las configuraciones necesarias para el despliegue automático del plano de control de la red SDWN, se ha realizado una prueba de concepto para comprobar el correcto funcionamiento de un *handover* efectivo entre 2 AP (desplegados en sus correspondientes Pod). Esta prueba ha verificado la idoneidad de la propuesta permitiendo asimismo establecer las conclusiones iniciales y futuras líneas de trabajo.

IV. Conclusiones y líneas futuras

El trabajo presentado constituye un primer paso en el desarrollo completo de una red virtualizada inalámbrica desplegada automáticamente y orquestada mediante Kubernetes. Se ha verificado la automatización del despliegue del plano de control garantizando la conectividad entre los puntos de acceso y el controlador, a través de la propia red gestionada por Kubernetes, lo que garantiza la flexibilidad de la propuesta. Como prueba de concepto, la realización de un handover de un terminal de usuario entre AP, con las mismas prestaciones alcanzadas con la red SDWN no virtualizada, garantiza el mantenimiento de la funcionalidad previa, abriendo así la puerta al desarrollo de nuevas funcionalidades, facilitando su ubicación oportuna y la automatización de los procedimientos de gestión.

Así, el trabajo en proceso plantea las siguientes líneas futuras: Por una parte, identificadas las necesidades de adecuación del despliegue a Kubernetes, se propone la ampliación de funciones inteligentes en los AP y controlador, elementos clave de la SDWN, mediante un rediseño más adaptado al contexto de automatización propuesto. Por otra parte, con el objetivo de integrar la SDWN en un escenario de conectividad global, se contempla la virtualización de los planos de datos y gestión, incluyendo aquellas funciones cuya programación y ubicación dinámica incrementen la flexibilidad y escalabilidad de la propuesta, aprovechando, para ello, las capacidades propias de Kubernetes.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos NEWLAN (PID2022-136476OB-100) del gobierno de España, CeNIT (T31_23R) del Gobierno de Aragón y NeWLAN (UZ2022-IAR-08) de la Universidad de Zaragoza.

Referencias

[1] F. Bouhafs et al., "Wi-5: A Programming Architecture for Unlicensed Frequency Bands," in IEEE Communications Magazine, vol. 56, no. 12, pp. 178-185, December 2018, doi: 10.1109/MCOM.2018.1800246.

[2] J. Saldana et al., "Unsticking the Wi-Fi Client: Smarter Decisions Using a Software Defined Wireless Solution," in IEEE Access, vol. 6, pp. 30917-30931, 2018, doi: 10.1109/ACCESS.2018.2844088.

[3] J. Saldana et al., "Attention to Wi-Fi Diversity: Resource Management in WLANs With Heterogeneous APs," in IEEE Access, vol. 9, pp. 6961-6980, 2021, doi: 10.1109/ACCESS.2021.3049180

[4] J. Schultz, R. Szczepanski, K. Haensge, M. Maruschke, N. Bayer and H. Einsiedler, "OpenGUFI: An Extensible Graphical User Flow Interface for an SDN-Enabled Wireless Testbed," 2015 IEEE CIT; IUCC; DASC; PICOM, Liverpool, UK, 2015, pp. 770-776, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.113.

[5] Diferencias entre los contenedores y las máquinas virtuales. https://www.redhat.com/es/topics/containers/containers-vs-vms. Consultado en 15/06/2023

[6] *¿Qué es Kubernetes?* https://kubernetes.io/es/docs/concepts/overview/what-is-kubernetes/ Consultado en 15/06/2023

[7] ¿Qué es Docker y cómo funciona? https://www.redhat.com/es/topics/containers/what-is-docker. Consultado en 15/06/2023

[8] K3s - Lightweight Kubernetes. https://docs.k3s.io/ Consultado en 15/06/2023

[9] The Kubernetes Networking Guide. CNI. Flannel https://www.tkng.io/cni/flannel/ Consultado en 15/06/2023

[10] Flannel CNI. https://github.com/flannel-io/flannel Consultado en 15/06/2023.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

The application of Digital Twins in Network Virtualization

Amir Hossein Banisadr, Xavier Hesselbach
Department of Network Engineering (ENTEL)
Universitat Politècnica de Catalunya (UPC)
C/ Jordi Girona, 31. 08034 Barcelona
{banisadr.amir.hossein, xavier.hesselbach}@upc.edu

In this paper, digital twins (DT) are being actively explored to introduce the novel notion of a virtual network twin (VNT), opening up new horizons in this domain. To achieve this goal, the critical requirements in terms of the architecture and communication are pointed, and also the benefits from the AI to efficiently control the operations and management between networks. Network requirements, especially ultra-low latency, and intelligent control of the actions are studied in order to guarantee an effective network twin.

***Keyword*- Digital twin, network virtualization, ultra-low latency, AI.**

I. Introduction

Over the last decades, the Internet has been developed by diverse infrastructure technologies which provide distributed applications and protocols. Unfortunately, this diversification has become an impediment to Internet development. So, this paradox is called the ossification problem [5]. Owing to the fact that Internet service providers (ISPs) are the owners of the current architecture and without their agreement, it is impossible to change or adopt a new one. These limitations have prevented the development of many possible applications and services.

Network virtualization (NV) is a promising solution to solve this issue. NV permits sharing the physical network (PN) resources among various virtual networks requests. Network virtualization is composed of a number of virtual networks which co-exist over the same PN. As such, NV includes two subsets, a physical network which is owned and operated by an infrastructure provider and composed by physical nodes connected by physical links that form the physical topology and a virtual network which is composed of a set of virtual nodes, each hosted on a physical one and virtual links, each established over a physical one. Based on architectural perspective, network virtualization allows [6]: Coexistence of many virtual networks on the same physical one, flexibility among them in order to their routing and forwarding functions and control protocols from the specific physical network, manageability, scalability and security. These characteristics can be held by suitable solutions.

Digital twin (DT) is proposed in this work to providing a network virtualization twin regarding all its features, DT will exploit the expected enhancements envisioned from 6G, the next generation of networks, especially regarding the ultra-low latency and hyper-connectivity. DT can be described as a digital encounter of a physical system with repeated processes through a data connection that changes the physical network to virtual one with high levels of synchronization, security, and flexibility.

DT consists of three key pillars: the physical, virtual and connection pillar. The physical pillar represents the real applications and services of the physical network. The virtual pillar is known as virtual twin, and includes historical data, decision support systems, artificial intelligence, capabilities and visualizations, this pillar can send control commands to the physical one. The connection pillar is considered as the communication bridge between these two parts and enables synchronizing, monitoring, and managing actions between them [9].

Based on the above definition, digital twin method enables to resolve all concerns in network virtualization, it means that synchronization, monitoring, and decision-making independently on one hand and scalability and highly security on other hand make DT the optimum approach to consider all NV´s aspects.

In this paper, the concept of Digital twin and network virtualization are defined in Section II considering the related relevant studies. In section III, the proposal and the open issues will be described and commented. Finally, Section IV summarizes the conclusions reached so far within this work and indicates the next steps.

II. Background

One of the issues that is faced by the current Internet is ossification, which denotes the network's inherent resistance to change [1]. Network virtualization (NV) has been recognized as a potential approach for addressing the current network ossification [2]. NV is designed to make numerous virtual networks on a shared physical network (PN), enabling independent deployment and management of each, which can be the solution to the issue of network rigidity as enabling the potential for dynamicity and diversity within the network [3]. This brilliant technology utilizes resources to the substrate network (SN) efficiently, provides a significant opportunity to implement and evaluate new architecture designs and serves to hinder useless expansion of network infrastructure [4]. Within the NV, coexisting multiple virtual networks constructed of different resources on the same PN, capability to create a hierarchical structure among virtual networks, ability of revisitation means that a single physical node to accommodate multiple virtual nodes belonging to the same virtual network, enhancing flexibility and manageability each virtual network independently and scalability to support an increasing number of virtual networks are happened. [5]. With respect to reviewing numerous studies, it seems that existing solutions could not reach optimal approaches for network virtualization in all aspects. For instance, reduction of complexity management and avoiding bandwidth performance issues between logical and physical networks is considered in this article without considering other characteristics such as security, flexibility, or scalability [6].

To achieve an optimum solution for network virtualization, digital twin is proposed. The DT was initially introduced by Michael Grieves at the University of Michigan in 2003 which is described as a comprehensive software representation of a physical object (PO), including its properties, conditions, and behaviors of the real-life completely [7]. A digital twin can be defined as a virtual representation of a physical asset that encompasses both a historical record of the asset's state and real-time one on its current state. This asset can refer to an object, process, network, or even a system. It is important to note that a DT surpasses the functionalities of an avatar, surveillance system, simulation, or a simple model [8]. Furthermore, the bidirectional connection with the PT makes a DT more advanced and capable compared to a surveillance system [9]. The modularity of digital twins allows for the creation of individual DTs for each component of a physical object. These smaller DTs can then be interconnected to form a mega-DT. This characteristic enables the development of hybrid simulation and prototyping systems without any concern about scale of that [10]. DT in other words, is defined as a self-adapting, self-regulating, self-monitoring, and self-diagnosing system including following properties: (1) symbiotic relationship between PO and its digital encounter (2) flexibility to select model and technology to achieve fidelity, rate of synchronization and other critical properties (3) supporting different services in various aspects of physical entity [11]. With respect to the definitions of DT, the major functions include prediction prior to actual runtime by analyzing the historical data of real system, real time monitoring by collecting and processing real time data including current state, performance and condition, diagnosis after operation by analysis the behavior of the real system, optimization and testing and decision support [12]. In order to achieve these functions, a certain set of requirements are discerned as functional requirements and service ones [13]. The functional requirements outline specific features describing system behavior that DT should have to optimally carry out its task, some of which are determined as: data collection policies and tools, data repository, data model and life cycle management. The service requirements are specified characteristics that a DT should possess based on different user demands, these features are compatibility with numerous types of models and PO elements, flexibility to meet demands of PO's applications with single or multiple purposes, privacy to guarantee data protection for all users during entire life cycle, scalability to handle and replicate physical entity of any scale, security to secure data throughout its life cycle, synchronicity to represent the real-time counterpart with acceptable latencies, repeatability and reproducibility to enable testing and validation of new technologies on real-time data [14]. Sixth-generation (6G) is envisioned to be determined by ultra-low latency, widespread connectivity and high security [15], which is the best choice to cover all functions of DT.

According to the viewpoints presented above, we describe a definition of DT to apply in NV. Digital twin is an intelligent and developing system which monitors, controls, diagnoses, tests, protects and optimizes the physical network throughout its life cycle.

III. proposal

At the present time, the access to a specific network is usually done by means of a tunnel, using VPNs especially regarding security issues.

In this article, we introduce the innovative concept of a digital twin for network virtualization, aimed at elevating access experiences to a level where users enjoy seamless connectivity that mirrors the authentic, original network experience, providing the same feeling from any connection inside the original network.

The virtual network twin (VNT) is a virtualized replica of the original, operating in the same manner, and keeping the synchronism with the real network. The proposal requires a VNT manager (Fig. 1) in order to coordinate the actions required, synchronize and maintain the virtualized network. Hence, any modification made is automatically synchronized by means of a VNT manager, ensuring the continuous preservation of system-wide consistency.

Therefore, the virtual network twin allows to deploy a functional replica wherever is required.

Fig. 1. VNT manager

Regarding the VNT manager, the main target is the communication protocol based on the requirements. Some of these critical requirements in network virtualization fall in the synchronicity of resources.

The following issues has been identified in order to define a set of rules and protocols running in the DT manager:

- Acceptable latency values required between physical network with its virtual counterpart.
- Procedures to manage every virtual networks independently.
- Scalability to handle and replicate networks of any scale.
- Security to provide mechanisms to secure data throughout its life cycle and to avoid poisoned traffic or other risks and attacks.

In short, the suitable suite of protocols must include the capability to synchronize physical network and all virtual counterparts in real-time considering flexibility, scalability, and security.

When addressing the requirements in terms of network performance (latency, throughput, resilience) and the capacity to efficiently handle the requests using intelligent strategies, AI with 6G networks emerges as a promising team, ready to offer tailored strategies to provide the appropriate usage experience.

AI is used to learn and optimize the behavior between physical and virtual domains.

The proposal will be analyzed under the following situations:

The impact of the latency in the service provided as requested from the virtualized network twin.

The protocols in terms of the topology and services of the network.

The AI strategies in order to protect the networks against wrong operation or attacks from others. In this field, strategies based in reinforcement learning are envisioned, regarding the capacity to take actions to optimize the reward function. Samples will be generated synthetically from selected defined use cases and from basic testbeds.

In this phase, it is expected to check the performance results in early basic proposals, and optimize the strategy to reach comprehensive protocol.

IV. Conclusions and future works

In this paper, we introduced the early ideas behind the digital twin concept for network virtualization, their envisioned features, and functions. The novel concept of virtual network twin is defined. The need of a prompt protocol is pointed to cover some critical requirements including low latency, flexibility, scalability, and security. The analysis, proposal and design of the protocol in the VNT manager will define the next steps to get initial results.

V. Acknowledgements

This work has been supported by the Agencia Estatal de Investigación of Ministerio de Ciencia e Innovación of Spain under project PID2022-137329OB-C41/ MCIN/ AEI/10.13039/501100011033.

References

[1] *Chau, Phanvu, and Yang Wang. "Security-awareness in network virtualization: A classified overview." 2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems. IEEE, 2014.*

[2] *Zhang, Hongjing, et al. "Network operation simulation platform for network virtualization environment." 2015 17th Asia-Pacific Network Operations and Management Symposium (APNOMS). IEEE, 2015.*

[3] *Xingtao, Liu, et al. "Network virtualization by using software-defined networking controller based Docker." 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference. IEEE, 2016.*

[4] *Nguyen, Khoa TD, Qiao Lu, and Changcheng Huang. "Rethinking virtual link mapping in network virtualization." 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall). IEEE, 2020.*

[5] *Belbekkouche, Abdeltouab, Md Mahmud Hasan, and Ahmed Karmouch. "Resource discovery and allocation in network virtualization." IEEE Communications Surveys & Tutorials 14.4 (2012): 1114-1128.*

[6] *Seddiki, M. Said, and Mounir Frikha. "A non-cooperative game theory model for bandwidth allocation in network virtualization." 2012 15th International Telecommunications Network Strategy and Planning Symposium (NETWORKS). IEEE, 2012.*

[7] *Grieves, Michael. "Digital twin: manufacturing excellence through virtual factory replication." White paper 1.2014 (2014): 1-7.*

[8] *Minerva, Roberto, Gyu Myoung Lee, and Noel Crespi. "Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models." Proceedings of the IEEE 108.10 (2020): 1785-1824.*

[9] *Ahmadi, Hamed, et al. "Networked twins and twins of networks: An overview on the relationship between digital twins and 6G." IEEE Communications Standards Magazine 5.4 (2021): 154-160.*

[10] *Jacoby, Michael, and Thomas Usländer. "Digital twin and internet of things—Current standards landscape." Applied Sciences 10.18 (2020): 6519.*

[11] *Mihai, Stefan, et al. "Digital twins: a survey on enabling technologies, challenges, trends and future prospects." IEEE Communications Surveys & Tutorials (2022).*

[12] *Mertes, Jan, et al. "Development of a 5G-enabled Digital Twin of a Machine Tool." Procedia CIRP 107 (2022): 173-178.*

[13] *Lichtzinder, B. Ya, and S. A. Chernysheva. "Digital Twins." 2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). IEEE, 2022.*

[14] *Kuruvatti, Nandish P., et al. "Empowering 6G Communication Systems with Digital Twin Technology: A Comprehensive Survey." IEEE Access (2022).*

[15] *Bhat, Jagadeesha R., and Salman A. Alqahtani. "6G ecosystem: Current status and future perspective." IEEE Access 9 (2021): 43134-43167.*

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Server-Side GNSS Spoofing Detection Challenges for Vehicle Tracking Applications

José Jesús Sánchez Gómez, Isaac Agudo.
Departamento de Lenguajes y Ciencias de la Computación,
Universidad de Málaga (UMA)
sanchezg@uma.es, isaac@lcc.uma.es

This paper focuses on the risks linked to the Global Navigation Satellite System (GNSS) and introduces a scenario involving a data-transmitting device connected to a cloud-based service. We explore potential attackers and the diverse attacks possible during the communication and data-processing stages of the scenario. Additionally, we categorize current detection methods based on the information they employ to detect spoofing attacks and discuss their limitations concerning Server-Side detection. Ultimately, we propose solutions and future lines of work to mitigate these problems.

***Keywords*—GNSS Spoofing, Location Spoofing, Server Verification, Security**

I. INTRODUCTION

Currently there are numerous satellite-based navigation systems that provide services to users. Although people commonly use the term GPS (Global Positioning System) to refer to all systems able to determine the location of a device, this usage is incorrect. GPS refers to the U.S. Global Positioning System, which was first developed by the United States Government in the 1970s [1]. The accurate term used to refer to the collection of systems to determine different location parameters is GNSS (Global Navigation Satellite System). This encompasses different systems, such as the U.S. GPS, the European Galileo System, the Russian GLONASS (GLObal NAvigation Satellite System) or the Chinese BDS (Beidou Navigation Satellite System), among others [2].

GNSS reception modules are commercially available for purchase. These modules come in a wide range of prices, and while more affordable options provide basic positioning and navigation functionalities, higher-end, more expensive modules offer enhanced features and greater accuracy. These modules can be integrated into various systems and applications, enabling a wide range of functionalities and being used for tasks such as obtaining accurate time synchronization or determining the precise location of the system itself. This location information can then be utilized by a variety of applications. For instance, in the case of vehicular navigation systems, the accurate positioning data obtained from these modules enables drivers to receive turn-by-turn directions, real-time traffic updates, and optimized route suggestions. Additionally, competitive applications like Strava use location information to allow users to track and analyze their performance in activities such as running or cycling, facilitating the comparison of results with others.

The widespread adoption of GNSS reception modules has led to an increasing number of companies using vehicle tracking for various purposes. For example, fleet management has significantly benefited from the integration of GNSS modules, not only in terms of real-time monitoring and control but also by enabling businesses to impose restrictions on vehicle usage. With the advanced capabilities of GNSS modules, companies can implement several policies and regulations, such as the enforcement of time-based restrictions or geofenced areas in order to restrict vehicle access to certain zones or regions. Insurance companies also leverage GNSS reception modules to offer pay-as-you-drive insurance programs. By tracking vehicles' location and driving behavior, insurers can assess risks more accurately and provide personalized plans. Safe driving habits can be rewarded with lower premiums, incentivizing responsible behavior on the road.

In 2021, the revenues of the GNSS market exceeded €200 billion. It is expected that by the next decade, GNSS-related revenues will reach €500 billion, with more than 10 billion GNSS devices in use [3]. The potential advantages gained from carrying out these attacks have motivated malicious actors to engage in GPS attack due to the relative ease with which such attacks can be executed.

The following section provides an overview of the architecture of a location-based service while also outlining the potential attackers to the system. Subsequent to this characterization, a range of possible attacks that can be carried out are addressed. Section III furnishes a comprehensive insight into various spoofing detection

mechanisms, categorizing them based on the data used to detect an attack and addressing their limitations. Lastly, section IV presents a conclusion about the detection methods and briefly presents future lines of work.

II. SYSTEM TOPOLOGY AND SECURITY CONCERNS

The typical data flow in a cloud-based service using GNSS is as shown in Fig. 1. Satellites broadcast information to all potential receptors. Then, the receivers compute and send its location to the server in order to get some service, e.g. Weather information.

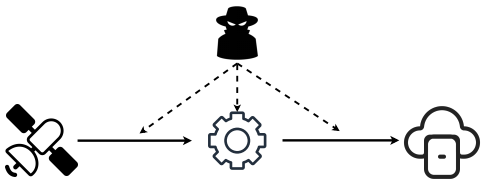


Fig. 1. Location-Based Architecture.

In this flow there are three critical steps:

- **GNSS signal**. The GNSS signal is broadcasted over the air. Since these signals are emitted without making use of any authentication method, these are vulnerable to spoofing attacks as described in the next subsection.
- **Vehicle processing.** The signal received from the satellites is processed by the in-vehicle device in order to determine the location of the vehicle. If the device is compromised, the computed location could be modified.
- **Server reception.** After computing its location, the device transmits it to the server. If this connection is not secure, the transmitted data being sent could be compromised.

Considering these steps, we can focus on the possible types of attacks depending on the attacker:

- **External.** An external attacker would have no access to the device itself. Consequently, attacks carried out by this type of attacker must target the GNSS signal received by the device.
- **Internal.** An internal attacker, e.g. the user of the device, would have physical access to the device, being able to act on all three steps described.

The next subsections outline various potential attacks depending on the attacker.

A. External malicious

The GNSS system makes use of different constellations of medium earth orbit satellites. Reception modules are capable to track satellites from different constellations, which means that different GNSS systems can be jointly used [2]. GNSS satellites make use of accurate atomic clocks to transmit signals to Earth. These signals are used by GNSS receivers to estimate the distance to the satellite by measuring the TOA (time of arrival) of the signal. In order to achieve that, both the satellites and the receiver clocks must be synchronized [4]. The exact position of each of the satellites of the GNSS system is included in the Ephemeris, a set of data periodically broadcasted by satellites or posted on Internet. These can be used by receivers in order to obtain information about the satellites that are in their line of sight. While the accuracy and availability of the system have significantly evolved since its inception, the implementation of publicly-available integrity methods is only beginning to take shape. GNSS systems offer different services, some being authenticated e.g. GPS PPS (Precise Positioning Service) or Galileo PRS (Public Regulated Service). However, the authenticated services are encrypted, being the access to these services limited to the military or authorized government personnel. The first public-authenticated GNSS service has been deployed by Galileo OSNMA (Galileo Open Service Navigation Message Authentication) and has just started this year (2023). The system uses the TESLA (Timed Efficient Stream Loss-tolerant Authentication) protocol to broadcast authentication data. Nevertheless, although this system solves the integrity problems of the GNSS system, at the moment there are few GNSS modules supporting this technology. The U.S. GPS system is also developing its own authentication system: Chimera.

GNSS satellites typically transmit signals at a power level of around 44 dBm. However, due to considerable distance between the satellites and the Earth's surface, which exceeds 20km, the power received at the Earth surface reaches approximately -130dBm considering a clear view of the sky [2]. The weak power of the received signal exposes the system to signal spoofing attacks, which can be easily executed using SDRs (Software-Defined Radios) or similar devices. For instance, Markgraf showcased at OsmoCon the utilization of a modified €5 USB to VGA card that was capable of executing such attacks [5]. This serves as evidence of the system's susceptibilities and emphasizes that performing these attacks does not necessarily require a significant budget. Additionally, there is well-documented Open-Source software available for generating fake GNSS messages [6].

Attacks against UAVs, boats and vehicular systems are well documented. Academics have demonstrated that it is possible to perform these kind of attacks [7]–[9]. In real situations, attackers were able to modify the estimated location of all cars in a motor show [10] or spoof the location of several boats in the black sea[11].

Although most of these attacks can be solved using signal authentication methods, it has been demonstrated that the system would still be vulnerable to relay attacks [12], which consist on capturing real traces in a location, in order to relay them to the victim receiver.

B. Internal

The user of the device containing the GNSS receiver module can also be regarded as an attacker. This attacker would benefit from the attack's possible consequences, such as overriding Geo-Fences or emulating driving behaviours. An internal attacker could launch an attack to the

GNSS signal, the same way as an external attacker would. An example of this situation could involve a user with a sealed device, unable to access it to alter its hardware or software. When an external attacker executes an attack on the GNSS signal, all of the devices in an area are affected, since the attack signal must cover a wide area to affect the attacked device. In contrast to this, an internal attacker would not need to affect other devices, since the target device would be close enough to emit the signal solely to itself.

If an internal attacker gains access to the device, they could modify the software of the device to manipulate the computed parameters or alter the data sent to the server. Additionally, this attacker could change the legitimate location data provider to a compromised one. For example, in the context of attacking the Pokemon Go application, it was common to utilize the Android developer API, which enables the emulation of the device's location. Similarly, data obtained from various sensors could be emulated or manipulated using similar techniques.

III. OVERVIEW OF SPOOFING DETECTION MECHANISMS

In the previous section, we described the scenario under examination within the automotive environment. This scenario encompasses two data flows intrinsic to providing the client location to the server. One of these flows pertains to the GNSS signals that the GNSS receiver obtains, while the other pertains one to the location data-sets that are sent to the server.

Various researchers have developed methods to detect spoofing attacks. We classify these detection methods in Fig. 2, based on the parameters and data used to detect GPS spoofing attacks. Despite the extensive documentation on these types of attacks, most of them consider an external attacker, who would focus on GNSS Signal Spoofing, whereby the signal broadcasted by the satellites is overridden by the signal transmitted by the attacker. Nonetheless, in our scenario, we also consider the possibility of an internal attacker who might target the data transmitted to the server instead of the GNSS signal.

Methods based only in GNSS use the received parameters from the GNSS receiver, such as the PDOP (position dilution of precision) or the physical characteristics of the received signal such as its SNR (signal-noise ratio) and power. These methods have been proven to work, obtaining a detection percentage of almost 99%, as described in [13]. Several implementations have been developed in the latest years, using ML (Machine Learning) and neural networks, evaluating the correlation and variation of the different parameters [14].

Despite these methods have been proved to work on the client side in most cases, its feasibility when deployed on the server side is not demonstrated. These would require the client to send the received GNSS traces and parameters to the server in order for it to verify them. Since we consider that the client is able to send illicit traces to the server, the traces could be generated by the client, emulating the correlations and expected variations of the parameters, in order for the algorithm running on the server-side to not detect the spoofing. In addition, despite these attacks have been proven to work in simple scenarios, attackers using multiple well-located antennas are able to bypass the security system.

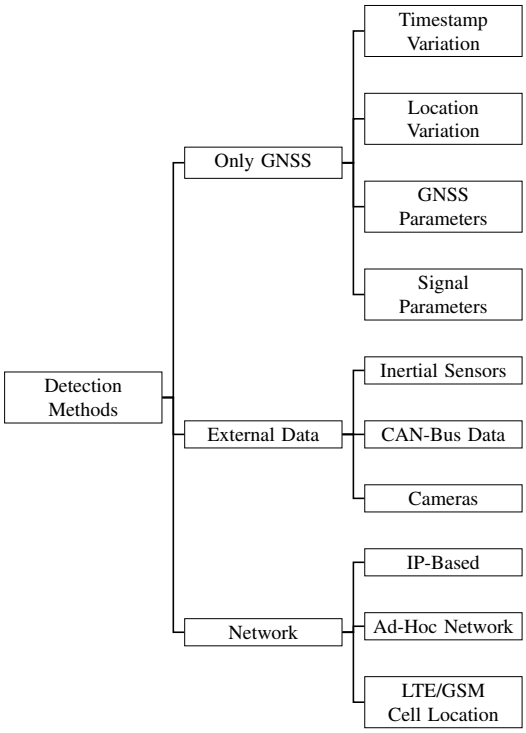


Fig. 2. Spoofing Detection Methods depending on the method used

Systems using external data, such as inertial sensors or obtained CAN-BUS data from the vehicle share a common objective: utilize this data to estimate the path followed by the vehicle, obtaining the possible turns, speed and distance traveled. Then, these estimates are compared to the data obtained using the GNSS module to verify its authenticity. Additionally, the integration of cameras allows to use ML for the same purposes. For example, captured frames can be compared with globally available images, allowing the server to verify the location. These methods induce higher complexity to the security system, thereby increasing the difficulty of a successful attack. As a way for the server to verify the location, these data-sets must be transmitted to the server. However, this leads to on a increase on the required bandwidth and processing power, in order for the server to run the mentioned verification methods. Despite that, determined attackers could attempt to generate synthetic sensor data or manipulated images in order to bypass the system.

Finally, we can consider systems that use network-obtained parameters in order to validate the location. Despite the diversity and limitations of the methods in this category, they are well-suited for a server-based verification. Some of these methods require the server use the IP of the client or the localization of the cell tower the device is connected to estimate the position of the

client. However, the primary drawback of these methods lies in the range of possible locations, since each LTE cell tower typically covers a radius of up to 3km, while a GSM tower can theoretically cover up to 15km [15]. Additionally, methods employing ad-hoc network facilitate communication among clients, enabling the comparison of received signal characteristics to detect attackers. The received parameters can be sent to the server, enabling it to compare the parameters obtained from the different clients to identify a client which is sending discrepancies in order to the determine if the reported location is genuine. Moreover, Ad-hoc networks can be utilized to enable clients to verify the location of each others, being able to report compromised devices to the server. These network-based methods may require to be combined with any of the others mentioned above in order to present a feasible solution.

IV. CONCLUSIONS AND FUTURE WORK

Detecting Location Spoofing on the server side proposes some challenges to overcome.

As already mentioned in Section III, there are many works on how to prevent GNSS signal spoofing. We have also mentioned the risk of compromised on-board devices, either by an external attacker or the user itself in order to circumvent access restrictions in location-based services.

Some solutions in this area focus on sensor fusion technologies, but the emerging use of AI to generate deep fakes in other fields makes feasible the possibility of an attacker using these tools to generate artificial data from sensors that would seem real.

A promising line of work in this area is the use of HSM (Hardware Secure Module) to offload GNSS information processing from the GNSS signals to a trusted element in the car, avoiding being tampered by an attacker: for example, the DRACONAV project aims to develop a secure GNSS module able to detect attacks using multi-constellation, a secure MCU and motion sensors, being able to deliver signed data . A combination of such module with OSNMA or Chimera would be a nice approach to a feasible solution. Another interesting research topic is the use of complex ML algorithms to detect synthetic traces received on the server using previous training data. Additionally, the data received from different clients and the network can be analyzed to identify the discrepancies produced by an attack.

V. ACKNOWLEDGEMENTS

This work is part of the research project SECUREDGE (Security Services Platform for the Protection of Edge Scenarios), with PID2019-110565RB-I00, funded by AEI/10.13039/501100011033.

REFERENCES

[1] NASA. “Global positioning system history.” T. May, Ed. (Oct. 27, 2012), [Online]. Available: https://www.nasa.gov/directorates/heo/scan/communications/policy/GPS_History.html (visited on 05/17/2023).

[2] E. Kaplan and C. J. Hegarty, *Understanding GPS/GNSS : Principles and Applications, Third Edition, Principles and Applications, Third Edition*. Artech House Publishers, 2017.

[3] EUSPA: European Union Agency for the Space Programme, “Euspa eo and gnss market report,” 2022. [Online]. Available: https://www.euspa.europa.eu/2022-market-report.

[4] J. L. B. Valero, N. G. Villen, and R. C. Romá, *GNSS GPS, Galileo, Glonass, Beidou. Fundamentos y métodos de posicionamiento*. Universitat Politècnica de València, 2019.

[5] S. Markgraf. “Osmo-fl2k.” (2018), [Online]. Available: https://osmocom.org/projects/osmo-fl2k/wiki/Osmo-fl2k (visited on 05/18/2023).

[6] T. Ebinuma, *Gps-sdr-sim*, https://github.com/osqzss/gps-sdr-sim.

[7] J. Gaspar, R. Ferreira, P. Sebastião, and N. Souto, “Capture of UAVs Through GPS Spoofing Using Low-Cost SDR Platforms,” *Wireless Personal Communications*, vol. 115, Dec. 2020.

[8] J. Bhatti and T. E. Humphreys, “Hostile Control of Ships via False GPS Signals: Demonstration and Detection,” *NAVIGATION*, vol. 64, no. 1, pp. 51–66, 2017.

[9] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, “Unmanned Aircraft Capture and Control Via GPS Spoofing,” *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.

[10] J. Torchinsky. “There’s something very weird going on with cars’ gps systems at the geneva motor show.” (Mar. 8, 2019), [Online]. Available: https://jalopnik.com/theres-something-very-weird-going-on-with-cars-gps-syst-1833138071 (visited on 05/18/2023).

[11] H. Lied. “Gps freaking out? maybe you’re too close to putin.” (Sep. 18, 2017), [Online]. Available: https://nrkbeta.no/2017/09/18/gps-freaking-out-maybe-youre-too-close-to-putin/ (visited on 05/18/2023).

[12] M. Motallebighomi, H. Sathaye, M. Singh, and A. Ranganathan, “Cryptography Is Not Enough: Relay Attacks on Authenticated GNSS Signals,” arXiv, Tech. Rep., Nov. 2022.

[13] T. T. Khoei, S. Ismail, and N. Kaabouch, “Dynamic selection techniques for detecting GPS spoofing attacks on UAVs,” *Sensors*, vol. 22, no. 2, p. 662, Jan. 2022.

[14] M. Nayfeh, Y. Li, K. A. Shamaileh, V. Devabhaktuni, and N. Kaabouch, “Machine learning modeling of GPS features with applications to UAV location spoofing detection and classification,” *Computers’ Security*, vol. 126, Mar. 2023.

[15] P. K. Sharma, D. Sharma, and A. Gupta, “Cell coverage area and link budget calculations in LTE system,” *Indian Journal of Science and Technology*, vol. 9, no. S1, Dec. 2016.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Cryptographic approaches for confidential computations in blockchain

Daniel Morales, Isaac Agudo
Network, Information and Computer Security Lab (NICS)
Departamento de Lenguajes y Ciencias de la Computación
Universidad de Málaga
damesca@uma.es, isaac@uma.es

Blockchain technologies have been widely researched in the last decade, mainly because of the revolution they propose for different use cases. Moving away from centralized solutions that abuse their capabilities, blockchain looks like a great solution for integrity, transparency, and decentralization. However, there are still some problems to be solved, lack of privacy being one of the main ones. In this paper, we focus on a subset of the privacy area, which is confidentiality. Although users are increasingly aware of the importance of confidentiality, blockchain poses a barrier to the confidential treatment of data. We initiate the study of cryptographic confidential computing tools and focus on how these technologies can endow the blockchain with better capabilities, i.e., enable rich and versatile applications while protecting users' data. We identify Zero Knowledge Proofs, Fully Homomorphic Encryption, and Secure Multiparty Computation as good candidates to achieve this.

***Palabras Clave*—blockchain, privacy, confidentiality, secure multi-party computation, zero knowledge proofs, fully homomorphic encryption**

I. INTRODUCTION

Blockchain technologies have emerged as a great solution for integrity, transparency, and decentralization. Broadly speaking, a blockchain network is a set of nodes with a P2P topology, which collaboratively maintain a unified ledger. Despite being conceived to manage cryptocurrency transactions (Bitcoin), other solutions have built a secure and distributed computing platform on top of the network, e.g., Ethereum. The key technology that has made such a secure ledger possible is Byzantine Fault Tolerant Consensus, in which a set of distributed and distrusted nodes can agree on what data is recorded in the ledger each time, resulting in a unified view of the ledger.

Since its conception, many use cases have been proposed [1], [2], e.g., financial, health, supply chain, or government.

Despite the benefits of blockchain, the lack of privacy hinders its adoption. Although it provides pseudonymity, it has been shown that users can be deanonymized [3]. Private connections, e.g., TOR [4], are recommended to mitigate this, at the expense of losing usability. Accessing blockchain data can also be a problem, because the most of the end-users do not own a blockchain node but delegate the access to a node provider[1], making them to become trusted third parties that can cheat on data provided, because end-users do not store all the blockchain data and cannot verify correctness. Also, the provider can perform a profiling attack, tracking all the activity by the user. Such issues directly ballast a real decentralization, which is the main contribution of blockchain.

Another issue is the lack of confidentiality. The evolution of blockchain has drifted towards programmable platforms, e.g., the Ethereum's Virtual Machine, which allows secure general-purpose computations. However, this approach loses its meaning when dealing with confidential data, as data must be decrypted to contribute to an on-chain computation. Different use cases, e.g., financial, or biometric data computation do not fit well with this public model. Finally, regulations such as GDPR can also contribute to restricting use cases.

Although blockchain's lack of confidentiality has been partially addressed, there are some misconceptions. One of the most trendy confidential computing technologies are Non-Interactive Zero Knowledge Proofs (NI-ZKP), which allow verifying that a computation has been performed correctly using specific data, without exposing them. However, NI-ZKP are mainly used in the blockchain ecosystem to achieve succinctness, e.g., in Layer 2 solutions [5]. While they can really help to acquire more capabilities while retaining more confidentiality, it is important to note that NI-ZKP must be computed directly on the plaintext data somewhere. This implies an overhead on

[1]https://www.infura.io/

the data owner's side, or a delegation to a trusted party to compute the proof if there are many data providers involved. There are other cryptographic solutions, e.g., Secure Multi-Party Computation (MPC) and Fully Homomorphic Encryption (FHE), that allow distrusted parties to compute on confidential data without exposing it. Unlike NI-ZKP, these technologies enable delegated computations on confidential data, as will be discussed in next sections.

In this paper we initiate research on blockchain's confidentiality problem, where our main contribution is a gathering of different technologies that can contribute to solve it. We briefly discuss their main features and argue to what extent they can actually achieve a confidentiality-preserving blockchain.

The rest of the paper is organized as follows: Section II gathers some surveys on blockchain privacy. Next, Section III analyzes three key characteristics of blockchain and their relation to confidentiality. The main technologies available for confidential blockchain solutions are briefly described in Section IV, and later discussed in Section V, emphasizing their relations and caveats w.r.t. blockchain. Finally, some conclusions and future work are presented in Section VI.

II. RELATED WORK

Blockchain privacy has been addressed in different works [6], [7], [8], mainly distinguishing between private payments and confidential computations ([8] also covers function privacy). However, private payments have been much more covered than confidential computations, mainly due to the maturity of the solutions. In addition, [8] states that confidential computations are much more difficult to achieve than private payments.

To achieve confidential computations in blockchain, the three works above claim NI-ZKP, FHE and Trusted Execution Environments (TEE) as the most extended building blocks, but [7], [8] also consider (briefly) MPC. In fact, [8] is the only work that deeply covers usability and interoperability of these techniques, identifying as open problems the handling of multi-user inputs (partially solved by MPC or multi-key FHE) and the development of case-specific cryptographic primitives to achieve more efficient solutions.

III. BLOCKCHAIN AND PRIVACY

This Section introduces some concepts that provide an understanding of how a standard blockchain (with public data) works and how confidential data can be related to it.

A. Blockchain state model

Roughly speaking, each node (or most of them) in a blockchain maintains a state S, which is computed from all recorded data. Each time new data x arrives on the blockchain, the state is re-computed using a state transition function $S' \leftarrow Transition(S, x)$. This is typically implemented in batches (a set of transactions forms a block) and the "checkpoints" of the state are computed using hash functions, which also link the blocks together. The specific details vary from blockchain to blockchain (in Bitcoin hashing transactions is enough, while Ethereum also maintains accounts and smart contracts). Although chaining blocks by hashing is the classic and most widespread option, there are new solutions that compact the whole state to constant size thanks to recursive ZKP[2].

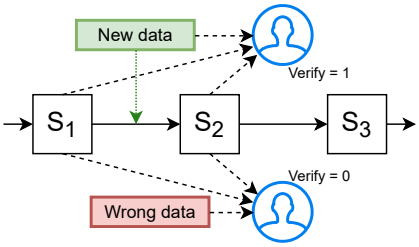


Fig. 1. The blockchain state model

Figure 1 depicts how the blockchain state evolves when new data is stored. More precisely, the new data (a block) triggers the transit from S_1 state to S_2 state. Any node in possession of S_1, S_2, and the block data can verify the correctness of the transition. In contrast, an incorrect block (due to an error or a modification attack) does not pass verification.

As for confidential data, the way the blockchain state is computed presents a first barrier, since every piece of data included in a state transition phase must be available in the verification process. Given a confidential value, including it locally in the owner's state leads to a different state from the rest of the network, losing the sense of consensus, while making it available to everyone means losing confidentiality. Confidential data can be added to the state using ciphertexts, however it is interesting to consider what value that actually adds versus storing data off-chain.

B. Blockchain storage model

Blockchain storage is problematic by nature, due to its high cost, as the ledger view must be the same for each node (data replication enables availability and eliminates deletion). Figure 2 compares a centralized storage system with a decentralized one. The centralized system allows deploying a central computer with a large amount of memory (in contrast to constrained clients) more cheaply than the decentralized one, where each node must store the same amount of information.

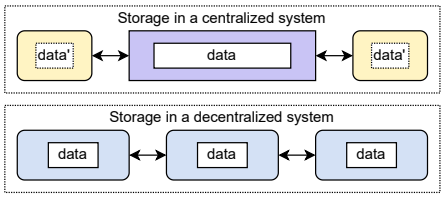


Fig. 2. Storage model in a centralized and decentralized system

In practice, different types of nodes can be deployed depending on the amount of data they store, e.g., Ethereum distinguishes between full nodes (which store all data and

[2]https://minaprotocol.com/

can verify states) and light nodes (which only store block headers and have to request data from full nodes).

Sharding [9] is a recent idea that aims to minimize the problem of replicated storage by dividing the network into logical subnets with independent data and validators, which are synchronized through a main network.

Despite sharding minimizes the exposure of data to network nodes, it does not really aim at confidentiality, but at performance, as specific data fragments can be requested if needed. In fact, the replicated storage does not pose a problem for confidentiality when using, e.g., ciphertexts, despite they will be publicly available as long as the blockchain lives, which increases the attack surface.

C. Blockchain computation model

Ethereum introduced a computational model that allows the use of data on the blockchain. Roughly speaking, its virtual machine accepts data and smart contract opcodes that enable general-purpose computations. The main difference with the centralized model is that data, contracts, and computation must be managed by each node, i.e., a node must re-compute a function to verify its correctness, leading to a secure, reliable, and expensive system.

In general, there exist two models (see Figure 3) regarding how a computation is executed in a blockchain:

On-chain. The computation is executed by the blockchain, i.e., any node executes two phases: (1) $result \leftarrow Compute(x)$, and (2) $\{0,1\} \leftarrow Verify(result, state)$.

Off-chain. The computation is not replicated, i.e., inputs and outputs can be stored on-chain, but the computation cannot be verified by the blockchain nodes.

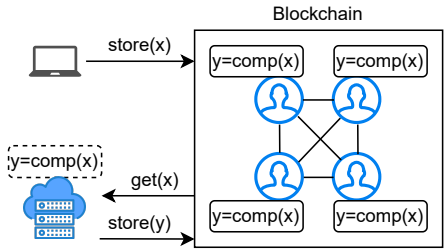


Fig. 3. On-chain (solid) vs off-chain (dashed) computation

It is easy to realize that on-chain computations are more expensive, but very secure, since to alter the result a malicious adversary must corrupt most of nodes. On the other hand, the off-chain computation model implicitly assumes a trust relation in the delegated computational party, but is cheaper. It is this area that NI-ZKP has contributed most, by storing publicly verifiable proofs of correctness on the chain.

Finally, as far as confidential computing is concerned, the on-chain model does not allow data to be protected by default, as it must be publicly available to allow verification of correctness. The only option available is to perform the computation on the user side, or to delegate it to a trusted third party, assuming they will not expose the data.

IV. TECHNOLOGIES FOR CONFIDENTIAL BLOCKCHAIN

In this Section, we gather a set of technologies that enable confidential computing and can be deployed in blockchain scenarios.

A. NI-ZKP

A NI-ZKP [10] allows a prover to convince a verifier (with one message) that a statement is true using some confidential data and without exposing it. These protocols can be formalized as follows:

$\pi \leftarrow Prove(setup, st, x, w)$: the prover generates a proof π using public data x and private data w that computes the statement st.

$\{0,1\} \leftarrow Verify(st, x, \pi)$: the verifier checks whether the statement is true when computed on x and w.

B. MPC and Proactive-MPC

MPC protocols allow a set of distrustful parties $\{P_1, ..., P_N\}$ to compute a function f on some private data $\{w_0, ..., w_N\}$ without exposing w_i to a party P_j with $j \neq i$. At the end, the computation outputs $y \leftarrow f(w_0, ..., w_N)$ as if it had been computed in clear. There exist different approaches for MPC, e.g., Garbled Circuits [11] for 2-party and Secret Sharing Schemes [12] for N-party, where security relies on the adversary inability to corrupt $t < N$ nodes. We remark that FHE (explained below) is typically understood as a specific form of MPC.

Proactive-MPC [13] is a variation in which every m operations of the computation the secret-shares are moved from a committee of holders C_1 to C_2 using a handover and re-sharing protocol. This approach limits the adversary time to corrupt parties, as secrets will not always reside in the same place.

MPC solutions for blockchain [14], [15], [16] tend to coordinate the computation on-chain and execute it off-chain, using a pool of designated nodes.

C. FHE

Roughly speaking, an FHE scheme allows to compute on ciphertexts as if it was computed on plaintexts, i.e., given $c_1 = Enc_k(m_1)$ and $c_2 = Enc_k(m_2)$, it is possible to compute $c_{add} = Enc_k(m_1 + m_2)$ or $c_{mul} = Enc_k(m_1 \cdot m_2)$. FHE was inefficiently introduced in [17], but it has been largely improved [18]. The most extended FHE settings work on public key cryptography and allow multiple clients to delegate the computation to a single server.

FHE solutions for blockchain [19] enable on-chain computations on encrypted data, however they struggle with multi-user inputs.

D. Multi-prover NI-ZKP

A multi-prover NI-ZKP [20] allows a set of parties $\{P_1, ..., P_n\}$, each with a private witness w_i, to compute a NI-ZKP in a collaborative way. More specifically, they run an MPC to compute $\pi \leftarrow Prove(st, x, \{w_1, ..., w_n\})$, where no party other than P_i learns w_i.

E. Threshold-key FHE and Multi-key FHE

Threshold-key FHE (Th-FHE) [21] is similar to public key-based FHE, but the secret key sk is secret-shared to a set of holders. The decryption process is computed interactively through an MPC, where $t \leq n$ key shares are needed to recover the plaintext. On its part, in multi-key FHE (Mk-FHE) [22] each party holds a different key pair and decryption is also done interactively using MPC.

V. DISCUSSION

The main difference between blockchain and cloud computing is that the former enables public verifiability, so achieving verifiable confidential computations should be considered. Verifiable computation is still novel, but NI-ZKP and its multi-prover version seem to be useful to add public verifiability to MPC and FHE. As for how confidential computations relate to state, we note that fully on-chain computations [19] are possible, however their difference from publicly available ciphertexts that are computed off-chain lies in additional issues, e.g., control and verification of computation steps (also input commitment and output disclosure). Th-FHE and Mk-FHE, e.g., rely on MPC for output disclosure, and key handling is not straightforward ([19] leads the blockchain to handle the decryption key, so security relies on majority honesty, i.e., FHE is reduced to MPC). On the other hand, in solutions like [15], [16], the on-chain overhead is avoided, but relating computation to state is more difficult and the benefit of replicated storage is lost.

As a summary of this discussion, we could offer an informal definition of what confidential data means in the context of blockchain: *confidential blockchain data is only accessible by designated parties, linked to the blockchain state, and verifiable by publicly available mechanisms in relation to the computation executed*. We note that this is a broad definition, difficult to achieve in its entirety and highly dependent on the specific technologies used.

VI. CONCLUSIONS

In this work, we have reviewed the blockchain model with respect to confidential data, and outlined the main lines of research and barriers to bring closer these two scenarios, which seem opposed by design. We have presented the main technologies for blockchain confidential computations (NI-ZKP, MPC, FHE, and some advanced variations), and briefly discussed their pros and cons.

As future work, we focus on providing a formal model for confidential computing in blockchain that gathers the main requirements and links them with the specific enabling technologies. We envision that it will be necessary to combine different cryptographic tools to provide sufficiently secure and usable solutions.

REFERENCIAS

[1] A. Alketbi, Q. Nasir, and M. A. Talib, "Blockchain for government services — use cases, security benefits and challenges," in *2018 15th Learning and Technology Conference (L&T)*, 2018, pp. 112–119.

[2] P. Gonczol, P. Katsikouli, L. Herskind, and N. Dragoni, "Blockchain implementations and use cases for supply chains-a survey," *IEEE Access*, vol. 8, pp. 11 856–11 871, 2020.

[3] A. Biryukov and S. Tikhomirov, "Deanonymization and linkability of cryptocurrency transactions based on network analysis," in *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2019, pp. 172–184.

[4] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous connections and onion routing," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 482–494, 1998.

[5] L. T. Thibault, T. Sarry, and A. S. Hafid, "Blockchain scaling using rollups: A comprehensive survey," *IEEE Access*, vol. 10, pp. 93 039–93 054, 2022.

[6] Q. Feng, D. He, S. Zeadally, M. K. Khan, and N. Kumar, "A survey on privacy protection in blockchain system," *Journal of Network and Computer Applications*, vol. 126, pp. 45–58, Jan. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804518303485

[7] R. Zhang, R. Xue, and L. Liu, "Security and Privacy on Blockchain," *ACM Computing Surveys*, vol. 52, no. 3, pp. 51:1–51:34, Jul. 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3316481

[8] G. Almashaqbeh and R. Solomon, "SoK: Privacy-Preserving Computing in the Blockchain Era," 2021, publication info: Published elsewhere. Minor revision. IEEE Euro S&P 2022. [Online]. Available: https://eprint.iacr.org/2021/727

[9] H. Dang, T. T. A. Dinh, D. Loghin, E.-C. Chang, Q. Lin, and B. C. Ooi, "Towards scaling blockchain systems via sharding," in *Proceedings of the 2019 International Conference on Management of Data*, ser. SIGMOD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 123–140.

[10] E. Ben-Sasson, A. Chiesa, E. Tromer, and M. Virza, "Succinct non-interactive zero knowledge for a von neumann architecture," Cryptology ePrint Archive, Paper 2013/879, 2013, https://eprint.iacr.org/2013/879.

[11] O. Goldreich, S. Micali, and A. Wigderson, "How to play any mental game," in *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, ser. STOC '87. New York, NY, USA: Association for Computing Machinery, 1987, p. 218–229.

[12] V. Goyal, Y. Song, and C. Zhu, "Guaranteed output delivery comes free in honest majority mpc," in *Advances in Cryptology – CRYPTO 2020: 40th Annual International Cryptology Conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17–21, 2020, Proceedings, Part II*. Berlin, Heidelberg: Springer-Verlag, 2020, p. 618–646.

[13] A. R. Choudhuri, A. Goel, M. Green, A. Jain, and G. Kaptchuk, "Fluid MPC: Secure Multiparty Computation with Dynamic Participants," 2020, report Number: 754. [Online]. Available: https://eprint.iacr.org/2020/754

[14] F. Benhamouda, C. Gentry, S. Gorbunov, S. Halevi, H. Krawczyk, C. Lin, T. Rabin, and L. Reyzin, "Can a Public Blockchain Keep a Secret?" in *Theory of Cryptography*, ser. Lecture Notes in Computer Science, R. Pass and K. Pietrzak, Eds. Cham: Springer International Publishing, 2020, pp. 260–290.

[15] M. de Vega, A. Masanto, R. Leslie, A. Yeoh, A. Page, and T. Litre, "Nillion network: Whitepaper," Tech. Rep., 2022. [Online]. Available: https://docsend.com/view/7bkgvzagr6ifhwrc

[16] P. Blockchain, "Partisia blockchain: Whitepaper," Tech. Rep., 2022. [Online]. Available: https://drive.google.com/file/d/1_doKDtMuY1YPDJ8LgKCiOqZvjoYkTmx4/view

[17] C. Gentry, "Computing arbitrary functions of encrypted data," *Commun. ACM*, vol. 53, no. 3, p. 97–105, mar 2010.

[18] I. Chillotti, N. Gama, M. Georgieva, and M. Izabachène, "Tfhe: Fast fully homomorphic encryption over the torus," Cryptology ePrint Archive, Paper 2018/421, 2018, https://eprint.iacr.org/2018/421.

[19] M. Dahl, L. Demir, and L. Tremblay Thibault, "Private smart contracts using homomorphic encryption," 2023.

[20] A. Ozdemir and D. Boneh, "Experimenting with Collaborative zk-SNARKs: Zero-Knowledge Proofs for Distributed Secrets," 2022.

[21] D. Boneh, R. Gennaro, S. Goldfeder, A. Jain, S. Kim, P. M. R. Rasmussen, and A. Sahai, "Threshold Cryptosystems from Threshold Fully Homomorphic Encryption," in *Advances in Cryptology – CRYPTO 2018*, 2018, pp. 565–596.

[22] P. Ananth, A. Jain, Z. Jin, and G. Malavolta, "Multi-key fully-homomorphic encryption in the plain model," Cryptology ePrint Archive, Paper 2020/180, 2020, https://eprint.iacr.org/2020/180. [Online]. Available: https://eprint.iacr.org/2020/180

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Detección de fraude en transacciones Blockchain usando procesos de Machine Learning, una Aproximación al Estado del Arte

Raúl Ocaña, Isaac Agudo, Javier López
Network, Information and Computer Security (NICS) Lab,
Universidad de Málaga, 29071
{roa, isaac, javier}@lcc.uma.es.

Las aplicaciones financieras basadas en tecnología blockchain son cada vez más comunes en el día a día de la economía regulada global. Con este precedente, y en una contextualización de la detección de fraude mediante el uso de técnicas de Machine Learning, el siguiente artículo de investigación en curso presenta una revisión sistemática de la literatura en la que se intenta dar respuesta a cuáles son las técnicas de detección más usadas actualmente y sus rendimientos, así como cuál es –si existe– la tendencia de uso de nuevas técnicas en este mismo contexto.

***Palabras Clave*—machine learning, blokchain, fraude, revisión, estado del arte**

I. INTRODUCCIÓN

La tecnología blockchain nació como la base para dar soporte al desarrollo de redes que permitiesen el intercambio de activos digitales. Esta era la idea original propuesta en el *white paper* de Bitcoin [1], y en un amplio sentido, se mantiene hasta la fecha.

Una de las características más innatas de las redes de cadena de bloques es su naturaleza transaccional. En ella se basan nuevas funcionalidades que han ido surgiendo en estos protocolos, como podrían ser la tokenización, la ejecución de un código remoto en formato *smart-contract*, o el registro de documentos digitales. Como resultado de esta evolución, las diferentes redes en activo están viendo crecer su volumen de operaciones y capitalización, alcanzando actualmente el trillón de dolares americanos [21].

Ante este tipo de situaciones, gobiernos e instituciones gubernamentales están tomando iniciativas de regulación para estas redes, las cuales están en proceso de tomar efecto o lo harán en periodos de tiempo cercano. Ejemplos de esto pueden ser la Ley MiCA [2], ratificada el 20 de abril de 2023 en el Parlamento Europeo; o el preaviso del Ministerio de Hacienda Español de la obligatoriedad de presentar registro de las operaciones realizadas en las plataformas de criptodivisas a partir de 2024, expuesta en el BOE Num. 81 [3], también de abril de 2023.

Con el objetivo de aportar valor al ecosistema investigador, y centrándonos en el estudio de la literatura hasta la fecha, se propone el desarrollo de una revisión sistemática basada en la metodología de Aproximación al Estado del Arte [4], mediante la cual se pretende dar respuesta a dos preguntas de investigación propuestas, las cuales indagan en las técnicas de Machine Learning más usadas, así como la evolución sufrida por las mismas en el contexto de la resolución del problema en la detección de fraude en transacciones blockchain.

II. METODOLOGÍA

La metodología de revisión seleccionada es la de Aproximación al Estado del Arte, que presentada por Gómez Vargas et al. [4], propone un proceso metodológico dividido en cinco fases: Indagación, Identificación, Selección, Clasificación, y Análisis. Como resultado, y esta es una de las principales diferencias con otras metodologías de revisión, se espera producir una recuperación y descripción del estado actual de la literatura, con la finalidad de transcender reflexivamente, generando una crítica en el caso de ser correspondiente.

La sección de Indagación será la encargada de dar contextualización a la temática a tratar, generando en el lector e investigador los conceptos necesarios para la correcta y objetiva evaluación de la literatura que será revisada posteriormente.

En segundo lugar, encontramos la fase de Identificación, en la que se definirán los términos PICOC, así como las preguntas de investigación y la cláusula de búsqueda que será utilizada en las bases de datos seleccionadas.

Además, continuaremos con parte de este proceso prerevisión en la tercera etapa, la Selección. En ella se propondrán los criterios de selección y filtrado, así como

se generará el cuestionario de evaluación de calidad y el de extracción de datos.

En cuarto lugar, procederemos con la etapa de Clasificación, en la que se realizará la importación de los resultados obtenidos, la evaluación de los mismos mediante el cuestionario de calidad, así como el proceso de extracción de datos.

Por último, finalizaremos con el Análisis, fase en la que se tratarán los datos obtenidos y se manipularán para su consumo. Así mismo, se discutirán los resultados obtenidos tanto cualitativa, como cuantitativamente. Como resultado, se presentará en este trabajo unas conclusiones formadas a partir de las reflexiones alcanzadas.

Cabe destacar, además, que se utilizará la herramienta Parsifal [5] para el desarrollo de esta revisión.

III. APROXIMACIÓN AL ESTADO DEL ARTE

A. Indagación

Blockchain es principalmente y en su definición más objetiva, una base de datos distribuida en la que se almacenan registros de transacciones, y la cual es mantenida y validada por una red de ordenadores repartidos por todo el mundo [6]. Esta red es la que fundamenta los principios de confianza y escalabilidad, a través de protocolos como el de consenso, o el cifrado de clave público-privada.

Fundamentadas en esta capacidad transaccional nacen las criptomonedas, que son en esencia, una unidad monetaria digital cuyo valor, a diferencia de cualquier otro activo financiero, no depende de ningún activo físico que lo respalde. Fomentar la participación, así como colaborar en el mantenimiento de la propia red, fueron las principales razones para la creación de estos activos, que sin embargo, han excedido tales responsabilidades, alcanzando en impacto la economía global y sus procesos.

Movimientos político-económicos como las regulaciones comunitarias [2] o nacionales [3] están ocurriendo, demostrando un claro interés por mantener bajo control una economía de escala que podría convertirse en parte de una transición digital.

B. Identificación

Durante el desarrollo de esta fase se ha realizado la definición de los términos PICOC, de los cuales nos hemos ayudado posteriormente para la formulación de nuestras preguntas de investigación. Una vez obtenidas las mismas, se han expandido con la inclusión de algunas palabras claves, además de sinónimos de las mismas, con el objetivo de generar una cláusula de búsqueda más completa. Por último, se han seleccionado las bases de investigación que se utilizarán durante este estudio.

1) Términos PICOC: Provenientes de población, intervención, comparación, salida y contexto, nos ayudan a entender de forma aislada cuales son los 5 factores principales que afectan a nuestro estudio, y así, acotar de una forma más semántica el proceso de generación de las preguntas de investigación. En nuestro caso, la definición realizada ha sido la siguiente:

- Población: *Software Engineers*, *Researchers*
- Intervención: *Blockchain*
- Comparación: *Machine Learning techniques*
- Salida: *Fraud detection*
- Contexto: *Cripto networks*, *Decentrilized Finance*

2) Preguntas de investigación: A través de los términos PICOC, así como teniendo en cuenta el interés en la temática de estudio a tratar, se han formulado las siguientes preguntas de investigación:

- RQ1: *Which are the main Machine Learning techniques applied to blockchain technologies for fraud detection in cripto networks?*
- RQ2: *How have fraud detection Machine Learning techniques evolved on decentralized finance blockchains applications over the years?*

3) Extensión de palabras clave: Tras la extensión de nuestras palabras clave usando sinónimos de los términos PICOC, así como las redes de criptomonedas más prolíferas actualmente, se ha generado como resultando en el siguiente listado:

- *Software Engineers, Data Engineers, Machine Learning Engineers, Researchers*
- *Blockchain, Ledger, Bitcoin, Ethereum, Tether, Dogecoin, Cardano, Polygon, Solana*
- *Machine Learning techniques, Artificial Intelligence techniques*
- *Fraud detection, Fraud prevention, Anti-fraud*

4) Cláusula de búsqueda: Usando la herramienta Parsifal, y basándonos en los términos previamente extendidos, se ha generado la siguiente cláusula de búsqueda:

(“Researchers” OR “Software Engineers” OR “Data Engineer” OR “Machine Learning Engineers”) AND (“Bitcoin” OR “Cardano” OR “Dogecoin” OR “Ethereum” OR “Polygon” OR “Solana” OR “Tether” OR “Blockchain” OR “Ledger”) AND (“Machine Learning techniques” OR “Artificial Intelligence techniques”) AND (“Fraud detection” OR “Anti-fraud” OR “Fraud prevention”)

5) Bases de datos de investigación: Para este estudio, se ha decidido utilizar dos bases de datos de investigación a las que se tiene acceso gracias a la red nacional de investigación de la que participa la Universidad de Málaga, siendo las seleccionadas:

- Scopus
- Web of Science

C. Selección

En esta sección de selección se definirán los criterios de filtrado que determinaran la amplitud de los resultados obtenidos en nuestras búsquedas. Además, estos serán los encargados de permitirnos replicar de la forma más rigurosa posible las diferentes búsquedas en las bases de datos de investigación seleccionadas.

1) Criterios de selección: Además de la cláusula de búsqueda, las bases de datos de investigación nos permiten determinar parámetros de filtrado. En este estudio, se han decidido aplicar los siguientes criterios de selección:

- Área de estudio: Ciencias de la Computación o Ingeniería del Software

- Tipo de documento: Artículo
- Keyword(s) indexada(s): "Blockchain", "Machine Learning" o "Machine-Learning"
- Idioma: Inglés
- Política de disponibilidad: Acceso público

2) Cuestionario de evaluación: Una vez realizadas las búsquedas, es necesario evaluar los resultados obtenidos, con el objetivo de asegurarnos que estos aportan valor a la temática de estudio seleccionada. Para esto utilizaremos el formulario de evaluación de calidad, formado de las preguntas de evaluación, así como tres respuestas tipo puntuadas de forma ponderada. Las preguntas definidas son las siguientes:

- Q1: *Does the document expose any kind of fraudulent action, behaviour or pattern detection?*
- Q2: *Does the document contain relation with Blockchain, Ledger, or any other mix of related technologies?*
- Q3: *Does the document expose the usage of one or more Machine Learning techniques for fraud detection?*
- Q4: *Does the document contains direct relation with at least one criptocurrency or tokenize asset?*

En cuanto a las respuestas tipo, estas tienen una puntuación asociada entre 0 y 1 puntos, pudiéndose obtener un máximo de 4 puntos y un mínimo de 0 puntos. Se ha situado el umbral de corte en 2.0 puntos, de forma que solo los artículos que superen dicha puntuación pasarán a la fase de extracción de datos. Las respuestas definidas son las siguientes:

- Yes (1.0 puntos)
- Partially (0.35 puntos)
- No (0.0 puntos)

3) Formulario de extracción de datos: En el se detallan las diferentes preguntas que se realizarán durante una segunda revisión del artículo, de la que se esperará extraer algunos de los detalles clave necesarios para el correcto análisis del mismo. Las preguntas seleccionadas han sido:

- Q1: *What is the study objective regarding fraud detection?*
- Q2: *What is the study definition of fraud for the presented use case?*
- Q3: *Is the study focused on proactive/preemptive measures to avoid fraud, or counteracting measures to analyse intents?*
- Q4: *What are the Machine learning techniques used in the study?*
- Q5: *Which was the Machine Learning technique with the best results?*
- Q6: *Which was the accuracy (%) of the most performer Machine Learning technique used?*
- Q7: *Is the most performer Machine Learning technique a supervised or unsupervised one?*

D. Clasificación

Durante el proceso de clasificación se trabajará con los resultados obtenidos de la búsqueda en las bases de datos de investigación, los cuales serán depurados y canalizados a través del cuestionario de evaluación, así como el formulario de extracción de datos.

1) Importación y eliminación de duplicados: Tras la importación de los ficheros *BibText* con las referencias de los resultados de búsqueda en la herramienta Parsifal, se realizó una primera limpieza de duplicados, que decrementó nuestro dataset de trabajo en un 25,73%: de 171 artículos inicialmente, a 127.

2) Realización de la evaluación de calidad: Tras una primera revisión, y la resolución del formulario de calidad, encontramos que solo 34 de los artículos (un 26.77% con respecto a la fase anterior) pasan nuestro umbral de corte, continuando hacia la siguiente fase.

3) Extracción de datos: Durante el proceso de extracción de datos, se completaron los registros correspondientes a las respuestas recuperadas de los diferentes artículos revisados [22]. Estos fueron transformados en tres tablas a utilizar durante el proceso de análisis. En esta fase no se produce merma en el número de artículos.

E. Análisis

El proceso de análisis se ha dividido en dos grandes bloques, atendiendo el primero de ellos al análisis de las características cualitativas y los objetivos de estudio de los artículos revisados (cuestiones Q1-Q2 del formulario de extracción de datos), mientras que el segundo se centra en el análisis más cuantitativo de las técnicas utilizadas, sus rendimientos y las tendencias temporales observadas en las mismas (cuestiones Q3-Q7 del mismo).

1) Análisis de los objetivos de estudio: En este primer bloque, se detectan tres grandes categorías en cuanto a objetivos de estudio, quedando dentro de ellas el 88.23% de los estudios revisados. Estas son: Detección de comportamientos y transacciones anómalos (38.23%), detección de estafas (29.41%), detección de vulnerabilidades (20.59%).

En el caso de la detección de anomalías, se encuentra que algunas de las definiciones encontradas hace referencia de forma explícita al comportamiento (Chuyi Yan et al. [8], Ruchi Mittal et al. [9], Xiao Fan Liu et al. [10]) del usuario en la red, sin embargo, algunos estudios desligan esta componente del patrón transaccional, atendiendo únicamente a los fines u objetivos de dichos intercambios. En este sentido, el 60% de los estudios centrados en la detección de transacciones anómalas tenían el fin de detectar patrones de lavado de dinero (Wai Weng Lo et al. [11], Johrha Alotibi et al. [12], Steven Farrugia et al. [13]).

En referencia a la detección de estafas, se encuentra un especial interés (70% revisiones) en los tipos de estafa piramidal o en planes de inversión con altos rendimientos. En estos, los estafadores aprovechan el pseudoanonimato de esta tecnología para implementar fraudes financieros [14], que sustentados en el desarrollo de contratos inteligentes, mantienen el proceso de estafa autónomamente.

El 57.14% de los estudios en esta categoría encuentra más rendimiento en la resolución de este problema a través de la clasificación mediante el uso de Redes Neuronales (Emad Badawi et al. [9], Lingyu Bian et al. [14], Yizhou

Chen et al. [15], Shuhui Zhang et al. [16]); sin embargo, en el caso del estudio de Xiajiong Shen et al. [17] se decidió transformar un claro problema de clasificación en uno de detección de anomalías.

En términos de detección de vulnerabilidades, se encuentra de especial interés en las contribuciones realizadas por Kabla et al. [18], en su implementación de un Sistema de Detección de Intrusiones con el fin de reducir las amenazas de la red Ethereum; o la de Khan et al. [19] en su investigación en la detección de ataques DDoS que amenacen con congestionar el ancho de banda transaccional de una red blockchain.

2) Análisis de las técnicas seleccionadas y sus rendimientos: Durante el estudio analítico de los resultados de los artículos revisados se detecta que estos quedan contenidos entre los años 2018 y 2023, siendo el año 2022 el que más contribuciones nos aporta: un 52.94% del total. Esto nos muestra la creciente tendencia de la temática entre la comunidad investigadora.

Es también interesante destacar que el 65.62% de los estudios aportan una solución preventiva, siendo además un 86.2% del total soluciones basadas en técnicas de aprendizaje supervisado.

Tras analizar las técnicas reportadas como mejores candidatas en cada uno de los artículos revisados, podemos ver como Random Forest (RF) se coloca en primera posición, seguida de Redes Neuronales Convolucionales (CNN), y con XGBoost (XGB) en tercera posición. El rendimiento promedio de RF es del 92%, mientras que el de XGBoost se acerca al 98%. También vemos como las CNN, que tienen un rendimiento algo superior al 96%, empiezan a jugar un papel fundamental apareciendo principalmente en artículos del año 2021 (Yizhou Chen et al. [15], Lingyu Bian et al. [14]), 2022 (Shuhui Zhang et al. [16]) y 2023 (Zijian Zhang et al. [20]), es decir, en la última mitad del periodo de estudio.

IV. CONCLUSIONES

Tras la recopilación analítica realizada, se observó que el objetivo principal del uso de las técnicas utilizadas era resolver un problema de clasificación. Dando respuesta a la primera de las preguntas de investigación (RQ1), se encuentra que Random Forest es la técnica de Machine Learning más usada; así como XGBoost, la cual presenta resultados de media un 5% mejores que Random Forest, y llegando a precisiones cercanas al 98% de acierto; o las Redes Neuronales Convolucionales, las cuales con precisiones medias cercanas al 96%, han demostrado ser muy efectivas en la extracción automática de características de clasificación.

Con respecto a la segunda de nuestras preguntas de investigación (RQ2), se encuentra la posible existencia de una tendencia en el uso de técnicas basadas en Redes Neuronales, la cual se obtiene de la distribución temporal encontrada en los artículos que utilizan esta solución de implementación. Dichos artículos suponen el 57% de la categoría de detección de vulnerabilidades, y todos ellos se encuentran aglutinados en la segunda mitad del periodo de estudio (2021-2023); aún así, parece necesaria una revisión futura para confirmar esta tendencia, asumiendo un eje temporal más largo.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por la Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía a través del proyecto BIG-PrivDATA (UMA20-FEDERJA-082) del Programa Operativo FEDER Andalucía 2014-2020.

REFERENCIAS

[1] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized business review, 21260.
[2] Legislative Observatory - European Parliament, Procedure File: 2021/0241(COD), URL: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?reference=2021/0241(COD)&l=en (visitado 29-04-2023).
[3] Agencia Estatal Boletín Oficial del Estado, BOE Num. 81, URL: https://www.boe.es/boe/dias/2023/04/05/ (visitado 29-04-2023).
[4] Gómez Vargas, M., Galeano Higuita, C., & Jaramillo Muñoz, D. A. (2015). El estado del arte: una metodología de investigación.
[5] Parsifal: About Parsifal. 2023. URL: https://parsif.al/about/ (visitado 29-04-2023)
[6] Sarmah, S. S. (2018). Understanding blockchain technology. Computer Science and Engineering, 8(2), 23-29.
[7] CoinMarketCap: Global Cryptocurrency Charts. 2023. URL: https://coinmarketcap.com/charts/ (visitado 29-04-2023)
[8] Chuyi Yan et al. Blockchain abnormal behavior awareness methods: a survey. En: Cybersecurity 5.1 (2022), pág. 5
[9] Ruchi Mittal y Mahinder Pal Singh Bhatia. Detection of Suspicious or Un-Trusted Users in Crypto-Currency Financial Trading Applications. En: International Journal of Digital Crime and Forensics (IJDCF) 13.1 (2021), págs. 79-93.
[10] Xiao Fan Liu et al. Characterizing key agents in the cryptocurrency economy through blockchain transaction analysis. En: EPJ Data Science 10.1 (2021), pág. 21.
[11] Wai Weng Lo et al. Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin. En: Applied Intelligence (2023), págs. 1-12.
[12] Johrha Alotibi et al. Money Laundering Detection using Machine Learning and Deep Learning. En: International Journal of Advanced Computer Science and Applications 13.10 (2022)
[13] Steven Farrugia, Joshua Ellul y George Azzopardi. Detection of illicit accounts over the Ethereum blockchain. En: Expert Systems with Applications 150 (2020), pág. 113318
[14] Lingyu Bian et al. Image-based scam detection method using an attention capsule network. En: IEEE Access 9 (2021), págs. 33654-33665.
[15] Yizhou Chen et al. Improving Ponzi scheme contract detection using multi-channel TextCNN and transformer. En: Sensors 21.19 (2021), pág. 6417.
[16] Shuhui Zhang et al. Ethereum Ponzi Scheme Detection Based on PD-SECR. En: Security and Communication Networks 2022 (2022).
[17] Xiajiong Shen, Shuaimin Jiang y Lei Zhang. Mining bytecode features of smart contracts to detect Ponzi scheme on blockchain. En: Computer Modeling in Engineering & Sciences 127.3 (2021), págs. 1069-1085.
[18] Arkan Hammoodi Hasan Kabla et al. Applicability of intrusion detection system on Ethereum attacks: a comprehensive review. En: IEEE Access (2022).
[19] Zulfiqar Ali Khan y Akbar Siami Namin. A Survey of DDOS Attack Detection Techniques for IoT Systems Using BlockChain Technology. En: Electronics 11.23 (2022), pág. 3892.
[20] Zijian Zhang et al. A Multi-Dimensional Covert Transaction Recognition Scheme for Blockchain. En: Mathematics 11.4 (2023), pág. 1015.
[21] CoinMarketCap - Capitalización total del mercado de criptomonedas. URL:https://coinmarketcap.com/es/charts/
[22] Referencias de artículos revisados - Tablas de clasificación. URL:https://github.com/raulocana/crypto-fraud-detection-using-ml-slr/blob/58c6feb4e4f6b397838df35809c2400f9f3a4caf/apendix_and_references.pdf

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Análisis de las vulnerabilidades en Smart Contracts: desafíos CTF para mejorar la seguridad

José María Tapia[1], Marco López[1], Isaac Agudo[1], José Carlos Ramírez[2]
[1]NICS Lab, Universidad de Málaga (Málaga), 29010.
[2]OAK Security, Munich (Alemania), 80331.
jmtapiacatena@uma.es, marcolopezg26@uma.es, isaac@lcc.uma.es, jcr@oaksecurity.io

En un mundo cada vez más digitalizado y orientado hacia la información, la seguridad se ha convertido en una prioridad crítica, especialmente en el contexto de las tecnologías emergentes, como la Web 3.0 y los SC (*Smart Contracts*, en español Contratos Inteligentes). Para abordar este reto, hemos llevado a cabo un análisis de desafíos CTF (*Capture The Flag*, en español Capturar la Bandera), con el propósito de identificar y clasificar las vulnerabilidades comunes que afectan a los SC utilizados en producción. Inspirados por el enfoque del *OWASP TOP TEN* para aplicaciones web, hemos desarrollado una clasificación basada en el Registro SWC. Esta clasificación se presenta como una herramienta introductoria para la comunidad de desarrolladores, facilitando la creación de código seguro y la prevención de vulnerabilidades conocidas y recurrentes. El propósito de este trabajo es fortalecer la seguridad en el dinámico entorno de los SC, contribuyendo así al resguardo de activos digitales y al fomento de la confianza en la tecnología blockchain.

***Palabras Clave*—Seguridad, Smart Contracts, Vulnerabilidades, Capture The Flag, Web 3.0**

I. INTRODUCCIÓN

El rol de las nuevas Tecnologías de Información (IT, del inglés *Information Technology*), está experimentando una expansión a un ritmo vertiginoso, provocando una enorme disrupción en la forma en que vivimos y trabajamos. Los avances tecnológicos y paradigmas emergentes, tales como la Web 3.0 [1], están revolucionando los servicios web y aplicaciones, marcando una era de cambio constante. La digitalización y la información se han vuelto esenciales en la modernización de industrias, el progreso económico y el bienestar social.

Sin embargo, esta transformación también ha expuesto desafíos críticos, particularmente en lo que respecta a la seguridad. Es por ello que la conciencia sobre la ciberseguridad crece día a día, y, cada vez, se van adoptando en mayor medida, acciones que dificulten a actores maliciosos llevar a cabo amenazas que comprometan la seguridad. Aun así, si la tarea de garantizar la seguridad en tecnologías de la información ampliamente adoptadas, como Web 2.0, es ardua, esto se intensifica aún más al aplicarlo a la tecnología emergente de Web 3.0 [2] , donde nos enfocamos en la utilización de SC [3]. Estos contratos, siempre en el punto de mira de atacantes, pueden gestionar el flujo de activos y el control de datos críticos, lo que los convierte en elementos fundamentales en el ecosistema. Fortalecer la seguridad en este entorno es necesario para incrementar la confianza, y proteger los activos y datos de los usuarios actuales.

En el presente artículo se pretende resaltar las vulnerabilidades asociadas a los SC [4], estableciendo una clasificación de algunas de ellas basada en el Registro SWC [5] (*Smart Contract Weaknesses and Common Vulnerabilities Registry*, en español Registro de Puntos Débiles y Vulnerabilidades Comunes de los Contratos Inteligentes), de manera análoga a como el pentesting tradicional ha abordado las diez principales vulnerabilidades de OWASP [6] (*Open Web Application Security Project*, en español Proyecto Abierto de Seguridad de las Aplicaciones Web), enumeración reconocida globlamente como *OWASP TOP TEN* [7]. Para ello, hemos centrado nuestra atención en el análisis de CTFs [8], donde hemos observado una correlación entre los desafíos recurrentes y las vulnerabilidades comunes en SC en producción.

El objetivo principal de este trabajo es proporcionar una lista de las vulnerabilidades más frecuentes en el desarrollo de SC. Proporcionando una información esencial para que los nuevos desarrolladores eviten replicar estos problemas peligrosos y comunes. Al hacerlo, no solo fortalecemos la seguridad en el dinámico entorno de los SC, sino que también contribuimos a incrementar la confianza en estas tecnologías emergentes.

II. CONTEXTO

Ethereum [9] se destaca como una de las principales plataformas blockchain en la actualidad. Esta plataforma ofrece la capacidad de ejecutar programas conocidos como Contratos Inteligentes, los cuales están escritos en el lenguaje de programación de alto nivel denominado Solidity [10]. Estos contratos se almacenan de manera inmutable en la cadena de bloques y se ejecutan cuando se cumplen condiciones predefinidas. La seguridad y confiabilidad de Ethereum se consigue mediante mecanismos de consenso [11] que establecen las reglas para que los nodos de la red puedan agregar bloques a la blockchain.

Estos mecanismos de consenso son fundamentales, ya que permiten que los SC sean ejecutados de manera confiable por una red distribuida de nodos que no necesitan confiar mutuamente entre sí, eliminando la necesidad de una autoridad externa de confianza en el proceso.

Sin embargo, la correcta ejecución de los SC no es suficiente para garantizar su seguridad, ya que no son seguros por defecto. Como exploraremos posteriormente, Solidity, el lenguaje de progamación comúnmente utilizado para el desarrollo de SC, presenta sus propias vulnerabilidades de seguridad. Por lo tanto, es crucial que los desarrolladores estén al tanto de ellas, así como de las características inherentes del ecosistema blockchain, como la ejecución determinista y la transparencia de la red. Este conocimiento es esencial para desarrollar SC seguros.

Una de las características que otorgan a los SC su atractivo para posibles atacantes es su capacidad no sólo para controlar el flujo de criptomonedas, sino también para almacenarlas. Esto implica que, mediante la correcta explotación de una vulnerabilidad, un atacante podría potencialmente hacerse con todos los activos depositados en el SC. Además, una vez que los SC han sido desplegados en la red, se vuelven accesibles para cualquier usuario. Esta accesibilidad aumenta la importancia de garantizar la seguridad de los SC.

Por lo mencionado, consideramos crucial que los desarrolladores adquieran un profundo conocimiento de los principios de seguridad del dinámico y emergente ecosistema blockchain. Esto implica no solo dominar las prácticas que permiten el desarrollo de código seguro, sino también comprender la necesidad de la seguridad en este contexto. Mediante la elaboración de nuestra lista de vulnerabilidades aspiramos a contribuir al desarrollo del conocimiento en materia de seguridad.

No obstante, también promovemos la participación en desafíos CTF para aquellos interesados en el desarrollo de SC. Un CTF, enfocado en el ámbito de la ciberseguridad, consiste en un desafío o competición cuyo objetivo primordial es resolver una serie de problemas con el fin de obtener una "bandera" o indicador específico. Los CTFs, particularmente aquellos centrados en los SC, se enfocan en la explotación de una o varias vulnerabilidades que los participantes deben identificar y aprovechar para completar cada nivel. Estos desafíos brindan un entorno propicio para aplicar habilidades de seguridad y mejorar la capacidad de identificar y mitigar vulnerabilidades en entornos reales.

III. METODOLOGÍA

Siguiendo la misma línea de pensamiento que impulsó la enumeración de *OWASP TOP TEN*, nuestro objetivo es elaborar una lista similar que se enfoque en las vulnerabilidades más frecuentes presentes en los SC. Para lograrlo, nos centraremos inicialmente en CTFs, ya que hemos identificado una correlación intrínseca entre las vulnerabilidades presentes en estos desafíos y las explotadas en SC en producción.

Para generar las estadísticas que se presentarán a continuación, hemos centrado nuestra atención en los CTFs de *Ethernaut* [12], *Damn Vulnerable DeFi* [13] y *Capture The Ether* [14], cada uno con su enfoque y complejidad distintos.

Ethernaut, seleccionado por la destacada importancia de OpenZeppelin [15] como una de las empresas líderes en materia de seguridad de SC. Esta empresa, se especializa en el desarrollo de bibliotecas y herramientas para la creación de contratos inteligentes seguros en diferentes plataformas blockchain, como Ethereum. En cuanto al CTF que propone, trata de una composición de 27 desafíos acerca de áreas como el control de acceso, la gestión de pagos, el manejo de tokens y otras funcionalidades críticas en el desarrollo de contratos seguros.

Damn Vulnerable DeFi, centrado en aplicaciones financieras descentralizadas (DeFi, del inglés *Decentralized Finance*). Basamos esta elección en la crítica importancia de comprender y abordar las vulnerabilidades en las aplicaciones financieras descentralizadas que manejan activos digitales de alto valor. Las vulnerabilidades en estas aplicaciones podrían permitir a los atacantes robar fondos, manipular precios, realizar ataques de préstamos flash, entre otros vectores de ataque. Estos riesgos pueden tener consecuencias financieras significativas para los usuarios y la integridad del ecosistema DeFi en general.

Capture The Ether, enfocado en promover buenas prácticas de codificación en SC, este CTF es esencial para comprender las vulnerabilidades comunes en la implementación de contratos inteligentes. Es crucial comprender las vulnerabilidades y debilidades que pueden existir en dichos contratos, ya que una explotación exitosa podría dar lugar a la pérdida de fondos, la manipulación de resultados o incluso el colapso completo de una aplicación descentralizada.

Para llevar a cabo el análisis de los CTFs, hemos seguido un proceso detallado de revisión y evaluación de cada desafío. Nuestro objetivo ha sido identificar y categorizar las vulnerabilidades que se presentaron como objetivos de explotación en estos desafíos. Para la categorización de las vulnerabilidades, es importante mencionar que nos hemos basado en la clasificación conocida como SWC, incorporando algunas categorías adicionales para abarcar aquellas vulnerabilidades que no se encuentran dentro del marco establecido por dicha organización. Cada categoría específica refleja la naturaleza y el impacto potencial de la vulnerabilidad.

IV. RESULTADOS PRELIMINARES

En la figura 1, se muestra una representación de los resultados obtenidos tras realizar el análisis de los desafíos presentes en Ethernaut. Este análisis revela que las vulnerabilidades más recurrentes que se han identificado en hasta 3 desafíos distintos, son las siguientes: datos privados no cifrados en la cadena, abuso de interfaces mediante la herencia entre contratos y desbordamiento aritmético de enteros [16]. En relación al desbordamiento de enteros, se ha señalado con un asterisco debido a la necesidad de aclaración.

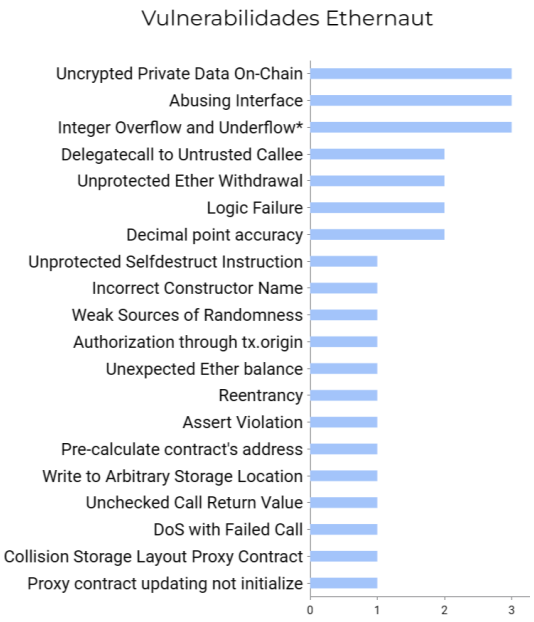


Fig. 1. Frecuencia de las vulnerabilidades encontradas en los 27 desafíos Ethernaut

En versiones anteriores a la 0.8 del compilador de Solidity, era relativamente frecuente toparse con la vulnerabilidad de desbordamiento de enteros. A menudo, esta vulnerabilidad pasaba desapercibida tanto para los programadores como para los equipos de desarrollo. Conscientes de esta problemática, los desarrolladores de Solidity implementaron una funcionalidad denominada "aritmética de enteros segura por defecto" (*default secure integer arithmetic*). Esta característica provoca una reversión en caso de que se produzca un desbordamiento de enteros, impidiendo que los atacantes puedan aprovecharse de esta vulnerabilidad. A pesar de estas medidas, todavía existen escenarios en los que el desbordamiento de enteros puede ocurrir, incluso en versiones posteriores a la 0.8 del compilador de Solidity [17]. En el caso de los desafíos presentados en la figura 1, todos ellos utilizan versiones del compilador de Solidity anteriores a la 0.8, por lo que el desbordamiento de enteros no es controlado de forma por defecto, causando una falsa sensación de seguridad para aquellos desarrolladores que usen versiones posteriores a la 0.8. Por tanto, resulta notable la ausencia de desafíos que aborden esta vulnerabilidad en versiones más recientes del compilador.

En la figura 2 se exponen las vulnerabilidades presentes en Damn Vulnerable Defi junto a su frecuencia. Es relevante resaltar el uso recurrente de Flash Loans como medio de explotación de vulnerabilidades, presente en hasta 5 de los desafíos analizados, así como su potencial utilización para la manipulación de oráculos, que se observa en 4 de los desafíos examinados.

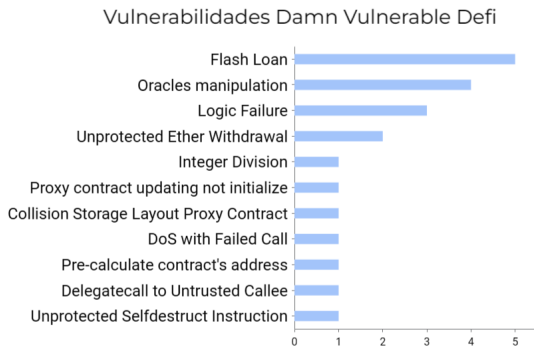


Fig. 2. Frecuencia de las vulnerabilidades encontradas en los 15 desafíos Damn Vulnerable DeFi

La utilización de Flash Loans ha ganado una creciente popularidad en los ecosistemas financieros descentralizados [18], permitiendo a particulares y empresas un acceso rápido a grandes sumas de capital sin necesidad de proporcionar garantías colaterales. Esta práctica, que sería impensable en el ámbito de las finanzas tradicionales, es factible gracias a la naturaleza determinista de Ethereum y al uso de SC. Con estos préstamos, los prestatarios pueden recibir fondos que deben reembolsarse inmediatamente a la plataforma de préstamos antes de finalizar la transacción. Si no se satisfacen las condiciones acordadas, como el reembolso mínimo del préstamo, la transacción se cancela de manera automática, sin producir ningún resultado, a excepción de la comisión de la red Ethereum, que se cobra al remitente de la transacción en forma de gas [19].

Es relevante destacar que los Flash Loans no son una vulnerabilidad por sí mismos, más bien, son una característica innovadora en el ámbito de las finanzas descentralizadas y la tecnología blockchain. En este contexto, las vulnerabilidades surgen cuando determinadas funcionalidades de un SC no han sido diseñadas para resistir el impacto de estos préstamos, lo que significa que carecen de las medidas de seguridad apropiadas para gestionar de manera segura esta interacción. Dichas vulnerabilidades pueden manifestarse debido a una falta de comprensión del ecosistema blockchain y sus especificidades, lo que a menudo conduce a la exposición de brechas en la seguridad.

En la figura 3, se ilustran las frecuencias de aparición de las vulnerabilidades identificadas en los desafíos de Capture The Ether. Los resultados ponen de manifiesto que la vulnerabilidad que se repite con mayor frecuencia, presente en hasta 6 de los desafíos distintos, es la relacionada con el "pre-cálculo de números pseudoaleatorios debido a fuentes débiles de aleatoriedad a partir de los

atributos de la cadena" [20] (del inglés, Weak Sources of Randomness from Chain Attributes).

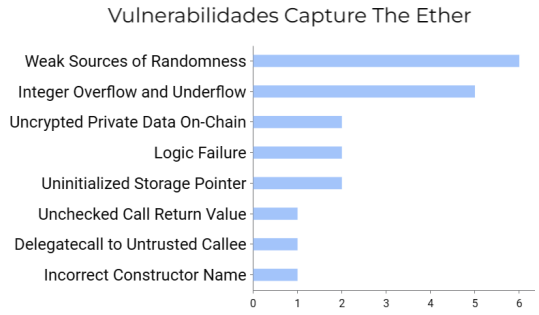


Fig. 3. Frecuencia de las vulnerabilidades encontradas en los 17 desafíos de Capture The Ether

Esta vulnerabilidad se deriva de la accesibilidad pública de la información contenida en los SC. Al fundamentar la lógica del cálculo de un número aleatorio en valores fácilmente accesibles, como el hash de la transacción anterior o el hash del bloque, un atacante puede utilizar estos valores para anticiparse y pre-calcular el número que se generará. Como resultado, el atacante puede obtener ganancias injustas, aprovechándose del funcionamiento del SC.

V. CONCLUSIONES Y LÍNEAS FUTURAS

Con el objetivo de salvaguardar los activos digitales y fomentar la confianza en la economía y los servicios que se sustentan en tecnologías blockchain, es esencial la categorización y comprensión de las vulnerabilidades comunes e inherentes a los SC. En consecuencia, nuestra labor al elaborar esta lista de vulnerabilidades busca contribuir al desarrollo del conocimiento en materia de seguridad en esta dinámica y emergente tecnología. Fomentamos activamente la realización de desafíos CTF para aquellos interesados en el desarrollo de SC. No obstante, consideramos que un reenfoque de los desafíos es esencial para que reflejen adecuadamente las vulnerabilidades predominantes en el contexto actual. Por ejemplo, la vulnerabilidad relacionada con el desbordamiento aritmético de enteros, que solo se aborda en versiones anteriores a la 0.8 del compilador de Solidity en los desafíos analizados, y no en las versiones recientes. De esta forma se minimiza la probabilidad de crear una falsa sensación de seguridad que podría surgir si las vulnerabilidades actuales no se tratan en los desafíos. Por consiguiente, nuestras futuras líneas de trabajo se centrarán en fortalecer la enumeración de vulnerabilidades a través de la realización de concursos de auditoría en plataformas como Code4rena [21] y Sherlock [22]. Esto nos proporcionará una visión más integral de las vulnerabilidades presentes en el ecosistema.

VI. AGRADECIMIENTOS

Deseamos expresar nuestro más sincero agradecimiento a la empresa OAK Security por financiar parte de este trabajo en el marco de su programa de apoyo al curso de Tecnologías Blockchain, organizado por el grupo de investigación NICS Lab en la Universidad de Málaga.

REFERENCIAS

[1] Elena Bello, "Qué es la Web 3.0 y cómo cambiará el mundo tal y como lo conocemos". Thinking for Innovation. https://www.iebschool.com/blog/web-3-0-que-es-tecnologia/.
[2] Zhuo Zhang, Brian Zhang, Wen Xu, Zhiqiang Lin, "A Systematic Study of Recent Smart Contract Security Vulnerabilities". https://niothefirst.github.io/CESC22.pdf.
[3] Elena Bello, "Smart Contracts: Qué son, para qué sirven y ventajas". Thinking for Innovation. https://www.iebschool.com/blog/smart-contract-blockchain-tecnologia/.
[4] Daniel Perez, Benjamin Livshits, "Smart Contract Vulnerabilities: Does Anyone Care?". https://allquantor.at/blockchainbib/pdf/perez2019smart.pdf.
[5] "Overview · Smart Contract Weakness Classification and Test Cases". https://swcregistry.io/.
[6] "OWASP Foundation, the Open Source Foundation for Application Security", OWASP Foundation. https://owasp.org/.
[7] "OWASP Top Ten - OWASP Foundation". OWASP Foundation, the Open Source Foundation for Application Security. https://owasp.org/www-project-top-ten/.
[8] "¿Qué es CTF? — KeepCoding Bootcamps". KeepCoding Bootcamps. https://keepcoding.io/blog/que-es-ctf-capture-the-flag/.
[9] Vitalik Buterin, "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform", https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf, 2014.
[10] Z. Wang, X. Chen, X. Zhou, Y. Huang, Z. Zheng and J. Wu, "An Empirical Study of Solidity Language Features," 2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C), Hainan, China, 2021, pp. 698-707, doi: 10.1109/QRS-C55045.2021.00105.
[11] Xie, M., Liu, J., Chen, S. and Lin, M. (2023), "A survey on blockchain consensus mechanism: research overview, current advances and future directions", International Journal of Intelligent Computing and Cybernetics, Vol. 16 No. 2, pp. 314-340. https://doi.org/10.1108/IJICC-05-2022-0126
[12] "The Ethernaut". Ethernaut. https://ethernaut.openzeppelin.com/.
[13] "Damn Vulnerable DeFi". Damn Vulnerable DeFi. https://www.damnvulnerabledefi.xyz/.
[14] "Capture the Ether - the game of Ethereum smart contract security". Capture the Ether - the game of Ethereum smart contract security. https://capturetheether.com/ .
[15] "OpenZeppelin". OpenZeppelin. https://www.openzeppelin.com/ .
[16] Kamil Polak. "Hack Solidity: Integer Overflow and Underflow - HackerNoon". HackerNoon - read, write and learn about any technology. https://hackernoon.com/hack-solidity-integer-overflow-and-underflow/.
[17] Faizan Nehal, "How Solidity 0.8 protect against integer underflow/overflow and how they can still happen in Solidity 0.8". Medium. https://faizannehal.medium.com/how-solidity-0-8-protect-against-integer-underflow-overflow-and-how-they-can-still-happen-7be22c4ab92f.
[18] "Flash Loan Attacks: Risks & Prevention - Hacken". Hacken. https://hacken.io/discover/flash-loan-attacks/.
[19] Khan, MMA, Sarwar, HMA, Awais, M. Gas consumption analysis of Ethereum blockchain transactions. Concurrency Computat Pract Exper. 2022; 34(4):e6679. doi:10.1002/cpe.6679
[20] SlowMist, "Intro to Smart Contract Security Audits - Randomness". Medium https://slowmist.medium.com/introduction-to-smart-contract-security-randomness-792cf8997599.
[21] "Code4rena". Code4rena. https://code4rena.com/ .
[22] "Sherlock". Sherlock. https://www.sherlock.xyz/ .

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Cyber Ranges: incorporación de dispositivos de encaminamiento multifabricante

Arturo Moseguí, Alan Briones, Julia Sánchez.
Departamento de Ingeniería,
La Salle Campus Barcelona – Universidad Ramon Llull (URL)
Sant Joan de la Salle 42, 08022, Barcelona.
arturo.mosegui@students.salle.url.edu, alan.briones@salle.url.edu, j.sanchez@salle.url.edu.

La revolución tecnológica ha generado nuevos desafíos en ciberseguridad. Para capacitar nuevos profesionales que afronten dichos desafíos, han surgido los Cyber Ranges, plataformas de entornos simulados que permiten a los expertos practicar y mejorar en la rama de la ciberseguridad. En este artículo se exploran los Cyber Ranges con el objetivo de investigar si es posible mejorar la capacitación actual en redes y ciberseguridad mediante la incorporación de routers de red multifabricante. La herramienta utilizada en el estudio del despliegue es el KYPO Cyber Range Platform. KYPO se empezó a desarrollar en 2013 en la universidad de Masaryk, República Checa, y es utilizada en el proyecto europeo REWIRE. Como resultado, se propone un informe sobre las capacidades de distintos routers en la plataforma KYPO.

***Palabras Clave*- ciberseguridad, Cyber Range, KYPO**

I. INTRODUCCIÓN

Los ciberataques son una amenaza constante y sofisticada para las organizaciones de todos los sectores, y es fundamental contar con profesionales capacitados para defenderse de estas amenazas. Los Cyber Ranges son entornos simulados donde los estudiantes y profesionales pueden practicar y perfeccionar sus habilidades en un entorno seguro y controlado. Al brindar un entorno de entrenamiento realista, los cyber ranges permiten a los participantes enfrentarse a situaciones y desafíos reales que pueden encontrar en el mundo cibernético. El KYPO Cyber Range Platform (KYPO CRP) [1] es una herramienta innovadora y avanzada diseñada para ofrecer un entorno de entrenamiento y simulación de ciberseguridad de alto nivel. Desarrollado por el Instituto de Ciberseguridad de la República Checa (CSIRT.CZ), el KYPO CRP proporciona a los profesionales de la ciberseguridad una plataforma integral para practicar y mejorar sus habilidades en un entorno virtual seguro y realista. Con una amplia gama de escenarios y ejercicios, el KYPO CRP permite a los usuarios simular y enfrentarse a amenazas cibernéticas reales, lo que les ayuda a desarrollar estrategias de defensa efectivas y adquirir experiencia práctica en la gestión de incidentes.

La necesidad de seguir desarrollando KYPO a nivel técnico se fundamenta en la falta de variedad de dispositivos en los entornos de entrenamiento y simulación de ciberseguridad. A medida que la tecnología evoluciona y los ciberataques se vuelven más sofisticados, es esencial que los profesionales de la ciberseguridad estén preparados para enfrentar una amplia gama de situaciones y escenarios. Sin embargo, muchos entornos de entrenamiento existentes carecen de diversidad en términos de dispositivos y configuraciones. Esto limita la capacidad de los profesionales para practicar y adquirir experiencia en entornos realistas que reflejen la complejidad y diversidad de las infraestructuras digitales del mundo real.

El artículo presenta una introducción a los Cyber Ranges y las necesidades en la ciberseguridad. Más adelante, se aborda el contexto y se revisa la literatura relacionada con el tema, centrándose en KYPO, la automatización de red y los dispositivos de red virtuales compatibles con la automatización. Seguidamente, se explora el despliegue y uso de los dispositivos de red en un entorno de Cyber Range. Se describen los pasos de comprobación utilizados para el despliegue, así como el proceso de implementación utilizando herramientas como OpenStack [2] y Ansible [3]. Además, se incluyen también los tests realizados para evaluar la funcionalidad de los dispositivos en el KYPO CRP. La sección de Resultados presenta los hallazgos obtenidos a partir de los tests realizados, analizando los datos recopilados de manera objetiva. En las Conclusiones se resumen los puntos clave discutidos en el artículo, destacando las implicaciones del estudio y proporcionando posibles direcciones futuras de investigación.

II. CONTEXTO

KYPO es una plataforma de Cyber Ranges que se empezó a desarrollar en 2013 en la universidad de Masaryk, República Checa. El objetivo principal de esta tecnología es hacer que el conocimiento de ciberseguridad llegue al máximo número de personas posibles, siendo su plataforma *open-source*.

Esta plataforma es utilizada en los proyectos europeos *CONCORDIA [4]* y *REWIRE [5]*, cuyo objetivo, entre otros, es proveer de infraestructuras y conocimientos a futuros profesionales del sector tecnológico, incluyendo la implementación de ejercicios de nivel básico para los neófitos de la ciberseguridad que también buscan ese tipo de empleos. Concretamente, KYPO CRP, ejecutándose sobre un despliegue de OpenStack, ha sido diseñada para tener la máxima flexibilidad y escalabilidad posible, minimizando los costes asociados. La mayor diferencia que tiene este Cyber Range con su competencia es que es totalmente gratuito. Como consecuencia, la única limitación que existe a la hora de desplegar esta plataforma es la cantidad de recursos necesarios para desplegar uno o varios escenarios de entrenamiento.

KYPO se basa en el uso de herramientas de automatización que desempeñan un papel fundamental en su funcionamiento. Estas herramientas permiten el despliegue de escenarios de entrenamiento y la configuración automática de los dispositivos. La automatización se refiere al uso de la tecnología para realizar tareas con mínima intervención humana. Esta capacidad se extiende a diversos sectores, pero en este trabajo se enfoca exclusivamente en la automatización dentro de los Cyber Ranges. En el ámbito de los Cyber Ranges, la automatización es de gran ayuda para garantizar su correcto funcionamiento. Gracias a ella, estas plataformas pueden desplegar una gran cantidad de equipos con configuraciones predefinidas en tiempos muy reducidos y con una intervención humana mínima. La automatización se hace presente en diferentes aspectos de estos entornos, lo que los convierte en herramientas altamente eficientes y útiles. Algunas de las áreas donde se puede encontrar la automatización en estas plataformas incluyen el despliegue de los equipos, su configuración y la posterior recopilación de datos. Mediante la automatización, los Cyber Ranges logran optimizar sus procesos y ofrecer un entorno de entrenamiento eficaz y ágil. El proceso de despliegue en KYPO consta de tres fases distintas para garantizar un funcionamiento eficiente.

En la primera fase, se crean las instancias de todos los elementos que formarán parte del escenario. Esta etapa asegura que todos los componentes necesarios estén disponibles y listos para su uso.

En la segunda fase, se aplica una configuración inicial a los dispositivos que se despliegan de forma predeterminada. Esta configuración inicial garantiza que los dispositivos estén en un estado funcional y listos para las actividades de entrenamiento.

Finalmente, en la tercera fase, se aplican las configuraciones específicas definidas por el diseñador del laboratorio. Estas configuraciones personalizadas se adaptan a los objetivos y requisitos de cada ejercicio o escenario, lo que permite una experiencia de entrenamiento más precisa y realista.

III. HERRAMIENTAS DE AUTOMATIZACIÓN

Todas las herramientas vistas a continuación nacen de la necesidad de un avance tecnológico en el área de la automatización. Con el surgimiento de los Cyber Ranges, estas herramientas han sido adaptadas a sus entornos, facilitando en muy gran medida el desarrollo y el éxito de dichas plataformas. Las herramientas vistas a continuación son: Netmiko, Nornir, Ncclient, Terraform, Ansible, Chef y Puppet.

Netmiko es una biblioteca Python de código abierto que permite a los desarrolladores automatizar y administrar dispositivos de red a través de una variedad de protocolos de comunicación, como SSH y Telnet [6].

Nornir es una biblioteca de Python de código abierto diseñada para automatizar tareas de red en múltiples dispositivos [7]. Esta herramienta proporciona una interfaz simple y fácil de usar para interactuar con dispositivos de red, utilizando protocolos como SSH, NETCONF o API REST. Nornir hace uso de la programación orientada a objetos, lo que significa que los equipos de red se representan como objetos en Python. Esto hace que sea fácil escribir scripts y aplicaciones para interactuar con múltiples dispositivos de red al mismo tiempo.

Ncclient es una herramienta fácil de usar para interactuar con equipos de red utilizando el protocolo NETCONF [8]. NETCONF es un protocolo de gestión de dispositivos de red que utiliza XML para la comunicación entre el cliente y el dispositivo de red. Su funcionamiento se basa en el modelo cliente-servidor, donde el cliente se conecta a un servidor NETCONF en el dispositivo de red y envía solicitudes utilizando XML.

Terraform es una herramienta que hace uso del IaC (*Infrastructure as Code*) para gestionar infraestructuras a partir de archivos de configuración [9]. Dichos archivos son los que utiliza Terraform para crear, modificar y eliminar los recursos necesarios para implementar esa infraestructura. Dicha herramienta está desarrollada con código abierto, siendo respaldada por grandes comunidades de colaboradores. Esta herramienta IaC puede ser utilizada con cualquier proveedor de servicios de *cloud*, eliminando la posibilidad de cualquier incompatibilidad con la plataforma de otros proyectos. En añadido, Terraform suministra una infraestructura inmutable, lo que posibilita la opción de conservar configuraciones anteriores aun habiendo actualizado la plataforma a una nueva versión. Esto permite mantener la configuración inicial pese a la acumulación de cambios en los elementos de la infraestructura.

Ansible es una herramienta que logra la automatización mediante la definición de tareas e infraestructura a través de archivos de configuración YAML [3]. Estas tareas se recogen en los *playbooks* y se ejecutan de forma ordenada. Éstas describen el estado deseado de un sistema y los pasos necesarios para llegar a ello. Esta tecnología permite incluir desde configuración de software hasta gestión de servicios y redes. Ansible tiene dos modos principales de funcionamiento: ejecución en local y ejecución en remoto. La ejecución en local tiene como objetivo realizar tareas

en el propio controlador de Ansible, es de utilidad a la hora de realizar tareas de mantenimiento o configuración. Por contraposición, la ejecución en remoto son tareas enfocadas a los dispositivos finales.

Chef es una herramienta de automatización de infraestructura que permite a los desarrolladores automatizar la creación, configuración y administración de nodos en una variedad de plataformas y nubes [10]. Su funcionamiento implica tres componentes principales. Los nodos son los servidores o máquinas virtuales que se administrarán con Chef. Cada nodo necesita tener un cliente de Chef instalado y configurado para poder comunicarse con el servidor. El servidor es el punto central de control y configuración para todos los nodos que se administran con esta tecnología. Aquí es donde se almacenan las recetas y los perfiles de Chef, que son los archivos que describen la configuración que se aplicará a los nodos. Las recetas son archivos escritos en el lenguaje *Ruby* y *Erlang* que describen la configuración que se aplicará a un nodo. Las recetas pueden contener instrucciones para instalar y configurar software, actualizar paquetes, agregar usuarios y configurar archivos de configuración. El funcionamiento es muy simple habiendo visto sus componentes, primero se crean las recetas y se importan al servidor. Seguidamente, éste se conectará a los nodos para aplicar las configuraciones.

Puppet es una herramienta de automatización similar a Chef, con el que comparte el mismo objetivo: la completa administración de nodos en una gran variedad de plataformas [11]. Tiene los mismos componentes que la herramienta Chef, los cuales tienen el mismo funcionamiento. Éstos son los agentes, servidores y módulos. Los módulos se importan a los servidores para que éste los transmita a los distintos agentes.

Una vez se han identificado las tecnologías clave en la automatización de los Cyber Ranges, es importante comprender cómo KYPO las utiliza y combina para lograr un funcionamiento óptimo. En el despliegue de laboratorios, KYPO hace uso de Terraform en el primer paso. Terraform es responsable de establecer la topología de la red en OpenStack y garantizar su correcto funcionamiento. Utilizando la información proporcionada en un archivo que describe la topología del escenario, Terraform crea las instancias de los equipos. Todos los parámetros necesarios son extraídos directamente del archivo, lo que significa que Terraform no genera ningún parámetro adicional por sí mismo. Además de crear las instancias de los equipos, Terraform también se encarga de establecer los conmutadores de red, los routers y los enlaces que conectan todos estos dispositivos.

Ansible desempeña un papel fundamental en la automatización del despliegue de escenarios en KYPO, específicamente en las tres etapas de creación de los pools. A través de Ansible, se definen todos los parámetros que Terraform utilizará en la primera etapa para indicar a OpenStack cómo desplegar los elementos del escenario. Estos parámetros se definen en un archivo llamado *topology.yml*, escrito en lenguaje YAML, y se encuentra en el repositorio del laboratorio. Además, Ansible se utiliza para definir las tareas a implementar en las dos etapas siguientes para cada uno de los dispositivos desplegados.

IV. Dispositivos de red compatibles con automatización

Una vez estudiado KYPO junto con los elementos que componen su automatización, se estudia los dispositivos de red compatibles con la automatización. Se estudian conmutadores de red, pero principalmente routers, ya que proporcionan más potencial a los escenarios desplegados.

Los *Switches*, o también conocidos como conmutadores de red, son dispositivos esenciales en cualquier topología de red. Son los encargados de operar en la capa de enlace de datos del modelo OSI y utilizan tablas de direcciones MAC para enviar paquetes de datos a su destino. Cada equipo en la red tiene una dirección MAC única, y el conmutador de red utiliza esta dirección para enviar los paquetes de datos al dispositivo correcto. A continuación, se muestran los *switches* de los principales fabricantes que soportan automatización en escenarios virtuales. Dada la inviabilidad de listar cada uno de éstos, se escogerán los más importantes y se explicará el porqué de su elección. Los dispositivos elegidos son: Cisco Nexus 1000v, Cisco Nexus 9000v, Arista vEOS, Juniper EX2200, Juniper EX8200 y Microsoft Hyper-V.

El Cisco Nexus 1000v es una versión virtualizada de un *switch* que se ejecuta en entornos de virtualización y *cloud* basados en VMWare. Éste se integra con los hipervisores para proporcionar una solución completa a un conmutador de red. Este dispositivo está dotado de seguridad en la nube, automatización e infraestructura convergente lista para operaciones. También tiene la capacidad de integración con varias soluciones de infraestructura como *vBlock*, *Flexpod*, *Cloud Services Platform* 2100 y *UCSO*. Además, aborda la seguridad del *cloud* empresarial a través del control de acceso basado en roles y el control avanzado de firewall.

El Cisco Nexus 9000v es una alternativa más potente al 1000v diseñada para ejecutarse en máquinas virtuales con hipervisor KVM. Este dispositivo cuenta con funciones básicas de capa dos y capa tres, ya que se trata de un *switch-router*. El Nexus 9000v también cuenta con una versión de hardware física, la cual utiliza el mismo sistema operativo con la gran parte de sus funcionalidades. Entre las prestaciones que presenta su versión virtual se encuentran funciones básicas de capa dos y capa tres, como VLANs, STP, LACP, OSPF, y VXLAN. Además, es altamente escalable y puede manejar un gran número de dispositivos virtuales conectados al mismo tiempo.

El Arista vEOS es un *switch-router* virtual que se ejecuta en una gran variedad de entornos de virtualización, incluyendo VMware y KVM. Proporciona una funcionalidad avanzada en la capa dos y capa tres, con múltiples protocolos de enrutamiento y protocolos a nivel dos como VLANs o STP. Este dispositivo simula un sistema de hardware físico con unos requerimientos preestablecidos. Al tener un gran parecido con su versión física, tanto las configuraciones como las prestaciones son las mismas, facilitando su configuración al poder consultar manuales en internet de ambas versiones. Además, el dispositivo puede gestionarse a través de la *Arista*

CloudVision Platform, que proporciona una solución completa de gestión y monitorización de red.

El Juniper EX2200 es un *switch-router* de la serie EX2200 que se ejecuta en un entorno virtual. Puede ser ejecutado en la gran mayoría de plataformas de hipervisores y proporciona las mismas capacidades de conmutación que su versión física. Al tratarse de un *switch-router*, tiene funcionalidades en la capa dos y capa tres, ofreciendo características avanzadas de red y seguridad. Como característica a destacar, este dispositivo cuenta con un firewall incorporado, pudiendo filtrar paquetes y aportando una capa extra de seguridad al sistema. Este firewall no es de última generación, por lo que no se puede únicamente depender de él, pero es muy útil a la hora de utilizarlo en los escenarios, ya que aporta un punto extra de personalización al laboratorio.

El Juniper EX8200 es una versión más avanzada del EX2200. La característica más importante de este equipo es su alto rendimiento y gran escalabilidad, haciendo posible la interconexión de hasta cuatro dispositivos como uno solo. La creación de una sola identidad a partir de varias satisface una amplia variedad de necesidades de red, desde redes de campus empresariales hasta centros de datos y proveedores de servicios. La plataforma de gestión y automatización de Juniper, *Junos Space Network Director*, hace que las características avanzadas de enrutamiento y conmutación, como BGP, OSPF, MPLS y VRRP sean fáciles de configurar y administrar.

El conmutador Hyper-V Virtual de Microsoft es un *switch* de red que trabaja en la segunda capa proporcionando conexión entre máquinas virtuales a redes externas. Este equipo propietario de Microsoft ofrece funcionalidades para conectar máquinas virtuales a redes físicas o virtuales, dependiendo de la necesidad del escenario. Entre los servicios que Incluye se puede encontrar el de resolución de malware, aislamiento de hosts, conformación de tráfico y de resolución de conflictos. Gracias a su compatibilidad con los controladores de Windows, el conmutador permite a los fabricantes crear complementos para proporcionar mejoras en la seguridad y en la conectividad de la red. Para poder tener un mejor control de todas las personalizaciones realizadas, el dispositivo se administra des del propio administrador de Hyper-V, el cual cuenta con una interfaz gráfica para una mejor experiencia de usuario.

Vistos los conmutadores de red, se procede al estudio de la automatización de *routers* virtuales. Los *routers* son dispositivos de red utilizados para conectar distintas redes y permitir el enrutamiento de paquetes de datos entre ellas. Trabaja en la capa tres del modelo OSI, conocida con el nombre de capa de red. Estos dispositivos hacen uso del protocolo IP para el enrutamiento a la hora de tomar decisiones sobre cómo enviar los paquetes a través de las redes. Los dispositivos destacados son: Cisco csr1000v, Cisco Catalyst 8000v, Cisco XRv9000, Aruba OS CX y Aruba VSR1000.

El Cisco csr1000v es un router virtual compatible con las principales plataformas de virtualización que proporciona una puerta de enlace WAN en servicios tanto virtuales como en la nube. El software que utiliza es el Cisco IOS XE, el cual es muy similar al que utilizan los *routers* físicos de esta misma marca. Esto elimina prácticamente la dificultad a la hora de realizar configuraciones, ya que no es necesario familiarizarse con una forma distinta de programar el dispositivo. Dada su gran versatilidad, este router se puede utilizar tanto en centros de datos como en bloques de escalabilidad, ya que puede ser modificado para limitar los recursos virtuales que utiliza. El csr1000v ofrece una amplia gama de funciones, como optimización WAN, VPN i QoS. También es compatible con una amplia gama de protocolos avanzados de enrutamiento.

El Cisco Catalyst 8000v es un *router* virtual que ofrece funcionalidades de SD-WAN y de puerta de enlace tanto en entornos virtuales como en el *cloud*. Este dispositivo utiliza el mismo software que el router anterior, el csr1000v, proporcionando la misma facilidad de aprendizaje y configuración. Cada vez más empresas se decantan más por la virtualización de sus infraestructuras y servicios para ahorrar costes. El Catalyst 8000v se ha diseñado para solventar este tipo de necesidades, tanto del enrutamiento como de la seguridad de éste. Este dispositivo es más potente que el visto anteriormente, el csr1000v, pudiendo proporcionar servicio a un mayor número de conexiones.

El Cisco XRv9000 realiza la función de router virtual en múltiples hipervisores con el software Cisco IOS XR. Este software no tiene el mismo método de configuración utilizado en los dos routers anteriores ni en los físicos proporcionados por Cisco. Los dos softwares son del mismo fabricante, por lo que tendrán una naturaleza similar, pero hay una diferencia que se destaca entre ambos: la revisión de la configuración. Mientras que en el software IOS y IOS XE la configuración se aplica de forma instantánea, en el IOS XR se han de validar los cambios una vez configurados. Esta característica proporciona un seguido de ventajas como la revisión de cada cambio sin haberse aplicado, la activación de múltiples configuraciones a la vez y la opción de deshacer los cambios en caso de que el resultado no sea el deseado. Este router ofrece una gran agilidad y eficiencia junto con una gran capacidad para escalar, permitiendo a los proveedores mejorar su excelencia operativa mediante la sustitución de los equipos físicos a virtuales. Cuenta con virtualización de funciones de red y gestión de servicios, proporcionando una alta utilidad a aquellas empresas que deseen aprovechar su máximo potencial, separación entre el plano de control y el plano de datos, aportando una capa de seguridad extra y una ayuda en la escalabilidad gracias a su sistema multinúcleo y plano de datos con gran capacidad de conmutación, permitiendo un gran flujo de información a la vez que se ejecutan tareas para asegurar la calidad de servicio o filtros de tráfico.

El Aruba OS CX *Router* es un dispositivo virtual proveniente de la familia de Aruba VSR la cual ofrece soluciones para la capa dos y capa tres del modelo OSI. Este equipo se ejecuta como una máquina virtual en un gran número de plataformas de hipervisores. Este router ofrece funciones avanzadas tanto de *routing* como de *switching* al tratarse de un equipo que trabaja en ambos modos. Este hecho hace que este dispositivo sea ampliamente utilizado en las redes virtuales, gracias a eso

empresas ofrecen guías gratuitas de implementación, como es el caso del centro criptológico nacional, el cual ha publicado un documento sobre como emplear de forma segura dicho dispositivo.

El Aruba VSR1000 es un router virtual diseñado para realizar funciones similares a los *routers* físicos de la misma gama. Está enfocado para una amplia gama de necesidades dadas sus tres diferentes versiones: con uno, cuatro u ocho procesadores lógicos. Soporta un ágil despliegue en hipervisores KVM o VMWare tanto en centros de datos como en la nube. Sus servicios de seguridad avanzada, como la VPN o el Firewall, lo convierte en una opción muy atractiva para aquellas infraestructuras de tamaño pequeño o mediano, pero no para grandes ya que pese a poder llegar a tener 8 procesadores, no tendría una potencia suficiente para un gran número de clientes. Una de sus características más importante es la licencia libre para el uso del dispositivo, no teniendo que pagar nada para poder aprovechar al máximo todos los recursos que ofrece.

La serie vMX de Juniper proporciona servicios de enrutamiento en múltiples entornos de hipervisores tanto en empresas como en la nube. Su plano de control está ejecutado por el sistema operativo Junos, el cual es el mismo que en los equipos físicos de esta marca. Como novedad al resto de routers, la serie vMX esta optimizada para ser ejecutada en entornos x86, proporcionando un mayor rendimiento que el resto. La arquitectura utilizada ofrece a las empresas desplegar instancias de *routers* según convenga para suplir la demanda de la red y eliminar los posibles cuellos de botella. Este despliegue se realiza de forma que el tráfico no se interrumpe en ningún momento, resultando en un proceso invisible para el usuario final. Al realizarse este despliegue de forma automática, se elimina la posibilidad del error humano, reduciendo las probabilidades de fallo en la red, mejorando su calidad.

V. Despliegue de los dispositivos de red en un cyber range

En este apartado se plantea la implementación mediante la automatización de distintos equipos de red en los laboratorios de KYPO. Posteriormente se utiliza el resultado obtenido para estudiar su viabilidad en ejercicios reales. El estudio se realiza pautado, describiendo cada uno de los pasos. Los pasos seguidos serán iguales para todos los dispositivos que se hayan intentado desplegar. Primero se empieza seleccionando los dispositivos con los que se realizarán las pruebas, posteriormente se lleva a cabo su despliegue en las plataformas OpenStack y Ansible. Finalmente, se combinarán los resultados para llevar a cabo su implementación en la plataforma KYPO y se analizarán los resultados obtenidos.

Como primer paso, se realiza la selección de los dispositivos los cuales son estudiados para su despliegue mediante automatización en KYPO. Dicha selección se ve afectada por ciertas restricciones que limitarán en gran cantidad el número final de equipos. Dichas limitaciones son las siguientes:

Para el estudio del dispositivo es necesario tener el archivo con su software disponible. Los únicos equipos provenientes de los fabricantes que se estudian son aquellos que disponen de licencia gratuita, para disponer de soluciones totalmente open source.

La disponibilidad de recursos virtuales también es un factor clave en el momento de realizar la selección de dispositivos. El laboratorio de KYPO cuenta con recursos limitados para el despliegue de los escenarios, tanto de memoria RAM como de procesadores lógicos. Por ese motivo se descartan aquellos equipos que consuman una gran cantidad de recursos, ya que limitaría en gran cantidad el número de laboratorios que se puedan desplegar y tampoco se aprovecharía todo el rendimiento del equipo.

Si bien es cierto que se ha realizado el estudio de conmutadores de red y routers, KYPO no contempla actualmente la posibilidad de desplegar conmutadores de red en los escenarios, únicamente de equipos con capacidad de enrutar. Es posible que en un futuro KYPO sí que permita dicha implementación, pero al estar en la actualidad restringida, solamente se estudia la implementación mediante automatización de routers.

Los equipos que se despliegan han de ser compatibles con la forma que tiene KYPO de conectar con ellos para aplicar las configuraciones. La única forma que tiene KYPO de acceder a las instancias creadas y configurarlas mediante la automatización es a través del servicio de SSH. Cualquier otro método, como el uso de las APIs, no funciona. Los dispositivos que se eligen han de ser compatibles con este método de acceso. También ha de ser compatible la forma en la que los usuarios accedan a estas máquinas a través de las instancias del laboratorio. Es importante que los alumnos que realicen los ejercicios puedan acceder a todas las instancias de éste, incluidos los dispositivos que se estudian en este trabajo. KYPO acepta dos métodos de acceso: vía consola o vía escritorio remoto. Si el dispositivo elegido no cumple ninguno de estos dos métodos de acceso, el usuario no podrá entrar a él, por lo que quedan descartados de la lista.

Una vez estudiadas las limitaciones, se procede con la creación de la lista de dispositivos los cuales serán estudiados para su despliegue en KYPO. Los dispositivos estudiados serán: Cisco csr1000v, Aruba VSR1000, Aruba OS CX y Juniper vMX vCP.

Se ha escogido el Cisco csr1000v por ser el router virtual de referencia de Cisco. Al ser un dispositivo tan versátil y popular, Cisco lo ofrece de forma gratuita en su página web oficial. La instancia creada de este dispositivo requiere bastantes recursos: un procesador lógico y 4 GB de memoria RAM. Una cantidad adecuada que permite desplegar múltiples cantidades de laboratorios con este equipo. Sus interfaces van a un gigabit por segundo, suficiente velocidad para no causar ningún cuello de botella en el escenario. Tal y como se ha visto en el apartado que estudiaba este dispositivo, al utilizar el mismo sistema operativo que las variantes físicas de Cisco, éste será fácil de configurar para los diseñadores. Se han logrado obtener múltiples imágenes de este dispositivo con distintas versiones de sistema operativo, desde versiones de hace casi siete años, como la 9.03, hasta prácticamente actuales, como la 9.17. El estudio se realiza con todas las versiones disponibles.

El Aruba VSR1000 es un router optimizado para escenarios pequeños y medianos, justo aquellos que se

realizan en KYPO. Por consiguiente, no requiere una gran cantidad de recursos y se encuentra disponible de forma gratuita para estudiantes en su página web oficial. Para que funcione correctamente dadas las características de la red donde se desplegará, la cual no cuenta con un gran numeró de clientes, solo necesita un único procesador lógico y 1 GB de RAM, haciendo a este dispositivo el que menos recursos necesita de la lista. El software ofrecido tiene tres años de antigüedad, lo cual es lo suficientemente moderno como para incluir todas las características avanzadas necesarias, como protocolos de encaminamiento o filtrado de paquetes. Por contrapartida, tal y como se ha visto anteriormente en el apartado que estudiaba este dispositivo, el VSR1000 no utiliza el sistema operativo Aruba OS CX, el cual es el mismo que el de los equipos físicos, sino que utiliza una variante. Esta variante no tiene los mismos métodos de configuración, por lo que para su implementación se requiere un pequeño estudio previo.

El router Aruba OS CX ha sido elegido para su estudio en la implementación de KYPO dada su facilidad de uso y la posibilidad de asignarle pocos recursos sin dejar de lado ninguna de sus funcionalidades. En este caso, con un único procesador lógico y 2 GB de RAM. Se ha visto en el estudio anterior que este dispositivo tiene un software muy similar a los que tienen los routers físicos de Aruba, facilitando la configuración de éstos en caso de estar familiarizado con los equipos de esta familia. Pese a no haber trabajado nunca con dispositivos de esta empresa, existen manuales en internet de configuraciones básicas que facilitan el aprendizaje de los comandos principales.

Para terminar con la selección de dispositivos, se ha elegido un router de Juniper de la serie vMX, concretamente de la versión 21.2. Esta versión ha sido publicada a inicios del 2023, haciendo de este dispositivo el más nuevo de la lista. El equipo ha sido proporcionado por La Salle con una licencia de estudiante que incluye todas las funcionalidades disponibles. El dispositivo cuenta con interfaces a un gigabit por segundo, maximizando la velocidad en la red eliminando los posibles cuellos de botella. El vMW requiere de únicamente un procesador lógico y 2 GB de RAM, siendo altamente eficiente en este aspecto. Pese a no estar familiarizados con el lenguaje de programación de los dispositivos de esta empresa, Juniper proporciona en su página web oficial distintos manuales de fácil lectura donde están descritos paso a paso un gran conjunto de configuraciones. Se han subido a la plataforma un total de seis dispositivos diferentes. Un Aruba OSCX, un Aruba VSR1000, un Cisco csr1000v versión 9.03, un Cisco csr1000v versión 9.16, un Cisco csr1000v versión 9.17 y un Juniper Junos vMX vCP.

Escogidos los dispositivos, se explica de forma ordenada los pasos realizados para subir las distintas imágenes de los dispositivos a OpenStack. Posteriormente, se crearán los *flavors*, los cuales son los elementos que definen la cantidad de recursos que se le asignará a cada instancia desplegada.

El servidor OpenStack está desplegado sobre varios nodos, obteniendo un entorno cloud que combina los recursos de múltiples equipos físicos. Éste consta de una gran cantidad de recursos, permitiendo desplegar una gran cantidad de escenarios con una amplia variedad de imágenes. Es importante destacar que todas las imágenes que se deseen subir a OpenStack para un uso posterior en KYPO deben estar almacenadas en formato QCOW2. QCOW2 es un formato de almacenamiento para discos virtuales utilizado en QEMU. Su principal característica es que separa la capa de almacenamiento física de la capa virtual mediante el uso de bloques. Esta separación entre bloques lógicos y físicos es lo que permite la creación de las instantáneas de la máquina virtual representando los cambios realizados en el disco virtual situado en otro de los bloques. Al utilizar este método de división, las instantáneas y la gestión de almacenamiento se realiza de forma mucho más óptima. Por fortuna, todas las imágenes de los dispositivos de la lista están en formato QCOW2, por lo que no existirá ningún problema de compatibilidad.

Para crear las imágenes de los dispositivos se ha de seguir el mismo método utilizado para cualquier otro tipo de equipo. Únicamente es necesario establecer un nombre, indicar el formato de la imagen, el cual siempre es QCOW2 y seleccionar el origen del archivo. Opcionalmente, se agrega una pequeña descripción para aclarar de qué tipo de dispositivo se trata. Una vez las imágenes se han subido correctamente al servidor, se definen los recursos asignados a cada uno de ellos mediante la creación de *flavors*. Se crean tres *flavors* distintos siguiendo los diferentes requerimientos de los dispositivos seleccionados. Todos ellos tienen una única CPU y la misma cantidad de disco físico, 25 GB. La diferencia se encuentra en la cantidad de memoria RAM, que varia entre 1, 2 o 4 GB.

Para implementar los dispositivos en Ansible primeramente será necesario definir un laboratorio desde sus inicios dónde se incluirán los dispositivos de la lista. Para ello se crea el repositorio en GitLab que contenga todos los ficheros vistos en apartados anteriores, los cuales definen un laboratorio en KYPO. Configurar el fichero *topology.yml* donde se define una topología lógica que incluya los dispositivos seleccionados de la lista. Finalmente, configurar el fichero *playbook.yml* y los directamente relacionados para automatizar las tareas realizadas a los equipos del laboratorio, incluyendo los dispositivos seleccionados. Como el objetivo de este artículo es estudiar la automatización de los dispositivos seleccionados en los laboratorios de KYPO, no es necesario crear grandes escenarios. Se ha optado por definir un escenario rotatorio, el cual consta de dos dispositivos finales unidos por un router, el cual rota de la lista de dispositivos seleccionados. Es decir, el escenario para los seis dispositivos es el mismo, y el router varía según el orden definido de la lista de dispositivos seleccionados.

En la Fig. 1 se puede observar la topología del escenario. Mantener el mismo escenario permite estudiar los dispositivos en igualdad de condiciones. Para los dos dispositivos finales se eligen dos máquinas virtuales que no sean muy invasivas ni requieran grandes cantidades de recursos, ya que tienen poca relevancia en este estudio. Por consiguiente, se han escogido dos distribuciones de Ubuntu.

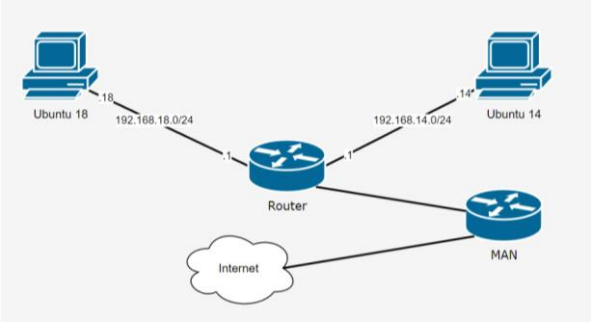


Fig. 1. *Topología del escenario*

Una vez definida la topología del escenario en el fichero *topology.yml* se define el archivo *playbook.yml*, el cual contiene todas las tareas que ansible automatizará en las instancias desplegadas en la tercera fase.

Estas tareas varían en función del router elegido, no obstante, aquellas tareas destinadas a los dispositivos finales permanecen siempre iguales. La tarea enfocada a los dispositivos finales es la que habilita acceso de los usuarios a éstos. Esta tarea la proporciona KYPO con el rol *kypo-user-access.* El rol establece una conexión mediante *cloud-init* a los equipos seleccionados y crea un usuario según los parámetros establecidos. Sin esta tarea sería imposible acceder a las instancias de los dispositivos finales, ya que éstos no tendrían ningún usuario definido. Una vez definida la tarea para los dispositivos finales, se definen aquellas destinadas al router, las cuales irán variando según el equipo.

En cuanto a la segunda fase de Ansible, no hay que realizar ninguna modificación ya que es la misma para todos los escenarios desplegados. Ésta se encarga de establecer una configuración inicial a todos los equipos y activar el funcionamiento de aquellos elementos esenciales para el escenario. Entre estos elementos se pueden encontrar los enlaces lógicos de las instancias y el router *MAN* encargado de conectar el escenario a internet.

Para automatizar las configuraciones de los routers se usan los módulos predefinidos de Ansible que se encuentran en el controlador.

Para los dispositivos Aruba se ha seleccionado el módulo *community.network.aruba_config* . Éste permite, mediante el uso de bloques simples, enviar comandos de configuración a los equipos Aruba. Este módulo está diseñado para todos los dispositivos Aruba, por lo que es utilizado para los dos equipos de la lista que tienen este fabricante. Con este módulo se ejecuta una simple tarea para comprobar el correcto funcionamiento e integración de los dispositivos Aruba con la plataforma KYPO.

Para los routers de la marca Cisco se ha elegido uno de los módulos predefinidos en el controlador de Ansible especialmente diseñado para estos dispositivos. El módulo llamado *cisco.ios.ios_config* se encarga de, mediante el uso de bloques simples, enviar configuraciones a los dispositivos en forma de comandos. Este módulo es el utilizado para los tres routers Cisco, al ser los tres del mismo modelo, pero con diferentes versiones. Ambos módulos, tanto el de Aruba como el de Cisco, parten de la misma base, la cual es Python con la librería *paramiko*. Al hacer uso del mismo lenguaje de programación y de la misma librería, la configuración de cara al usuario es muy similar, al igual que el método de conexión. Con este módulo se ejecuta una simple tarea para comprobar el correcto funcionamiento e integración de los dispositivos Cisco con la plataforma KYPO.

Igual que en el caso de los dispositivos Aruba, si la tarea termina con éxito, se valida el módulo y la correcta conexión entre Ansible y la instancia del router.

Para el dispositivo Juniper de la lista de routers seleccionados, se ha elegido un módulo de Ansible predefinido instalado en el controlador. Este módulo está especialmente diseñado para los dispositivos Juniper de la serie Junos. La función de este módulo, llamado *junipernetworks.junos.junos_command*, es encargarse de enviar comandos de configuración mediante bloques simples a los dispositivos Juniper de la serie Junos. En caso que Ansible ejecute correctamente la tarea y se vea la versión del dispositivo en el registro de Ansible, se validará el módulo y la conexión entre Ansible y la instancia desplegada.

A partir de las configuraciones preparadas para cada uno de los dispositivos, se ha procedido a realizar el estudio del despliegue de cada uno de ellos en la plataforma KYPO. Cada uno de los dispositivos ha mostrado respuestas distintas al despliegue con el módulo inicial. Gracias a dichas respuestas se ha ido variando la configuración establecida para los dispositivos, realizando distintas pruebas para lograr desplegarlos en KYPO mediante el uso de la automatización. Dado el gran número de pruebas realizadas para obtener los resultados de este trabajo, no se explica en este documento, por lo que se exponen directamente los resultados obtenidos.

VI. Resultados

Para hacerlo más visual y mejorar la comprensión, se define una tabla (Tabla I) con cada una de las conclusiones extraídas de cada dispositivo de la lista. Cada dispositivo es comparado en las cinco categorías principales de este trabajo.

Tabla I
Resultados Obtenidos

Dispositivo	Aruba OSCX	Aruba VSR1000	Cisco csr1000v	Junos vMX vCP
Compatible con OpenStack	Sí	Sí	Sí	Sí
Compatible con Ansible	Sí	Sí	Sí	Sí
Automatizable desde Ansible	No	No	No	Sí
Configurable desde KYPO	Sí	Sí	No	No
Enruta en KYPO	No	Sí	No	No

La primera categoría es la compatibilidad con OpenStack, indicando si el dispositivo es compatible con su uso en OpenStack. La siguiente es la compatibilidad con Ansible, indicando si Ansible da soporte a la imagen. El dispositivo es soportado si existen módulos, tanto en internet como preinstalados, enfocados a la versión del

equipo. La automatización desde Ansible indica si Ansible ha sido capaz de establecer conexión con el dispositivo para aplicar configuraciones de forma automática. El apartado de configuración desde KYPO indica si, una vez desplegada la instancia del router en el escenario del laboratorio, ha sido posible acceder a ella para aplicar configuraciones. Para finalizar, el enrutamiento en KYPO indica si, después de configurar el router para que realice su correcta función, éste ha sido capaz de realizarla correctamente, cumpliendo con su función como *router*.

Observando los resultados de la Tabla I se obtiene que el dispositivo que mejor se ha adaptado al entorno de KYPO es el router Aruba VSR1000, el cual ha sido capaz de enrutar paquetes a través del escenario e internet pese a no haber sido compatible con la automatización mediante Ansible.

VII. CONCLUSIONES

Se han observado los precedentes para el surgimiento de los Cyber Ranges. Por un lado, el aumento del número de ciber ataques y su complejidad han requerido la creación de una mejor herramienta para preparar a los ingenieros ante este ciber crimen. Por otro lado, las tecnologías existentes para el aprendizaje de ciberseguridad se estaban empezando a quedar obsoletas.

El objetivo del artículo, una vez definidos los Cyber Ranges y visto su origen, es el de estudiar las tecnologías de automatización que utilizan los Cyber Ranges para desplegar los escenarios. Estas tecnologías proporcionan una clara ventaja respecto a sus competidores, facilitan el uso del Cyber Range y lo sitúa a la vanguardia de la tecnología. Una vez vistas las tecnologías, se ha estudiado la implementación de distintos routers en los escenarios virtuales de KYPO.

Se analizaron los precedentes de los Cyber Ranges, estudiando los puntos positivos y negativos de cada uno de ellos. Con estos datos se ha visto porqué éstos no podían evolucionar al mismo ritmo en cuanto a complejidad que los ciber ataques reales, motivo por el cual se desarrollan los Cyber Ranges.

KYPO ofrece una solución Cyber Range sin ningún coste, aprovechando tecnologías open-source. Gracias a ello, el coste de la implementación de la plataforma se enfoca únicamente en el servidor en el que va a ser desplegado. Cuantos más recursos se le quiere asignar, más costes conlleva.

A fin de conocer más el funcionamiento de KYPO, se han analizado las tecnologías de automatización involucradas. Se ha visto como estas trabajan en sincronización para lograr el correcto funcionamiento del Cyber Range. Posteriormente y, a modo de parte práctica, se ha realizado el estudio de la implementación de distintos routers de fabricantes de tecnologías de red, como Cisco, Aruba y Juniper. Mediante este estudio no solo se han visto las capacidades de los routers en los escenarios desplegados, sino que también se ha analizado la capacidad de KYPO para integrar estos dispositivos.

En conclusión, es posible mejorar la capacitación actual en redes y ciberseguridad mediante la incorporación de dispositivos de red multifabricante en plataformas Cyber Range al existir la capacidad de implementar dichos dispositivos en los escenarios. Con el despliegue de estos equipos, se logra un escenario más completo y realista.

Como líneas futuras del trabajo, se seguirán investigando nuevos dispositivos virtuales que vayan saliendo al mercado para descubrir si son adecuados para su incorporación al Cyber Range de KYPO. También se considera interesante no dejar de lado el estudio de switches virtuales automatizables a la espera de que puedan ser implementados en KYPO o por si se opta por una nueva plataforma de Cyber Range que los soporte.

AGRADECIMIENTOS

Este trabajo se ha llevado a cabo gracias a la financiación recibida por el proyecto *Cybersecurity Skills Alliance – A New Vision for Europe (REWIRE)* - Grant Agreement 621701-EPP-1-2020-1-LT-EPPKA2-SSA-B, el cual tiene como objetivo principal construir un *Blueprint* para la industria de la Ciberseguridad y una estrategia europea concreta para la adquisición de habilidades en Ciberseguridad.

REFERENCIAS

[1] KYPO Docs. KYPO. KYPO Cyber Range Platform (muni.cz)

[2] Openstack Software. Openstack. Open Source Cloud Computing Infrastructure - OpenStack

[3] How Ansible works. Ansible. How it works (ansible.com)

[4] CONCORDIA service board. CONCORDIA. Home : CONCORDIA (concordia-h2020.eu)

[5] REWIRE project. REWIRE. REWIRE (rewireproject.eu)

[6] Netmiko Library. PyNet. Python for Network Engineers - Netmiko Library (twb-tech.com)

[7] An Introduction to Nornir. PyNet. An Introduction to Nornir (twb-tech.com)

[8] Ncclient. Read The Docs. Welcome — ncclient 0.6.9 documentation

[9] Terraform. IBM. ¿Qué es Terraform? | IBM

[10] Emulab Chef Tutorial. Emulab. 15 Emulab Chef Tutorial

[11] Puppet Overview. Puppet. Introduction to Puppet

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVIJornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Análisis forense de conversaciones de WhatsApp

Xabiel G. Pañeda, David Melendi, Victor Corcoba, Dan García, Roberto García, Alejandro G. Pañeda
Departamento Informática
Universidad Oviedo
Campus de Gijón, Gijón/Xixón, Asturies, España
xabiel@uniovi.es, melendi@uniovi.es, corcobavictor@uniovi.es, garciadan@uniovi.es, garciaroberto@uniovi.es

No hace mucho tiempo era normal ver cómo en los procesos judiciales se aportaban copias impresas de evidencias digitales. Obviamente, estas impresiones no se acompañaban de algún respaldo técnico que corroborara su veracidad. Sin embargo, en los últimos dos años esta costumbre se ha invertido, pidiendo los abogados que todas evidencias se incorporen en un informe pericial que garantice su veracidad y proceso de extracción. En este artículo analizamos el proceso de extracción de una conversación de WhatsApp para su presentación en un proceso judicial, indicando las posibles situaciones anormales e incongruencias que se pueden dar en ciertas ocasiones y que el experto deberá manejar. El objetivo es estudiar qué cuestiones pueden ser certificadas por el analista forense como no falsificables, qué discrepancias o incongruencias podrían aparecer entre dos extracciones y sobre qué elementos no se puede garantizar la veracidad. El resultado es que el texto de una conversación no es manipulable, pero sí es posible que haya habido mensajes que ya no aparezcan en una exportación, difiriendo la captura obtenida de las evidencias posteriores que se puedan recoger. Además, para algunos elementos, como las votaciones, no se puede garantizar que el resultado que se ve en una exportación sea el resultado que se produjo en su momento.

***Palabras Clave*- WhatsApp, forensia, evidencia digital, informe pericial, incongruencias en evidencias**

I. Introducción

La forma de presentación de las evidencias digitales en los procesos judiciales ha cambiado en los últimos años. De presentar copias impresas sin ningún aval técnico, se ha pasado a tener que aportar las evidencias justificando su autenticidad y proceso de extracción. Este cambio se debe a que el número de impugnaciones de evidencias digitales se ha disparado y los abogados no quieren verse en la situación de tener que actuar en el último momento, sin garantías de poder corroborar lo previamente aportado. Por ello, buscan presentar la evidencia con un buen refrendo que evite cualquier puesta en tela de juicio. El correo electrónico, las publicaciones en redes sociales y los servicios de mensajería instantánea son medios de comunicación muy utilizados en la actualidad. Por ello, tienen una amplia presencia en el contexto judicial.

Dentro de este tipo de medios de comunicación digital destaca el servicio WhatsApp. El importante número de usuarios (con una cuota de mercado en España que supera el 90%), su facilidad de uso y la forma en la que se utiliza para conversar, discutir o criticar lo han vuelto un elemento estrella en el contexto judicial. Un número muy relevante de procesos incluye conversaciones de WhatsApp que se utilizan como evidencia de algún acontecimiento, opinión o acuerdo.

En este artículo analizamos el proceso de extracción forense de las conversaciones de WhatsApp. Este servicio ya cuenta con un sistema diseñado para poder hacerlo, que permite una extracción sencilla. Sin embargo, el proceso puede dar resultados inesperados y otros que lleven a contradicciones entre extracciones realizadas desde diferentes dispositivos participantes en una misma conversación. El analista forense debe de saber interpretar la información obtenida, ser capaz de determinar qué puede certificar como auténtico, a qué se deben las posibles discrepancias y qué es lo no puede garantizar que sea fidedigno.

El resto del artículo está organizado de la siguiente forma. En la Sección II se hace un recorrido por los principales trabajos relacionados. La sección III describe el proceso de extracción de una conversación de WhatsApp. La sección IV analiza las situaciones inesperadas y posibles incongruencias que se pueden generar entre dos extracciones de la misma conversación. Por último, la sección V presenta las principales conclusiones del estudio.

II. Trabajos Relacionados

En la actualidad WhatsApp es una de las aplicaciones de mensajería instantáneas más utilizadas en el mundo [1]. Esto implica también que sea un elemento presente en muchos crímenes y que, por tanto, sea importante disponer de herramientas para obtener evidencias sobre su uso. Muchos investigadores han analizado y propuesto diversas

metodologías para extraer información. Sin embargo, debido a las constantes actualizaciones de este software y de los sistemas operativos donde se ejecuta es necesario disponer de estrategias actualizadas.

La mayor parte de los trabajos que encontramos en la literatura se centran en examinar qué información se puede extraer de la base de datos SQLite que usa WhatsApp en Android para almacenar la configuración y los datos del programa [2]. En [3] los autores exploran las herramientas existentes para extraer información y realizar un análisis forense. Los investigadores destacaron la necesidad de hacer un preprocesamiento de los datos debido a su gran tamaño para poder posteriormente completar el análisis forense.

Una gran parte de los usuarios de esta aplicación utilizan la versión web para acceder al servicio desde su ordenador. En [4] los autores realizaron un análisis forense para identificar qué evidencias se podían extraer de WhatsApp en su versión web y compararlo con la aplicación móvil para Android. En la aplicación móvil consiguieron obtener la fecha y hora de las conexiones, el registro de llamadas, los mensajes de texto, fotos y videos. Sin embargo, las evidencias en la aplicación web fueron más limitadas y sólo lograron extraer la fecha y hora de las conexiones realizadas por el usuario. En [5] los investigadores lograron recuperar más información usando la versión web. Concretamente fueron capaces de obtener las fechas y hora de las conexiones, el historial de mensajes, los archivos de audio y video y el número de teléfono de la víctima y el atacante analizando la memoria RAM.

WhatsApp añadió recientemente una funcionalidad que permite eliminar mensajes que han sido enviados tanto en el emisor como en el receptor. En [6] los investigadores propusieron e implementaron una herramienta denominada SQLWSP para recuperar este tipo de mensajes. La solución se basa en análisis de la estructura de una base de datos SQLite, que es la que utiliza WhatsApp para almacenar sus datos tales como los mensajes. Para usar esta herramienta es necesario obtener previamente la base de datos de WhatsApp. En el trabajo los autores describieron los pasos seguidos obtenerla en teléfonos que estuviesen previamente *rooteados*.

En la mayor parte de la literatura uno de los requisitos para poder extraer evidencias es que el dispositivo este *rooteado*. Esto podría suponer un problema ya que implica la modificación de una evidencia. En [7] se propuso realizar una copia lógica de la base de datos de WhatsApp y posteriormente desencriptarla mediante las herramientas DB Extractor y WhatsApp Viewer. Con esta metodología, los autores del trabajo consiguieron evidencias sobre la fecha y hora de las conexiones, el texto de los mensajes y el teléfono del usuario. En [8] los autores desencriptaron las bases de datos revirtiendo a una versión anterior de WhatsApp para obtener la clave de cifrado. Una vez desencriptada la base de datos se centraron en analizar los datos almacenados en ella siendo capaces de reconstruir la lista de contactos e identificar cuales estaban bloqueados. También recuperaron otras evidencias tales como: los mensajes textuales y no textuales, su cronología, las llamadas de audio y video y el contenido mostrado usando la función “estado”.

Como conclusión, la mayor parte de los trabajos en este ámbito se basan en la base de datos que utiliza WhatsApp en su versión Android para almacenar la información. El acceso a este archivo está cada vez más restringido en las versiones más recientes del sistema operativo y del propio programa. Además, hay muchos usuarios que emplean smartphones con el sistema operativo IOS. Por lo tanto, es necesario desarrollar metodologías multiplataforma para la recuperación de evidencias y que no requieran la manipulación del dispositivo ni el uso de versiones obsoletas. Este es el objetivo de nuestro trabajo.

III. Proceso de Extracción

El proceso de extracción de una conversación de WhatsApp se realiza utilizando la opción “Exportar Chat” que la propia aplicación incluye tanto en sus versiones de iPhone como de Android. A día de hoy, la aplicación de ordenador no incluye esta funcionalidad. Esta forma de obtención de la evidencia será la más directa y recomendable, ya que una extracción accediendo a los ficheros en el teléfono tendrá el problema de que la base de datos de mensajes de WhatsApp se almacena en un formato cifrado. La extracción de estos ficheros y su decodificación tiene una gran de dependencia del modelo de teléfono, las adaptaciones del sistema operativo y las actualizaciones de WhatsApp, por lo que el acceso a los mensajes se convierte en una lotería. Esto hace que este método no pueda ser utilizado para procesos judiciales a nivel general.

La exportación de un chat puede realizarse únicamente del texto de la conversación o incluyendo los elementos multimedia transmitidos. Esto incluye fotografías, audios, vídeos, ubicaciones y contactos. La exportación de solo texto dará lugar a un fichero con formato txt que incluirá el texto enviado. En el caso de la exportación con multimedia, se generará un fichero comprimido que incluye un fichero txt con el texto de la conversación y los ficheros multimedia transmitidos. En su mayoría, los ficheros multimedia no contienen metainformación relevante, ya que la descripción EXIF es eliminada por WhatsApp antes de su envío. Sin embargo, los metadatos RDF sí permanecen en los ficheros PDF, al igual que la metainformación de los ficheros Office. Dentro de los ficheros multimedia exportados no estarán todos aquellos que hayan sido previamente eliminados para todos los participantes en la conversación o aquellos eliminados únicamente para el usuario desde cuyo móvil se realiza la exportación. Un ejemplo de exportación es el que se muestra a continuación.

```
[21/2/23, 19:00:26] Xabiel: <adjunto: 00000002-
PHOTO-2023-02-21-19-00-26.jpg>
[21/2/23, 19:53:02] Melendi: 👍👍
[21/2/23, 19:53:19] Xabiel: Mandevos el formatu
de contratu
[21/2/23, 21:30:54] AG Pañeda: podría buscar qué
info tengo
[21/2/23, 22:00:37] Xabiel: Nun fae falta. Yera
pa la vista
[21/2/23, 22:02:45] Melendi: miraste lo de les
llicencies?
[21/2/23, 22:03:02] Xabiel: Si
[21/2/23, 22:15:44] Xabiel: <adjunto: 00000048-
STICKER-2023-02-21-22-15-44.webp>
```

```
[21/2/23, 22:55:55] Xabiel:
https://www.microsoft.com/es-es/microsoft-
365/business/compare-all-microsoft-365-business-
products?&activetab=tab:primaryr2
```

En este ejemplo se puede ver cómo cada mensaje se muestra asociado a un identificador y a la fecha y hora en la que fue enviado. Una cuestión importante es que el servicio de WhatsApp adapta las horas y fechas a la franja horaria del teléfono en el que se registra la exportación en el momento que se realiza. Esto debe ser tenido muy en cuenta en situaciones con participantes que se encuentran en diferentes franjas horarias y teléfonos que se mueven de un país a otro. Otra circunstancia que es importante resaltar es que las exportaciones en Android y iPhone tienen pequeñas diferencias en el formato del archivo txt que no deben de ser interpretadas como una falsificación. En la figura 1 se muestran algunas de estas diferencias. Se ha resaltado el identificador de cada participante. Este elemento tiene relevancia porque puede ser modificado a lo largo del tiempo, apareciendo en la exportación tal y como estaba configurado en el momento de su generación. Por eso es importante referenciar los participantes en una conversación con el número de teléfono y no con el identificador, dado a que este podría haberse alterado. En el caso de que alguno de los participantes no esté registrado en la agenda del teléfono desde el que se hace la exportación, en ésta aparecerá su número de teléfono.

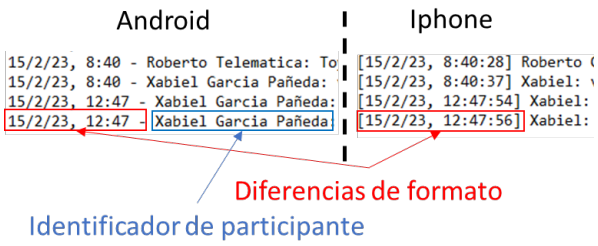


Fig. 1. Formato de fichero txt

En lo que se refiere a la exportación única del texto de la conversación, ésta incluirá un aviso de "imagen omitida", o similar dependiendo del tipo de fichero, en el punto donde se hubiera transmitido un elemento multimedia. En algunos casos, incluso podrá aparecer el nombre del documento omitido, tal y como se muestra en el siguiente ejemplo:

```
[30/12/22, 13:28:51] Roberto Garcia Fernandez:
Carta de invitación jornada de ciberseguridad
2023.docx documento omitido
```

En el caso de la exportación que incluya los elementos multimedia, se indicará el nombre del fichero transferido. El nombre de fichero incluirá su tipo y una serie de dígitos que representan la fecha de transmisión. A continuación, se muestra un ejemplo del envío de un *sticker*:

```
[21/2/23, 9:15:44] Xabiel: <adjunto: 00000048-
STICKER-2023-02-21-09-15-44.webp>
```

Ambas exportaciones tienen límites en cuanto al número de mensajes máximo que puede ser exportado [9]. La exportación del texto solamente está limitada a un máximo de 40.000 mensajes. Por otro lado, la exportación que incluye multimedia está limitada a un máximo de 10.000 mensajes. En caso de que la conversación supere estos límites, esta será amputada por la parte más antigua.

A. Exportación de encuestas

La exportación de las encuestas tiene algunas características especiales. Por un lado, a pesar de incluirse con su creación en el botón "+" donde se encuentran los elementos multimedia, su representación en la exportación es puramente textual. Los participantes podrán votar a todas las opciones, por lo que el número de votos máximo puede ser igual al doble del número de participantes. En la exportación se genera un mensaje denominado "ENCUESTA". A continuación, aparecerán los votos para cada una de las opciones sin indicar quién ha sido el votante, tal y como se muestra en el siguiente ejemplo.

```
[25/2/23, 9:22:00] Xabiel: ENCUESTA:
Vota una opción (quiero ver qué pasa)
OPCIÓN: opción 1 (1 voto)
OPCIÓN: opción 2 (1 voto)
```

No será posible ver qué participante emitió cada voto, a pesar de que en la aplicación sí es posible hacerlo a través de la opción de visualización de votos. Puesto que las encuestas no tienen instante final, los votos pueden ser modificados por los participantes todas las veces que quieran. Esto tiene una gran importancia desde el punto de vista forense.

B. Envío de imágenes y videos

WhatsApp permite enviar imágenes y videos que podrían ser utilizados para cometer algún tipo de delito como amenazas. El usuario puede eliminar el elemento multimedia que ha enviado o recibido en su teléfono usando la opción "Eliminar para mí". En el caso de ser el emisor del contenido multimedia también puede borrarlo para los destinatarios usando la opción "Eliminar para todos". Si se emplea la opción "Eliminar para mi" al exportar el chat no aparece ninguna información sobre el envío del elemento multimedia. Por el contrario, si se usó la opción "Eliminar para todos" se muestra:

```
[3/3/23, 19:26:21] Víctor: Eliminaste este
mensaje.
```

WhatsApp incorporó en agosto de 2021 una opción que permite compartir fotos y vídeos de manera que solo pueden ser visualizados por parte del receptor una sola vez. Los mensajes de este tipo son identificados con un "1" y no se almacenan en la galería del teléfono ni se pueden recoger mediante capturas de pantalla. Además, sólo se

pueden enviar y visualizar desde dispositivos móviles. En la figura 2 se observa el mensaje que se muestra en pantalla cuando se recibe una de estas imágenes. Por otro lado, en la figura 3 se observa el comportamiento del sistema cuando se intenta realizar una captura de pantalla en la que aparece una de estas imágenes.

Fig. 2. Mensaje cuando se recibe una foto o vídeo de visualización única desde la aplicación web

Cuando se realiza la exportación, aparece una indicación de imagen omitida, tanto en el modo de exportación de solo texto como en el modo multimedia. Si el envío de la foto o vídeo se hizo sin la opción de un solo uso y no se elimina el mensaje, en el modo de exportación multimedia aparecerá la referencia al archivo enviado. A continuación, se muestra el resultado de la exportación de un chat en modo multimedia, en el que se envió una imagen de visualización única y otra imagen sin esta opción:

```
[28/2/23, 17:34:29] Víctor: imagen omitida
[28/2/23, 17:42:43] Víctor: <adjunto: 00000003-
PHOTO-2023-02-28-17-42-43.jpg >
```

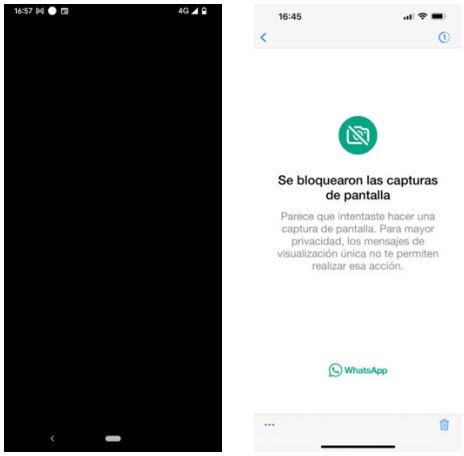


Fig. 3. Mensaje cuando se intenta hacer una captura de pantalla de una foto de visualización única en Android (izquierda) y IPHONE (derecha)

IV. SITUACIONES ANÓMALAS O INCONGRUENCIAS

A. *Mensajes de texto vinculados a elementos multimedia*

En muchas de las extracciones forenses, lo relevante es el texto que se ha intercambiado y no tanto los ficheros multimedia transmitidos. Para facilitar el manejo, es posible que se realice una extracción únicamente del texto de la conversación. Esta decisión tiene un aspecto que ha de ser tenido en cuenta. Recientemente, WhatsApp ha incorporado la posibilidad de asociar un texto al envío de un fichero multimedia. Este texto se presentará al destinatario justo debajo de la imagen o icono del fichero multimedia. Es relevante porque cuando se selecciona la opción de exportar solo texto, únicamente aparecerá la referencia <Multimedia omitido> purgando el fichero multimedia y también el texto asociado. Incluso es posible que el fichero multimedia no se haya exportado por superar el tamaño máximo de la exportación pero que sí lo haga su texto asociado. Esto puede generar incongruencias entre los dos tipos de exportación. A continuación se muestran dos fragmentos de una misma conversación exportados incluyendo los ficheros multimedia y sin incluirlos. Se puede observar que con el primer fragmento (exportación multimedia) se incluye el texto asociado a un vídeo, a pesar de que este fue omitido.

```
1/9/22, 16:33 - Victor: Se feliz
1/9/22, 16:33 - <video omitido>
Poco más que añadir
1/9/22, 16:33 - Victor: No crees?
```

```
1/9/22, 16:33 - Victor: Se feliz
1/9/22, 16:33 - <video omitido>
1/9/22, 16:33 - Victor: No crees?
```

B. *Mensajes borrados*

El borrado de mensajes es otra circunstancia que tiene relevancia a la hora de realizar la exportación de una conversación. Existen en WhatsApp dos opciones de borrado, una únicamente para el emisor del mensaje y otra para todos los participantes en una conversación. Ambas situaciones se verán reflejadas de diferente forma en la exportación.

Por un lado, el borrado para todos los participantes se reflejará con el texto “Eliminaste este mensaje”. En la figura 4 se muestra un ejemplo de la visualización de un mensaje eliminado para todos los participantes de una conversación. Adicionalmente, en el siguiente ejemplo se muestra el resultado de su exportación:

```
[13/2/23, 9:12:05] Xabiel: Llegue
[13/2/23, 9:37:06] Xabiel: Eliminaste este
mensaje.
```

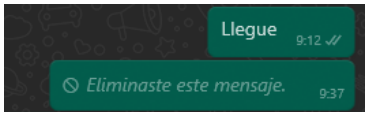


Fig. 4. Visualización de mensaje eliminado

Por otro lado, cuando el borrado se realiza únicamente para el emisor, su visualización desaparece completamente para este, mientras que permanecerá inalterado para el resto de participantes. Este efecto se mantendrá en la exportación. Es decir, en la exportación del emisor desaparecerá totalmente sin dejar rastro, mientras que para el resto de participantes aparecerá inalterado.

C. *Mensajes temporales*

La aplicación de WhatsApp dispone de una característica que le permite que en una conversación los mensajes tengan un tiempo de vida definido. Esta característica se activa para todos los mensajes de dicha conversación e incluye también a los elementos multimedia. El tiempo de vida puede ser de 24 horas, 7 días o 90 días. Una vez transcurrido el periodo definido, los mensajes dejan de visualizarse en las pantallas de los usuarios. Sin embargo, ese proceso no es tan automático cuando se realiza una exportación. Si al tener activada esta característica el usuario borra mensajes, se producen desincronizaciones que hacen que mensajes que deberían haber sido eliminados, y que de hecho no aparecen en la aplicación, sigan apareciendo en las exportaciones. A continuación se

muestran dos exportaciones realizadas simultáneamente (9:29 horas) por dos participantes en una misma conversación. Se puede observar cómo en el participante 1 aparecen tres mensajes, que han desaparecido totalmente en el participante 2.

```
[25/2/23, 8:53:40] Xabiel: Activaste la duración
de los mensajes temporales. Los mensajes nuevos
desaparecerán de este chat después de 24 horas
de haber sido enviados. Pulsa para cambiar.
[25/2/23, 8:55:27] Xabiel: Eliminaste este
mensaje
[25/2/23, 8:57:03] Xabiel: Mensaje temporal que
no voy a borrar
[25/2/23, 9:01:11] Xabiel: <adjunto: 0000009-
PHOTO-2023-02-25-09-01-12.jpg>
```

```
25/2/23, 8:53 - Xabiel García Pañeda: Xabiel
García Pañeda actualizó la duración de los
mensajes temporales. Los mensajes nuevos
desaparecerán de este chat después de 24 horas
de haber sido enviados. Pulsa para cambiar.
```

Estos mensajes siguen apareciendo en las exportaciones durante horas, hasta que la aplicación efectúe definitivamente su borrado. Por todo ello, pueden producirse incongruencias entre dos exportaciones realizadas desde dos participantes, incluso habiendo sido realizadas en el mismo momento.

V. Conclusiones

La extracción de conversaciones de WhatsApp para su uso en procesos judiciales no es un proceso complejo, ya que la propia herramienta dispone de una funcionalidad para hacerlo. Sin embargo, es importante tener en cuenta que se pueden dar una serie de situaciones que lleven a que dicha extracción no contenga ciertos mensajes y esto lleve a incongruencias con otras extracciones realizadas desde otro de los dispositivos implicados en la conversación. Estas situaciones deben ser reconocidas e interpretadas adecuadamente por los peritos, para evitar que se pueda concluir erróneamente que ha existido algún tipo de manipulación de las evidencias. Una de estas situaciones se da tras el borrado de un mensaje o multimedia para un usuario, que no deja rastro alguno en la exportación. Dos exportaciones desde dos participantes darían lugar a conversaciones con diferentes mensajes. El uso de mensajes temporales es otra característica que puede provocar incongruencias. Es cierto que la activación de esta funcionalidad queda reflejada en la exportación, pero a partir de ese momento la exportación variará según la hora de realización y la ejecución manual de borrados por parte de los usuarios. El caso de las votaciones es singular dentro del servicio WhatsApp. Su resultado puede ser modificado sin dejar rastro por lo que no se puede certificar que el resultado visible en la exportación sea el que se visualizó por los participantes en un momento dado.

Todas estas situaciones relevantes deben dar lugar a una serie de conclusiones. Todo lo que aparece en la exportación existe o ha existido y su contenido no está alterado. Es posible que además existieran otros mensajes. En algunos casos será posible ubicar estos mensajes en la conversación y en otros no. Exportaciones realizadas desde distintos participantes en una conversación pueden dar lugar a diferencias, pero los mensajes que están en ambas exportaciones deben ser necesariamente iguales. En todo caso, es importante alertar sobre la volatilidad del formato de exportación, ya que WhatsApp puede cambiar la información recogida en el mismo con frecuencia. Así que constantemente, y sin previo aviso, se pueden producir cambios que obliguen a la reinterpretación de los registros.

Referencias

[1] «Aplicaciones de mensajería: ranking según usuarios mensuales activos 2023», *Statista*. https://es.statista.com/estadisticas/599043/aplicaciones-de-mensajeria-mas-populares-a-nivel-mundial-de/ (accedido 3 de marzo de 2023).

[2] R. Z. Fathiyana, Yudiansyah -, N. Cahyadi, y D. J. Hidayat, «A Comparative Study and Analysis of Forensic Artifacts of WhatsApp and Telegram on Android Devices», *J. Inform. Commun. Technol. JICT*, vol. 4, n.º 2, Art. n.º 2, 2022, doi: 10.52661/j_ict.v4i2.143.

[3] K. Kaushik y Y. Katara, «Forensic Analysis of WhatsApp chat data», en *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, oct. 2022, pp. 1-6. doi: 10.1109/ICRITO56286.2022.9965028.

[4] A. S. M. Idris y N. H. A. Rahman, «A Comparative Analysis of Potential Digital Evidence in WhatsApp Web-Based and Mobile-Based Application», *Appl. Inf. Technol. Comput. Sci.*, vol. 3, n.º 2, Art. n.º 2, nov. 2022.

[5] S. D. Utami, C. Carudin, y A. A. Ridha, «ANALISIS LIVE FORENSIC PADA WHATSAPP WEB UNTUK PEMBUKTIAN KASUS PENIPUAN TRANSAKSI ELEKTRONIK», *Cyber Secur. Dan Forensik Digit.*, vol. 4, n.º 1, Art. n.º 1, jun. 2021, doi: 10.14421/csecurity.2021.4.1.2416.

[6] R. Khweiled, M. Jazzar, A. Eleyan, y T. Bejaoui, «Using SQLite Structure Analysis To Retrieve Unsent Messages On WhatsApp Messaging Application», en *2022 International Conference on Smart Applications, Communications and Networking (SmartNets)*, nov. 2022, pp. 01-06. doi: 10.1109/SmartNets55823.2022.9993988.

[7] L. A. A. Mohammad Shadeed Majdi Owda, «Forensic Analysis of " WhatsApp " Artifacts in Android without root», vol. 7, n.º 2, pp. 127-132, 2022, doi: 10.25046/aj070112.

[8] H. Fayyad-Kazan, S. Kassem-Moussa, H. J. Hejase, y A. J. Hejase, «Forensic Analysis of WhatsApp SQLite Databases on the Unrooted Android Phones», *HighTech Innov. J.*, vol. 3, n.º 2, Art. n.º 2, feb. 2022, doi: 10.28991/HIJ-2022-03-02-06.

[9] https://faq.whatsapp.com/1180414079177245/?locale=es_LA&cms_platform=android (Accedida 27/02/2023).

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Experiencias de consumo interactivo de video inmersivo VR360 en escenarios multi-cámara y multi-usuario distribuidos

M. Fernández-Dasi[1,2], M. Montagud[1,3], I. Fraile[1], J. Paradells[1,2], S. Fernández[1,2]
[1] Fundación i2CAT, Barcelona
[2] Universitat Politècnica de Catalunya (UPC), Barcelona
[3] Universitat de València (UV), Dpt. Informàtica, Burjassot (València)
{miguel.fernandez, mario.montagud, isaac.fraile, josep.paradells, sergi.fernandez}@i2cat.net

La producción y consumo de contenidos multimedia están en continuo crecimiento, y ello aplica particularmente a los formatos inmersivos, tales como el vídeo 360º o VR360. A pesar de haberse experimentado avances significativos en cuanto al procesado y distribución de vídeos VR360 en los últimos años, todavía quedan retos y oportunidades esenciales por abordar. Este artículo ofrece una visión general sobre cómo una plataforma extremo-a-extremo modular se ha ido extendiendo con tal de integrar contribuciones tecnológicas innovadoras en los ámbitos de la distribución adaptativa de baja latencia, procesado y distribución basados en campo de visión, sincronización multimedia inter-fuente e inter-cliente, y visionado social interactivo, todo ello con soporte para vídeos VR360 inmersivos de alta resolución. Asimismo, se presentan dos casos de uso para los que la plataforma se ha adaptado y desplegado, como ejemplo de su modularidad y alta aplicabilidad: retransmisión de eventos multi-cámara, y gestión colaborativa de emergencias, ambos para vídeos capturados en directo.

***Palabras Clave*- DASH, Sincronización Multimedia, Streaming Multimedia, Vídeo 360º, VR360, Visionado Social**

I. INTRODUCCIÓN

La producción, distribución y consumo de contenidos multimedia se están incrementando progresivamente en los últimos años. Asimismo, la resolución creciente de dichos contenidos, así como la necesidad de distribuirlos a través de entornos ubicuos y heterogéneos (ej. en cuanto a redes y dispositivos de consumo), no simplemente requiere de la necesidad de mayores capacidades de ancho de banda, almacenamiento y procesado, sino también de técnicas avanzadas y adaptativas para su codificación, distribución (ej. basadas en *streaming* adaptativo HTTP) y consumo. Estos requisitos, a su vez, son más estrictos cuando aplican a contenidos inmersivos o de Realidad Virtual (*Virtual Reality*, VR), tales como el vídeo 360º (en adelante, VR360), ya que incluyen mayor cantidad de información que los vídeos 2D y típicamente requieren resoluciones altas, así como ciertas capacidades de interacción, con tal de proveer niveles de calidad de experiencia (*Quality of Experience*, QoE) satisfactorios.

La comunidad científica y la industria han dedicado, y están dedicando, esfuerzos significativos con el objetivo de superar retos y limitaciones esenciales en el ámbito de los servicios de vídeo VR360 distribuidos (ej. [1, 2]). Este artículo ofrece una visión general de varias necesidades y objetivos siendo abordados en este campo, de manera coordinada, con tal de superar retos tecnológicos en cuanto al procesado, distribución y consumo interactivo de contenidos de vídeo VR360 de alta resolución (≥4K), incluyendo funcionalidades avanzadas tales como procesado y distribución basado en campo de visión, *streaming* de baja latencia, sincronización inter-fuente e inter-cliente, y soporte para escenarios de visionado social multi-usuario. Las contribuciones tecnológicas asociadas a cada uno de estos retos se integran en una plataforma extremo-a-extremo VR360 modular desarrollada previamente [3, 4], y se complementan con diferentes módulos para la medida y registro de métricas de calidad de servicio (*Quality of Service*, QoS), así como de consumo de recursos y de actividad de los usuarios. Una premisa esencial de las contribuciones tecnológicas siendo proporcionadas es que sean compatibles con los formatos y estándares existentes, así como que garanticen interoperabilidad con tecnologías y entornos web, así como con diferentes tipos de dispositivos de consumo.

Como prueba de evidencia de la modularidad y alta aplicabilidad de la plataforma, se presentan dos casos de uso relevantes para los que se ha adaptado y desplegado: retransmisión de eventos multi-cámara, y gestión colaborativa de emergencias.

II. RETOS Y OBJETIVOS

En esta sección se introduce la plataforma extremo-a-extremo de la que se parte en este trabajo y, a continuación, se

presentan cada uno de los módulos que se han integrado en la misma para abordar retos y objetivos específicos, incluyendo un breve repaso del estado del arte para cada uno de ellos.

A. Plataforma VR360 modular extremo-a-extremo

Un requisito esencial para abordar los objetivos y retos planteados consiste en disponer de una plataforma y marco de pruebas sobre el que integrar, evaluar y validar cada una de las contribuciones tecnológicas a diseñar. En este sentido, se parte de la plataforma presentada en [3, 4], compuesta de los siguientes bloques (Fig. 1):

Preparación de Contenidos. Por una parte, incluye módulos para la recuperación de vídeos almacenados y la ingesta de vídeos en vivo, ya sea a través de conexiones IP (ej. vía *Real-Time Streaming Protocol* (RTSP)) o de interfaces HDMI / USB. Por otro lado, soporta diferentes configuraciones de codificación, y conversión a/entre protocolos de *streaming*, tales como *Dynamic Adaptive Streaming over HTTP* (DASH), incluyendo su perfil de baja latencia *Common Media Application Format* (CMAF), que a su vez permite reaprovechar los segmentos multimedia generados para su distribución vía el protocolo HTTP Live Streaming (HLS), todo ello mediante uso de *ffmpeg*[1]. Además, soporta diferentes formatos de proyección para los vídeos VR360, tales como *Equirectangular Projection* (ERP) y *Cubemap Projection* (CMP).

Servidor de Contenidos y Señalización de Servicio. Servidor de contenidos para el almacenamiento de los contenidos generados y de la información de señalización sobre los mismos (ej. ficheros de manifiesto, catálogo de contenidos…), así como de recursos web para su acceso, catalogación y presentación. Este servidor puede configurarse como el punto de origen de una *Content Delivery Network (CDN)*, con tal de asegurar la provisión de contenidos a gran escala.

Consumo de Contenidos. Reproductor web VR360, desarrollado principalmente mediante los componentes *dash.js*[2], *three.js*[3] y *WebXR*[4], adoptado de [5].

B. Distribución adaptativa de baja latencia

Los retardos en los servicios de *streaming* multimedia son un factor crítico cuando se requiere o pretende una interacción o reacción en base a los contenidos consumidos, especialmente si son capturados y distribuidos en directo. El retardo en los servicios de *streaming* basados en HTTP ha sido tradicionalmente un obstáculo y reto a resolver [6], como por ejemplo cuando se pretende retransmitir eventos en directo, o de manera paralela y complementaria a servicios broadcast asociados [7]. Sin embargo, recientes avances en técnicas de codificación y *streaming* HTTP adaptativo, como el uso de CMAF (ej. [8]), permiten reducir drásticamente la magnitud de dichos retardos, incluso cuando se trata de vídeo VR360 de alta resolución (ej. [1, 2, 8]). Además, un reto adicional, y común, aplica cuando las cámaras a utilizar no son capaces de transmitir directamente vía DASH o HLS. En dichos casos, se requiere de un proceso de transcodificación y conversión de protocolos en un servidor intermedio. En la Tabla I se resumen resultados obtenidos en diferentes configuraciones de *streaming* DASH mediante la plataforma desarrollada y utilizando una cámara 4K VR360 (modelo *Kandao Qoocam 8K*). Se puede observar que cuando se tiene acceso a las tramas VR360 en crudo (*raw*) vía HDMI, y se configura el perfil CMAF con *chunks* de corta duración (ej. 100ms) y un buffer de reproducción corto (ej. 1s) en el cliente, se puede obtener retardos extremo-a-extremo (*glass-to-glass*) inferiores a 2s. Además, el uso de DASH *Low Latency*, con CMAF, permite configurar un umbral de retardo objetivo en los reproductores, de manera que estos adapten su tasa de reproducción y/o la calidad de vídeo (resolución, *bitrate*) a seleccionar con tal de ajustarse a este retardo objetivo.

C. Codificación y Distribución en base a Campo de Visión

Cuando se consumen contenidos VR360, en todo momento los usuarios sólo pueden ver una región del entorno 360º capturado, conocido como campo de visión (*Field of View*, FoV), que suele tener un tamaño de entre 90º-130º en los dispositivos de consumo actuales. Ello implica que la distribución de la esfera 360º a una calidad máxima resulta ineficiente en cuanto a consumo de ancho de banda y recursos de computación. Por este motivo, la comunidad científica ha propuesto varias estrategias de compresión y *streaming* que tratan de concentrar la resolución en los FoVs de una manera lo más transparente posible al usuario consumidor (ej. [1, 2]). Sin embargo, la mayoría de las propuestas existentes se basan en el uso de códecs y estrategias de *tiling* (estructuración del vídeo en una matriz con celdas a diferentes calidades) que pueden implicar restricciones de licencias y/o de soporte en entornos web. Así pues, los autores están explorando dos estrategias alternativas que garanticen la compatibilidad en entornos web (ej. uso del códec H.264), y minimicen los requisitos para los clientes finales mediante la entrega de un único flujo multimedia y de una simple instancia de decodificación.

La primera de ellas, motivada por los resultados prometedores obtenidos en [9], consiste en generar *N* versiones diferentes del mismo vídeo VR360, con diferentes ajustes de calidad en diferentes regiones del entorno 360º, como pueden ser las caras del cubo cuando se usa la proyección CMP (ver Fig. 2). De este modo, mediante una estrategia de señalización diseñada, el reproductor puede conmutar a la versión más adecuada de manera dinámica en base a los patrones de visionado. Resultados de pruebas preliminares han mostrado que se puede reducir el consumo de ancho de banda en el cliente en torno al 30% con respecto al uso de una estrategia de codificación uniforme. Con tal de confirmar los potenciales beneficios, se plantearán estudios de usuarios para diferentes configuraciones de codificación y vídeos.

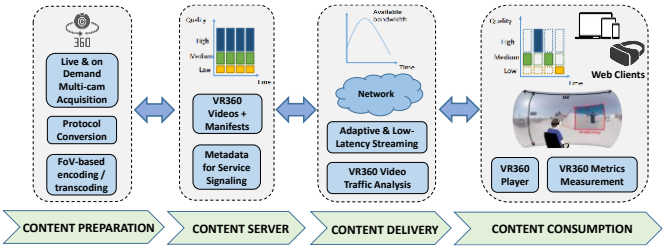


Fig. 1. Plataforma VR360 extremo-a-extremo [3, 4]

[1] Ffmpeg https://ffmpeg.org/ Último acceso en junio de 2023
[2] Dash.js https://github.com/Dash-Industry-Forum/dash.js Último acceso en junio de 2023
[3] Three.js https://threejs.org/ Último acceso en junio de 2023
[4] WebXR https://www.w3.org/TR/webxr/ Último acceso en junio de 2023

Tabla I
RETARDOS EN STREAMING DASH PARA VR360 4K EN DIFERENTES ESCENARIOS

Protocolo de Ingesta	Canal de Ingesta	Método DASH	# Representaciones DASH	Tamaño de Segmentos (s)	Tamaño de Chunks CMAF	Tamaño de Buffer de Reproducción (s)	Retardo (s)
RTSP	Ethernet	Traditional	1 (4K, 30fps)	1	-	1	~4
RTSP	WiFi	Traditional	2 (4K and 2K, 30fps)	1	-	1	~5
RTSP	Ethernet	Traditional	1 (4K, 30fps)	2	-	1	~5
Raw	HDMI	Traditional	1 (4K, 30fps)	1	-	1	~3.5
Raw	HDMI	CMAF	1 (4K, 30fps)	1	100ms	1	<2

La segunda estrategia consiste en activar un canal interactivo entre el reproductor y la instancia de codificación de vídeo, de modo que se concentre la resolución / calidad en el FoV instantáneo reportado por cada cliente. Esta estrategia presenta problemas de escalabilidad, aunque una potencial solución sería combinarla con la anterior, de modo que se genere un subconjunto delimitado de versiones del vídeo VR360 que se ajusten en gran medida a todas las combinaciones posibles de visionado.

Para cada una estas estrategias existen una serie de aspectos y factores relevantes a explorar con tal de maximizar el rendimiento y QoE, tales como: (i) la forma y tamaño (y margen de seguridad) de las regiones del vídeo VR360 en las que concentrar la calidad; (ii) los ajustes de degradación para la región exterior a esta ventana (ej. resolución, *bitrate*…); y (iii) los periodos de actualización de ventana y/o conmutación de versión de vídeo VR360.

D. Sincronización multi-cámara y multi-reproductor

Varios trabajos científicos han reflejado la relevancia de la sincronización multimedia en escenarios distribuidos e interactivos [10]. Por un lado, la sincronización inter-fuente (ej. multi-cámara) posibilita que se presenten de una manera alineada en el tiempo las escenas capturadas por diferentes cámaras en una sesión de *streaming* multimedia (ej. ver un gol al mismo tiempo desde diferentes cámaras desplegadas en un estadio de fútbol) [11]. Por otro lado, la sincronización inter-cliente (ej. multi-reproductor o multi-usuario) posibilita que se presenten de una manera alineada en el tiempo las escenas seleccionadas en todos los dispositivos cliente o usuarios en una sesión multimedia compartida [12]. Por su puesto, estas necesidades de sincronización aplican a cada una de las modalidades de contenidos de la sesión multimedia, como vídeo, audio, subtítulos, etc.

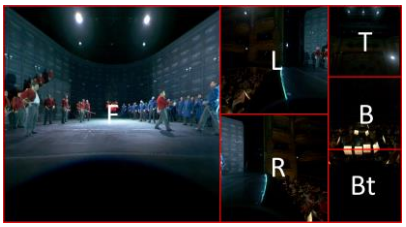


Fig. 2. Configuración no-uniforme en CMP, con diferentes resoluciones para diferentes caras del cubo (*F: Frontal; L: Left; R: Right; T: Top; B: Back; Bt: Bottom*)

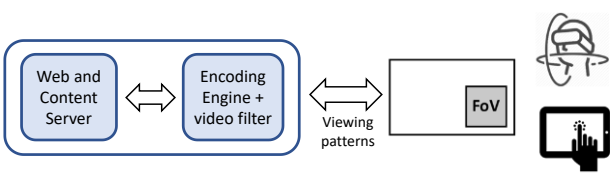


Fig. 3. Codificación dinámica en base al campo de visión reportado por clientes VR360

Como respuesta a estos requisitos, la plataforma ha sido extendida con tal de integrar una solución modular y escalable para proporcionar sincronización multi-cámara y multi-cliente, de manera precisa y adaptativa (ver Fig. 4). Por una parte, se ha diseñado y desarrollado un módulo de sincronización multi-cámara que permite la coordinación y sincronización de los procesos de codificación y *streaming* de varias cámaras distribuidas, independientemente de su localización, si están conectadas un equipo de procesado (ej. PC) con conectividad IP. Este módulo incluye un "Controlador de Cámaras y Sincronización" que se encarga de registrar todas las cámaras disponibles y deseadas en una sesión común, junto a sus capacidades (ej., tipo de vídeo, resoluciones y formatos que soportan…). En base a esta información, dicho controlador indica a cada equipo de procesado los ajustes de codificación y *streaming* DASH más adecuados para las cámaras que conecta, así como el momento exacto en el que lanzar dichos procesos, distribuyendo para ello una señal de temporización global a la que ajustarse. Durante el transcurso de la sesión, cada instancia de codificación informa al controlador sobre su estado y progreso, de manera que el controlador puede gestionar la actualización de un fichero de manifiesto que señalice los segmentos generados por cada instancia, englobando así los contenidos generados por las diferentes cámaras en una sesión única. Los segmentos generados por cada instancia de codificación y los ficheros de manifiesto asociados se pueden almacenar en servidores alojados en los PC distribuidos, o bien en un servidor central, y para ambos casos se pueden configurar dichos servidores como origen de una CDN si se requiere un servicio de *streaming* a gran escala. Esta solución de sincronización modular aporta escalabilidad y robustez frente a una solución centralizada en la que todos los flujos son codificados por una misma instancia de codificación, ya que: (i) la codificación de múltiples flujos de alta resolución requiere de *Graphics Processing Units* (GPUs) potentes y costosas; y (ii) problemas en la ingesta de una cámara podrían afectar a la continuidad del proceso de codificación global. Además, aunque se utilicen técnicas de paralelización y aceleración por hardware, el hecho de codificar varios flujos en una misma máquina resulta en un incremento del retardo de procesado y, por tanto, del servicio multimedia.

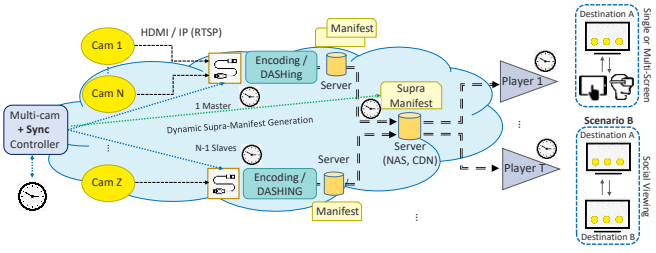


Fig. 4. Módulos de sincronización multi-cámara y multi-cliente

Por otra parte, se ha diseñado y desarrollado un módulo de sincronización inter-cliente o multi-usuario que permite registrar a los reproductores deseados en una sesión multimedia común y suministrarles tanto una señal de temporización como indicaciones de ajuste de reproducción con tal de conseguir una sincronización global en dicha sesión. Esta solución está basada en la filosofía del estándar *Digital Video Broadcasting Companion Screens & Streams* (DVB CSS), adoptado a su vez por el estándar *Hybrid broadcast broadband TV* (HbbTV) [10], pero adaptada para: (i) su compatibilidad en entornos web; (ii) contenidos en directo y bajo demanda; y (iii) entornos multi-pantalla y multi-usuario distribuidos. Esta solución también incluye los servicios correspondientes para la gestión de sesiones, y para la señalización y selección de contenidos en cada sesión.

Ambas soluciones de sincronización se pueden ofrecer como un *Software-as-a-Service* (SaaS), requiriendo simplemente de librerías y APIs basadas en Python o C (junto a la instalación de *ffmpeg*) para el control de las instancias de codificación y *streaming*, y basadas en Javascript para los reproductores web (o en el lenguaje correspondiente si se pretende integrar en reproductores nativos). De hecho, se han realizado pruebas de despliegue satisfactorias de estos módulos de control de sincronización inter-fuente e inter-cliente en instancias de *Amazon Web Services* (AWS).

Asimismo, ambas soluciones son capaces de proporcionar niveles de sincronización con precisión a nivel de trama (del orden de unos pocos ms), y se han validado funcionalmente con sesiones de hasta 5 cámaras simultáneas, de sesiones con hasta 10 dispositivos de consumo, con hasta 3 sesiones simultáneas e independientes, y con uso de diferentes dispositivos de consumo (ej. PCs, móviles, cascos RV…).

E. Visionado Social e Interactivo VR360

El consumo compartido de contenidos multimedia por parte de usuarios distribuidos puede fomentar la interacción social, *engagement*, e incluso posibilitar tareas colaborativas. En [12] se ofrece una revisión de las soluciones de visionado social para vídeo 2D propuestas hasta el momento, así como de los potenciales beneficios que estos escenarios pueden aportar. Asimismo, en [13] se analizan los retos y requisitos para poder proporcionar de manera efectiva estas experiencias multi-usuario cuando se pretenden consumir contenidos VR360. Así pues, en base a los resultados y recomendaciones en [12] y [13], se diseñado, desarrollado e integrado en la plataforma una herramienta de visionado social que incluye: (i) la gestión de sesiones compartidas; (ii) la selección de contenidos, incluyendo sesiones multi-cámara y en directo; (iii) canales de chat basado en texto y audio/video conferencia para interacción entre los usuarios, a través de *Web Real-Time Communication* (WebRTC); y (iv) funcionalidades asistivas, tales como zoom e inclusión de mini-ventanas mostrando las cámaras y regiones de visionado para cada uno de los usuarios en la sesión compartida (que a su vez pueden aplicarse al reproductor local con simplemente hacer *click* en las mismas).

III. CASOS DE USO

La plataforma VR360 desarrollada, incluyendo los nuevos módulos integrados en la misma, tiene una alta aplicabilidad en el sector de los servicios multimedia. A continuación, se presentan dos casos de uso en las que se ha adaptado y demostrado recientemente.

A. Eventos en vivo multi-cámara (ej. conciertos)

En el marco del proyecto RIS3 ViVIM, y en colaboración con TV3, la plataforma con el módulo de sincronización multi-cámara se ha desplegado para la retransmisión en directo de dos galas del concurso musical Eufòria en 2023. En concreto, se desplegaron 4 cámaras VR360 en diferentes emplazamientos (Fig. 5): junto al escenario; junto a la mesa del jurado; junto a los espectadores presenciales; y en la sala de realización. De este modo, la audiencia desde casa pudo seguir la gala de manera inmersiva, seleccionando la cámara de interés en cualquier momento, y desde el dispositivo disponible o preferido (ej. PC, móvil, tableta, gafas VR).

Cada cámara se conectó a un mini-PC o portátil con GPU integrada con prestaciones suficientes para poder codificar un flujo 4K VR360 en tiempo real. Los segmentos generados por cada instancia de codificación se almacenaron en un servidor compartido en red local, que se configuró como origen de la CDN de TV3. Configurando la generación de dos representaciones DASH con resoluciones de 4K y 2K para el flujo de cada cámara VR360, y buffers en los procesos de producción y consumo de contenidos de 2-3s de vídeo, el retardo total del servicio de *streaming* estuvo alineado con el retardo de la transmisión broadcast paralela.

Las retransmisiones estuvieron activas durante unas 3 horas para cada una de las galas, alcanzando una audiencia en torno a 20000 usuarios únicos durante dichos periodos. Ambas retransmisiones, así como grabaciones de actuaciones particulares de los concursantes, se pueden visualizar en la web de TV3[5]. Se utilizó *Google Analytics*, así como cuestionarios sobre la percepción de rendimiento, inmersividad y aplicabilidad de la tecnología.

Fig. 5. Cámaras VR360 instaladas junto al escenario (arriba) y junto a la mesa del jurado (centro). Interfaz del reproductor con cuatro botones para selección de cámara (abajo).

[5] Grabaciones de contenidos VR360 disponibles en la web de TV3: https://www.ccma.cat/tv3/euforia/360-immersiu/fitxa/146680/ Último acceso en junio de 2023

B. Gestión colaborativa de emergencias

En el marco del proyecto europeo H2020 Respond-A, y en colaboración con la Policía Local de Valencia, el Puerto de Valencia, la empresa alemana ATMOSPHERE y el cuerpo de bomberos de Chipre, la plataforma junto a los módulos de sincronización multi-cámara y visionado social se desplegaron en escenarios de simulacro de gestión colaborativa de emergencias, tales como incendios y accidentes químicos.

Mediante la instalación de las cámaras VR360 en drones, robots terrestres y/o emplazamientos estratégicos, la plataforma permitió a agentes de rescate (bien en el lugar de la emergencia o en remoto) inspeccionar de manera colaborativa la situación y entorno de la emergencia, dando soporte a la toma de decisiones. En estos escenarios, los flujos capturados por cada cámara se enviaron vía RTSP a una estación de procesado en el centro de *Command & Control* (C&C) terrestre desplegado en la zona de emergencia a través de una red 5G privada (detalles en [14]). La estación de procesado convirtió los vídeos recibidos a DASH para su distribución a los clientes, y también ejecutó los servicios de distribución y consumo interactivo de vídeo VR360, evitando así dependencias de conectividad a Internet (Fig. 6).

La plataforma se comportó de manera estable y satisfactoria en demonstraciones en Lárnaca (Chipre) y Valencia (España), obteniendo una valoración muy positiva por parte de los agentes de rescate que la vieron en acción.

Vídeos demostrativos de este caso de uso se pueden ver en: https://www.youtube.com/watch?v=C1TchguFFeQ y https://www.youtube.com/watch?v=hMOAj9Hh1uw

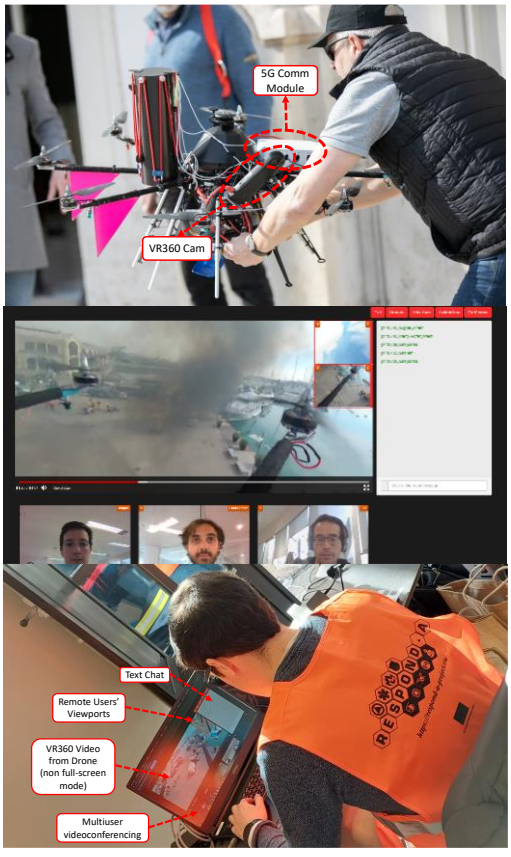


Fig. 6. Visionado VR360 social para gestión colaborativa de emergencias: (arriba) cámara VR360 montada en dron; (centro) aplicación web cliente; (abajo) agente de rescate usando la aplicación

Cabe mencionar que para ambos casos de uso se han realizado evaluaciones objetivas y subjetivas, obteniendo resultados satisfactorios y prometedores. Sin embargo, su presentación y análisis se omiten ya que el objetivo de este artículo es ofrecer una visión global de componentes tecnológicos en el ámbito del vídeo VR360 y de su potencial aplicabilidad en casos de uso relevantes.

IV. Conclusiones

En este artículo se ha ofrecido una visión global descriptiva de la extensión de una plataforma modular extremo-a-extremo VR360 para la integración de diferentes módulos innovadores con tal de proporcionar *streaming* adaptativo de baja latencia y con soporte multi-plataforma, sincronización multi-cámara y multi-reproductor, y funcionalidades de visionado social interactivo. Se han detallado las capacidades y rendimiento de dichas contribuciones tecnológicas, y se ha mostrado su aplicabilidad en dos casos de uso relevantes, como son la retransmisión de eventos multi-cámara en vivo (ej. conciertos), y la gestión colaborativa de emergencias, junto a evidencias de su despliegue efectivo y satisfactorio. Trabajos futuros reportarán sobre resultados de evaluaciones objetivas y subjetivas para las contribuciones tecnológicas presentadas, para/en ambos casos de uso.

Agradecimientos

Este trabajo ha sido financiado por la Agencia Estatal de Investigación (AEI), bajo un Proyecto Estratégico Orientado a la Transición Ecológica y a la Transición Digital (TED) 2021, con título REVOLUTION (Ref. TED2021-131690B-C32 y TED2021-131690A-C33), por ACCIÓ (RIS3CAT, Generalitat de Catalunya) en el marco del proyecto VIVIM (Ref. COMRDI18-1-0008), y por el programa H2020 de la Unión Europea en el marco del proyecto Respond-A (Ref. 883371). El trabajo de Miguel Fernández ha sido financiado por la *Secretaria d'Universitats i Recerca de la Generalitat de Catalunya* y por el Fondo Social Europeo (Personal Novel 2022FI-B1-00228). El trabajo de Mario Montagud Climent ha sido financiado por MCIN/AEI/10.13039/501100011033 (Ayuda Ramón y Cajal 2020, Ref. RYC2020-030679-I) y por "FSE Invierte en tu futuro".

Referencias

[1] C.L. Fan, W.C. Lo, Y.T. Pai, C.H. Hsu, "Survey on 360° Video Streaming: Acquisition, Tránsmission, and Display", ACM Comput. Surv. 52, 4, Article 71, Sept. 2019

[2] M. Xu, C. Li, S. Zhang, P.L. Cállet, "State of the Art in 360° Video/Image Processing: Perception, Assessment and Compression", IEEE JSAC, vol. 14, no. 1, pp. 5 26, Jan. 2020

[3] M. Fernández-Dasí, M. A. Torres, M. Montagud, M. García-Pineda, "Plataforma modular para la codificación y distribución interactiva de contenidos VR360 basada en campo de visión", JITEL 2021, A Coruña, Oct. 2021

[4] M. Fernández-Dasí, M. Montagud, J. Paradells, "Design, development and evaluation of adaptative and interactive solutions for high quality viewport aware VR360 video processing and delivery", ACM MMSys'22, Athlone (Ireland), June 2022

[5] M. Montagud, I. Fraile, E. Meyerson, S. Fernández, "Open-Source Accessibility-Enabled VR360 Player", ACM IMX 2020, Barcelona (Spain), June 2020

[6] F. Boronat, R. Mekuria, M. Montagud, P. Cesar, "Distributed Media Synchronization for Shared Video Watching: Issues, Challenges, and

Examples", Social Media Retrieval, Springer Computer Communications and Networks, ISBN 978-1-4471-4554-7, 2013

[7] F. Boronat, D. Marfil, M. Montagud, J. Pastor, "HbbTV-Compliant Platform for Hybrid Media Delivery and Synchronization on Single and Multi-Device Scenarios", IEEE Transactions on Broadcasting, 64(3), pp. 721-746, September 2018

[8] M. Uitto, A. Heikkinen, "Exploiting and Evaluating Live 360° Low Latency Video Streaming Using CMAF", 2020 European Conference on Networks and Communications (EuCNC), Dubrovnik (Croatia), 2020

[9] D. Gómez, J. A. Núñez, I. Fraile, M. Montagud, S. Fernández, "TiCMP: A lightweight and efficient Tiled Cubemap projection strategy for Immersive Videos in Web-based players", ACM NÓSSDAV'18, Amsterdam, June 2018

[10] M. Montagud, P. Cesar, J. Jansen, F. Boronat, "MediaSync: Handbook on Multimedia Synchronization", Springer-Verlag, ISBN 978-3-319-65840-7, 2018

[11] F. Boronat, D. Marfil, M. Montagud, J. Pastor, "Hybrid Broadcast/Broadband TV Services and Media Synchronization. Demands, Preferences and Expectations of Spanish Consumers", IEEE Transactions on Broadcasting, 64(3), pp. 52-69, August 2017

[12] F. Boronat, M. Montagud, P. Salvador, J. Pastor, "Wersync: a web platform for synchronized social viewing enabling interaction and collaboration", JNCA, 2020

[13] S. Rothe, A. Schmidt, M. Montagud, D. Buschek, H. Hußmann, "Social viewing in cinematic virtual reality: a design space for social movie applications", Virtual Reality, 25(3), 613-630, 2021

[14] Z. Paladin, Ž. Lukšić, N. Kapidani, M. Montagud, M. Fernández-Dasí, S. Srinidhi, T. Wöllert, G. Boustras, "The 5G-supported Unmanned Aerial Vehicles for Emergency Cases Response", 46th International ICT Convention, MIPRO 2023, Opatija (Croatia), May 2023

.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Distribuidor de carga para SMS Gateway

Óscar Palacín Grasa, Fernando Orús Morlans, Julián Fernández Navajas, María Canales Compés.
System One Noc & Development Solutions, S.A. / Departamento Ing. Electrónica y Comunicaciones
Universidad de Zaragoza
Edificio Ada Byron, C/ María de Luna,1. 50018, Zaragoza.
797021@unizar.es, forus@sonoc.io, navajas@unizar.es, mcanales@unizar.es.

Un SMS Gateway es un sistema que permite enviar y recibir mensajes de texto (SMS) entre dispositivos móviles y aplicaciones a través de una red telemática. IRISGW es el SMS Gateway de partida de nuestro trabajo, que proporciona una interfaz HTTP para recibir mensajes SMS de clientes y redirigirlos a los proveedores capaces de enviarlos a los terminales de los destinatarios finales. Se propone anteponer a IRISGW un balanceador de carga (IRISLB). Este módulo permitirá desacoplar la lógica interna IRIS relacionada con la facturación, enrutamiento, etc… de la gestión de las sesiones con los clientes y la recepción y envío de mensajes. Para solucionar las problemáticas en cuanto a fiabilidad (si se presenta un único punto de fallo) y escalabilidad (cuando se alcanza un número determinado de mensajes por segundo y la plataforma llega a su límite de capacidad), se propone desplegar IRISLB dentro de un grupo elástico de instancias en *cluster* Kubernetes.

***Palabras Clave*- SMS Gateway, Kubernetes.**

I. Introducción

La tecnología de SMS (Short Message Service) es un servicio disponible para los teléfonos móviles que permite el envío de mensajes cortos (con un límite de caracteres) a través de una amplia variedad de redes, incluidas las redes 4G/5G. Aprovechando el éxito de SMS, como vemos en [1], "GSMA (Global System for Mobile Communications Association) ha intentado fomentar múltiples tecnologías para el envío de mensajes MMS (Multimedia Message Service) que no han conseguido imponerse en el mercado". Sin embargo, SMS sigue teniendo demanda gracias a su ubicuidad e interoperabilidad (SMS es una tecnología universal que se puede utilizar en cualquier red móvil y entre diferentes operadores. Los mensajes SMS no requieren una conexión a Internet y se pueden enviar y recibir de manera rápida).

SMS A2P (Application-to-Person) se refiere al envío de mensajes SMS desde una aplicación o plataforma automatizada hacia un usuario o receptor final [2]. En este caso, "aplicación" puede referirse a una amplia gama de servicios o sistemas automatizados, como notificaciones de servicios, alertas, recordatorios, confirmaciones de transacciones o códigos de verificación, entre otros. El servicio A2P ha experimentado una alta demanda debido a las nuevas regulaciones de seguridad en pagos electrónicos y aplicaciones bancarias, así como al aumento en el uso de notificaciones vía SMS en diversas aplicaciones. A diferencia de los mensajes P2P (Person-to-Person), que son enviados y recibidos directamente entre usuarios individuales, los mensajes SMS A2P son generados por sistemas o aplicaciones con el propósito de ser entregados a múltiples usuarios o destinatarios.

El uso de SMS A2P [3,4] es común en diversos sectores, como servicios financieros, comercio electrónico, atención al cliente, servicios de entrega y logística, salud o entretenimiento. Estos mensajes A2P son enviados a través de pasarelas de mensajería, que son plataformas de software que permiten la conexión entre la aplicación o sistema automatizado y los operadores de telefonía móvil, facilitando la entrega de los mensajes de texto. Las empresas y proveedores de servicios suelen utilizar servicios de mensajería A2P para gestionar y enviar estos mensajes de manera eficiente y masiva.

En este contexto, surgen los Wholesale SMS, que son un modelo de negocio en el que una empresa compra grandes cantidades de mensajes SMS a un proveedor y luego revende esos mensajes a sus propios clientes a precios más bajos que el que ofrecen directamente las operadoras. Este modelo de negocio se utiliza comúnmente por empresas que ofrecen servicios de SMS a sus clientes, como proveedores de servicios móviles, proveedores de software de mensajería, proveedores de servicios de SMS y otros intermediarios. La compra de mensajes SMS a granel permite a las empresas obtener precios más bajos por mensaje, lo que les permite ofrecer precios competitivos a sus clientes y aumentar su margen de beneficio. Además, las empresas pueden personalizar los mensajes SMS que envían a sus clientes y gestionar sus campañas de marketing de manera más eficiente. El modelo de negocio de Wholesale SMS también es beneficioso para los proveedores de SMS, ya que les permite vender grandes cantidades de mensajes SMS a un solo cliente, reduciendo sus costos de administración y aumentando su volumen de ventas.

SONOC, la empresa con la que se colabora en el desarrollo del presente trabajo, es líder en el mercado de software para operadores mayoristas de telefonía (wholesale carriers). En este contexto, SONOC ha desarrollado una plataforma de SMS Gateway [5] con nombre IRIS Gateway como solución para gestionar mensajería A2P al por mayor sobre protocolo SMPP (Short Message Peer-to-Peer) [6]. Las diferentes funcionalidades del sistema IRIS se encuentran desplegadas en varias máquinas que pueden ser reales o virtuales. Dado el gran volumen de tráfico que debe gestionar la plataforma es necesario estar alerta, de forma continua, ante cualquier variación en la situación de demanda de recursos. En la actualidad, el uso de máquinas virtuales y contenedores virtuales [7,8] permite agilizar su funcionamiento, para lo que se utilizan orquestadores como es el caso de Kubernetes [9].

En vista de lo explicado anteriormente, los tres objetivos que se persiguen en el presente trabajo son: procesar los mensajes SMS para optimizar su envío, gestionar adecuadamente las conexiones SMPP de los clientes y adecuar la capacidad de procesamiento, con los recursos disponibles, a la evolución de la demanda.

Partiendo de la necesidad de dar solución a los objetivos propuestos, el presente artículo se organiza presentando en primer lugar el sistema desarrollado previamente por la empresa SONOC y que servirá de punto de partida. Sobre esta solución, se proponen una serie de modificaciones funcionales que permitan alcanzar los objetivos descritos anteriormente. A continuación se va a exponer el despliegue del sistema, utilizando la tecnología de Kubernetes. Para evaluar el resultado del desarrollo realizado se van a realizar diferentes pruebas de funcionamiento, comentando cómo se adapta el sistema a posibles situaciones de carga. Por último, se presentan las conclusiones y líneas futuras de trabajo.

II. Punto de partida del trabajo

A. *Protocolo SMPP.*

SMPP (Short Message Peer-to-Peer) es un protocolo de comunicación utilizado para enviar y recibir mensajes SMS (Short Message Service) entre dispositivos móviles y servidores de SMS.

El funcionamiento del protocolo SMPP se basa en un modelo cliente-servidor, donde el cliente SMPP se conecta al servidor SMPP a través de una conexión segura y envía mensajes SMS al servidor SMPP. El servidor SMPP procesa los mensajes y los envía a través de la red móvil al destinatario adecuado. La comunicación entre el cliente y el servidor se realiza a través de un canal dedicado y seguro, lo que garantiza la entrega confiable y segura de los mensajes SMS.

El protocolo SMPP permite la entrega de mensajes SMS en tiempo real, lo que significa que los mensajes se entregan al dispositivo móvil del destinatario casi al instante. También permite la entrega de mensajes en masa, lo que permite enviar mensajes a muchos dispositivos móviles de forma simultánea. La confirmación de la llegada del mensaje SMS al destino se realiza mediante el mensaje DLR. SMPP es un protocolo muy eficiente y escalable que se utiliza en todo el mundo para el envío y recepción de mensajes SMS. Es compatible con diferentes lenguajes de programación y se puede integrar con sistemas existentes, lo que lo hace muy flexible y fácil de usar.

Para el envío de SMS también puede utilizarse el protocolo HTTP de propósito general, y es muy común que convivan ambos protocolos.

B. *Sistema previo desarrollado.*

En la Figura 1 se observa el funcionamiento de envío de mensajes a través de la plataforma IRIS.

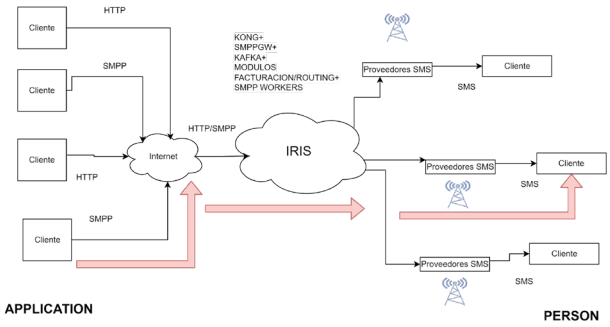


Fig. 1. IRIS. Esquema de uso

Como hemos mencionado anteriormente cuando se envían mensajes P2P (Person-to-Person), los mensajes se envían habitualmente a través de las redes de telecomunicaciones móviles. Sin embargo en nuestro caso en el envío de mensajes A2P (Application-to-Person) tenemos dos mundos: el primero es el mundo IP con la aplicación enviando los mensajes a través de internet y la posterior recepción por los proveedores donde comienza el mundo no IP con la entrega de los mensajes a través de su red de comunicaciones móviles a los terminales móviles personales.

Para tener una mejor visión del funcionamiento de IRIS, vamos a presentar en la figura 2 el despliegue de sus funciones, donde se incluyen en el recuadro punteado en rojo las modificaciones propuestas en el presente artículo. En la parte superior del esquema se observa que cada cliente de la plataforma IRIS tiene disponible un servicio web llamado hypersports donde puede dar de alta los usuarios a los que envía los mensajes y los proveedores que utiliza, además de configurar las rutas para los envíos, facturación y parámetros básicos, como número de binds disponibles para abrir. Esta configuración quedará guardada en una base de datos que será consultada posteriormente por módulos de enrutamiento de IRIS.

Actualmente los clientes sólo pueden enviar mensajes a través del protocolo HTTP hacia Kong [10] para su balanceo. Después del proceso de autenticación, los mensajes se envían hacia la API del IRISGW. Si la tasa de mensajes entrantes es mayor que la que puede procesar el sistema, la solución se resuelve utilizando unas colas distribuidas de mensajes que son consumidas a la tasa que acepta el proveedor. Como solución para la cola de mensajes distribuida se ha elegido Apache Kafka.

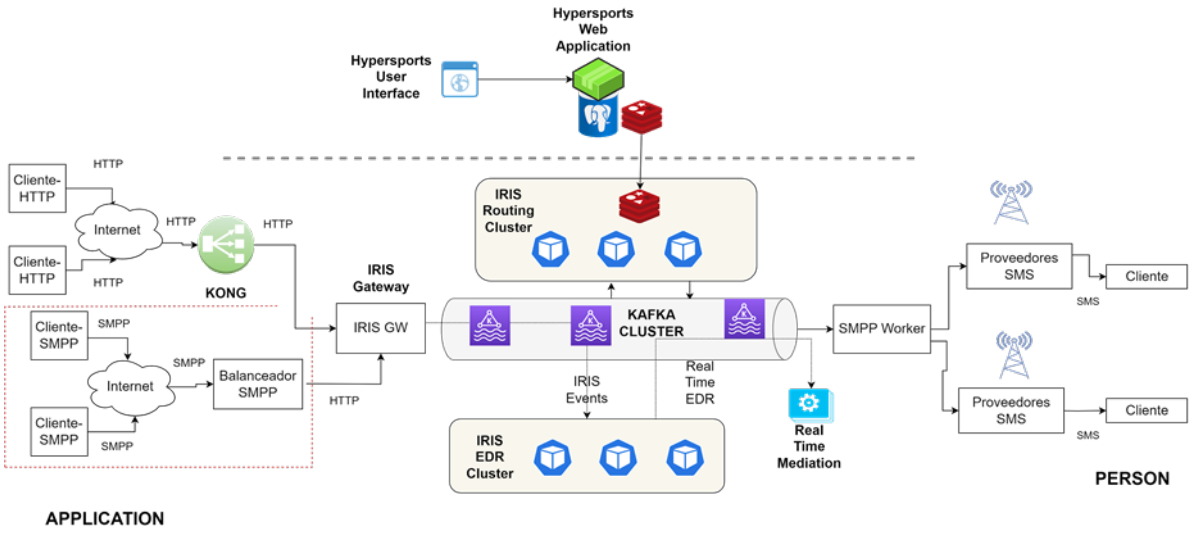


Fig. 2. Escenario inicial IRIS y propuesta

III. Desarrollo del Balanceador SMPP

A. *Respuesta a los objetivos.*

De acuerdo a los objetivos planteados, en primer lugar, para optimizar el envío y manejo de un gran flujo de mensajes SMS se propone el desarrollado de un módulo balanceador de carga para el protocolo de SMPP. Su función será enviar y repartir los mensajes SMS a un servicio encargado del envío de mensajes.

Para la gestión de sesiones debemos autentificar a los usuarios, gestionar que las conexiones cumplan los límites establecidos con SONOC y que cumplan la tasa de mensajes permitida. Gracias a esta gestión se conseguirá desacoplar la lógica interna IRIS relacionada con la facturación, enrutamiento, etc. de la gestión de las sesiones y la recepción de mensajes a través del protocolo SMPP.

Para obtener la máxima capacidad de procesamiento con los recursos disponibles se ha decidido utilizar virtualización mediante contenedores Docker. Estos utilizan sólo los recursos necesarios para ejecutar la aplicación, sin desperdiciar recursos en la duplicación de sistemas operativos completos como ocurre con las máquinas virtuales. Esto permite maximizar su uso y reducir los costos. Debido a que nuestra plataforma tiene clientes alrededor del mundo, será necesario desplegar múltiples instancias de ésta, y para ello, los contenedores nos ofrecen mayor facilidad a la hora de moverse y ser ejecutados en diferentes ubicaciones. Otra ventaja clave es la posibilidad de utilizar orquestadores de contenedores que nos permitan escalar la aplicación en función de las necesidades, lo que significa que se pueden agregar o eliminar contenedores en función de la carga de trabajo. Esto permite que la aplicación propuesta sea eficiente con los recursos utilizados, algo muy importante para un sistema que tendrá múltiples picos y valles de trabajo en cada parte del mundo, dependiendo de la hora local.

B. *Funcionalidades a desarrollar.*

Para que nuestros clientes puedan enviar mensajes SMS a través del protocolo SMPP de una forma eficiente, debemos desarrollar un módulo de balanceador de carga que realice una función similar a Kong pero para el protocolo SMPP. En caso de que los clientes quieran utilizar el protocolo SMPP para el envío de mensajes deberán iniciar una conexión con el Balanceador SMPP y enviar mensajes. Cuando se reciben los mensajes se realiza la autentificación (para el SMPP, el mensaje debe provenir de un cliente habilitado en la plataforma. Por ejemplo, si un BIND está activo, pero el cliente se ha bloqueado por un problema de crédito, el mensaje se debe rechazar) y un control de la conexión y de la tasa de mensajes permitida. Posteriormente si se completa correctamente las verificaciones anteriores se envían los mensajes a través de una petición HTTP que invoca la API de un único IRISGW.

A la hora de realizar el desarrollo se manejaron varias posibilidades. La primera de ellas consistía en que el módulo permitía establecer conexiones SMPP con clientes y cuando llegara un mensaje de un cliente se iniciara otra sesión SMPP con IRISGW, de forma que se mantuvieran las dos sesiones abiertas de forma simultánea y se reenviara los mensajes del cliente al IRISGW. Esta opción se ha descartado porque obliga a que IRISGW maneje dos protocolos el SMPP y HTTP y tenga que diferenciar entre los dos tipos de clientes y también por la dificultad de escalar la solución.

Como puede verse en la figura 2, La alternativa que se ha desarrollado es que el balanceador de carga SMPP establezca conexiones SMPP con los clientes para recibir los mensajes pero que sean enviados al IRISGW utilizando la API disponible con el protocolo HTTP. De forma inversa se mandará la confirmación DLR de vuelta al cliente. De esta forma se ha desplegado la funcionalidad básica de un servidor de SMPP: establecer conexiones SMPP, recibir

mensajes y mandar un DLR automáticamente a estos mensajes.

A esta funcionalidad básica ha sido necesario añadir otras tres funcionalidades: autentificación de usuarios y comprobación de que las peticiones se ajustan a lo contratado por el cliente, adaptación del protocolo SMPP al protocolo HTTP de forma transparente para el cliente a través de la creación de peticiones HTTP a partir de los mensajes recibidos y recepción de confirmación de llegada del mensaje al destino final transmitido desde IRISGW a través de HTTP y envío del correspondiente DLR a nuestro cliente a través de SMPP.

C. *Creación de imágenes Docker.*

El siguiente paso es crear imágenes de contenedores Docker, que contengan toda la funcionalidad expuesta anteriormente.

Inicialmente se pensó en tener todas las funcionalidades comentadas anteriormente en una única imagen. Sin embargo, se vio que sería conveniente desacoplar la gestión y control de las conexiones SMPP del envío del mensaje SMS. De forma que se puede dimensionar cada uno de los servicios de forma separada y dinámica dependiendo del número de mensaje a enviar y mejorando las prestaciones del sistema.

Para esta solución el módulo Balanceador SMPP implementado se divide en dos módulos (figura 3). El primero llamado proxy se encargaba de la gestión de las conexiones SMPP, autenticación de clientes y control de número de binds, de que se respete la tasa de mensajes por segundo y de la recepción de los mensajes a través de la conexión SMPP. El otro módulo llamado envio_mensaje está dedicado a la gestión del envío de mensajes y posterior recibimiento del DLR del mensaje por parte del IRISGW. Vemos como en nuestra solución esperamos que el módulo de envío de mensaje reciba un DLR a través de HTTP en una dirección y puerto establecidos. Este servicio se encarga de enviar el DLR al módulo que le envió el mensaje para que luego sea éste el que se lo envíe a través de SMPP a nuestro cliente. Esto es realmente importante porque sólo podemos cobrar a nuestros clientes por los mensajes que son enviados de forma correcta.

Por lo tanto, la mejora en el rendimiento se consigue colocando múltiples instancias del servicio de envío de mensaje por cada servicio de proxy. Se observó que como se conoce de antemano el número de clientes de la plataforma IRIS y que abrir y cerrar binds es poco costoso, se podía dimensionar el sistema y desplegar una imagen de proxy por cada cliente, de forma que no fuera necesario lanzarlos de forma dinámica, como el módulo de envío de mensajes.

Para la autentificación de los clientes se realiza una consulta a una base de datos Redis a través del cliente de java jedis de donde se extraen los usuarios registrados, sus contraseñas y el número de binds que tiene disponibles, posteriormente se compara si alguno de ellos coincide con el cliente y contraseña de la sesión que se inicia. Si no coincide, no se permite establecer las conexiones SMPP y se informa al cliente, en caso contrario se autentica el cliente y se comprueba que tiene disponibles binds para establecer una sesión con él. Se ha utilizado Redis ya que es necesario poder realizar modificaciones de forma rápida y dinámica.

También podemos ver cómo el módulo proxy es el encargado de establecer la conexión SMPP con el cliente, Después de la autentificación y verificación de la conexión, el cliente puede empezar a mandar mensajes. Cuando el proxy recibe un mensaje, lo manda a través de una conexión con sockets que tiene con el servicio de envio_mensaje, este servicio es el encargado de repartir entre las réplicas del módulo para que envíen los mensajes al IRISGW.

Para la comunicación entre los módulos se ha utilizado la biblioteca sockets de java que permite crear sockets y mandar mensajes con la estructura de protocolo SMPP a través de ellos pero sin la necesidad de establecer conexiones SMPP. Los pod de proxy actúan como clientes que se conectan y envían los mensajes de los clientes a los servidores sockets que serán los pod de envio_mensaje.

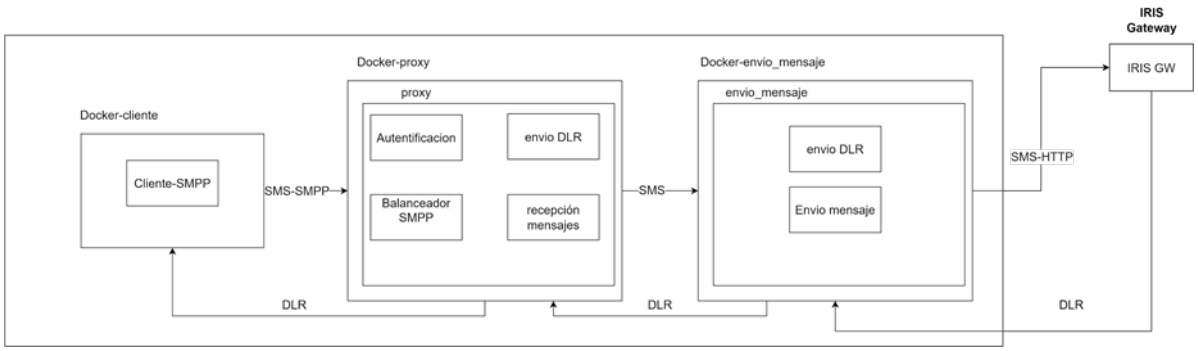


Fig. 3 imágenes de contenedores

IV. Despliegue con Kubernetes

A. *Despliegue básico.*

Una vez creadas las imágenes de los contenedores Docker que vamos a utilizar, podemos pasar a su implementación en un entorno de Kubernetes. Como es sabido, el entorno en el que se mueve Kubernetes se llama *cluster*, que es un conjunto de equipos reales con Kubernetes donde se alojan los recursos de una misma aplicación, de forma que ésta no se encuentra localizada en un único equipo, sino que está distribuida a través de múltiples nodos que forman parte de un *cluster*. En nuestro caso vamos a utilizar un *cluster* para simular un entorno de producción.

Dadas las distintas necesidades que tendrán los dos servicios (proxy y envio_mensajes) debido a su naturaleza, tendremos que incluir componentes de Kubernetes que actúen sobre los *deployment* de las dos aplicaciones que tenemos. Para el primer servicio de proxy, puesto que el número de clientes que abren conexiones es fijo, dimensionamos el sistema estableciendo un número fijo de réplicas. Hay que tener en cuenta que hay una etapa previa de acceso al sistema (iptable), independiente de nuestra aplicación, de tal forma que si la conexión IP no pertenece a ningún cliente previamente autorizado, nunca accederá a nuestro sistema Sin embargo, para la segunda etapa (envio_mensajes) necesitamos poder autoescalar el número de réplicas para el envío de mensajes conforme los clientes necesiten enviarlos. También debemos considerar que en el caso del primer servicio se deben mantener las conexiones con los clientes y no podemos eliminar contenedores ya que se podrían perder conexiones abiertas con los clientes.

B. *Autoescalado.*

Para la orquestación de los contenedores que ejecutan la aplicacion de enviar mensajes es necesario un componente de Kubernetes HPA (Horizontal Pod Autoscaler), que es un elemento que toma medidas de uso de los recursos hardware asignados a los *pod* de un deployment y los pasa al plano de control de Kubernetes para que éste decida si se deben tomar acciones de escalado de la aplicación.

Inicialmente se utilizó el uso de memoria como métrica predeterminada para realizar el autoescalado, ya que no requería de ninguna configuración adicional. Tras realizar algunas pruebas se observó que utilizando una métrica personalizada basada en el número de mensajes recibidos por segundo se lograba un autoescalado más preciso. Esto se traduce en una mejora en el rendimiento y en el consumo de recursos.

Nuestra métrica personalizada se debe exponer a través de Custom Metrics API para ser utilizada por el HPA. Para utilizar métricas personalizadas, es necesario que el *cluster* Kubernetes tenga una solución de monitoreo, como Prometheus, configurada y en funcionamiento.

C. *Networking.*

Los servicios de Kubernetes son una abstracción que permite el descubrimiento, equilibrio de carga y comunicación, mediante el protocolo IP, entre los *pod* de una aplicación.

Como puede verse en la figura 4, para la primera aplicación Proxy se ha utilizado un servicio NodePort, que expone la aplicación en un puerto en cada nodo del cluster y asigna un puerto externo para acceder al servicio a través de una IP pública única. En nuestra configuración se utiliza el puerto 8057 para acceder al servicio y se especifica que la conexión utiliza el protocolo TCP. Con ello ya disponemos de la IP pública y el puerto para acceder al servicio a través de la red.

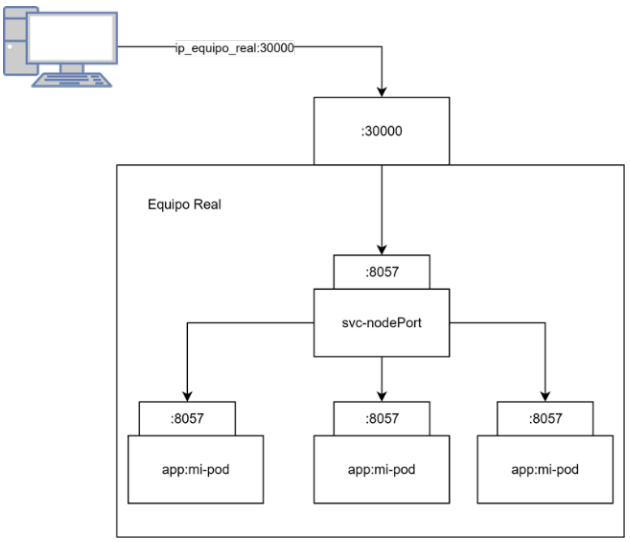


Fig. 4. Esquema de funcionamiento del nodePort.

Respecto a la comunicación entre los dos deployments, cuando un *pod* que ejecuta Proxy quiere comunicarse, mediante IP, con los *pod* de envio_mensaje, puesto que están en el mismo *cluster* de Kubernetes, utilizará el nombre del Service como un nombre de host DNS. El nombre del *Service* se resuelve automáticamente con la dirección IP del *Service* mediante el DNS interno de Kubernetes. Como hemos mencionado anteriormente este *Service* será el encargado de hacer llegar el mensaje, de forma transparente, a la instancia seleccionada.

Para permitir mandar las peticiones HTTP desde los *pod* del *cluster* hacia el IRIS GW debemos configurar el *networking* de Kubernetes, para ello se debe crear una NetworkPolicy que permita el tráfico saliente hacia cualquier dirección IP y puerto TCP fijo: 8088. En general, cuando los *pod* realizan peticiones salientes, utilizan la dirección IP del nodo en el que están programados como la dirección de origen en las conexiones salientes. Esto se debe a que los *pod* se ejecutan en el contexto del nodo y se comunican a través de la interfaz de red del nodo.

En la figura 5, vemos que el *Service* que distribuye la carga entre *pod* se encuentra localizable por todos los nodos, de forma que cualquier petición, interna o externa al

cluster pueda ser recibida por el *Service*. En el caso de querer servir peticiones, provenientes del exterior, se hace uso de un puerto seleccionado de las interfaces externas de los nodos del *cluster* para recibir paquetes. Si se quieren servir peticiones provenientes del interior del *cluster*, Kubernetes implementa un servidor de nombres conocido por todos los *pod* y con una dirección IP estática, y de esta forma se puede localizar el *Service* sencillamente sin conocer su IP mediante peticiones DNS. Cuando un servicio NodePort tiene múltiples réplicas, Kubernetes utiliza un balanceador de carga interno para distribuir el tráfico entrante entre las diferentes réplicas del servicio. De esta forma, las conexiones entrantes a través de un *socket* TCP se distribuirán automáticamente entre las réplicas del servicio.

En resumen, las comunicaciones se agrupan en el Service, que se encarga de distribuirlos a los Pod, que a su vez se comunican con otro Service único que los distribuirá, de nuevo, a sus Pod.

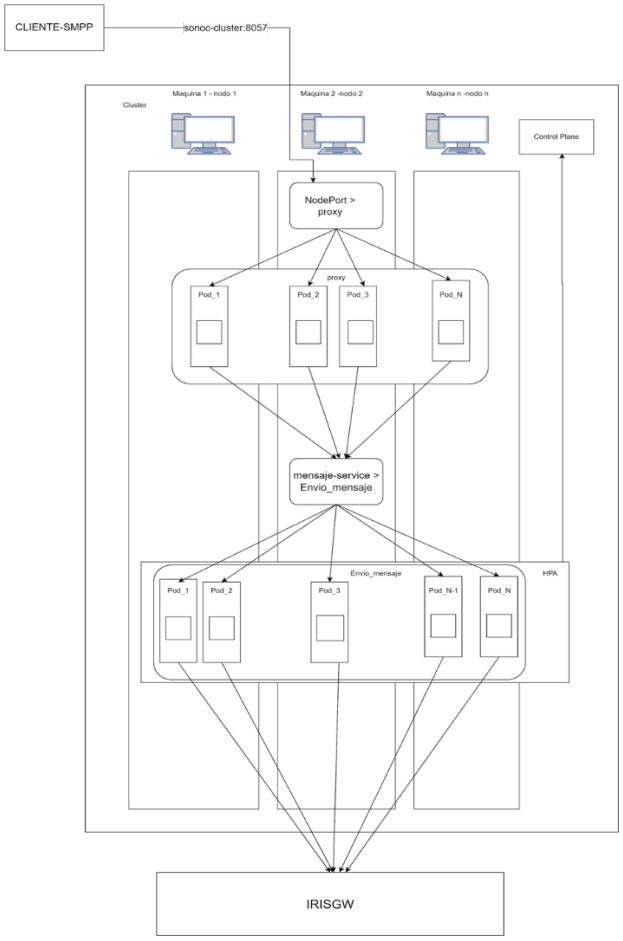


Fig. 5. Despliegue de la solución.

V. PRUEBAS REALIZADAS

Se han realizado pruebas de carga del sistema donde se han enviado simultáneamente 10, 100 y 1000 mensajes SMS por parte de un cliente. En estas pruebas hemos evaluado el tiempo total necesario para el envío de los mensajes que nos indica la capacidad de dar servicio a los clientes y el tiempo de *delay* de cada mensaje individualmente que nos permite analizar si hay mensajes que han tenido un tiempo de envío elevado. El tiempo de *delay* es acumulativo, por tanto en cada prueba realizada el último mensaje en ser procesado será el que tenga un mayor *delay*, puesto que tarda más tiempo en ser enviado.

Para esta prueba se ha desarrollado un programa que simula el envío masivo de mensajes SMS de forma simultánea. Cada prueba de envío de 10, 100 y 1000 se realizó 50 veces y se realizó una media de los resultados obtenidos. Se muestra en la tabla I la comparativa de resultados.

Tabla I
TABLA DE RESULTADOS DE TIEMPOS

Mensajes enviados	Tiempo total(ms)	Delay(ms)	Número de réplicas
10	2020	1135	10
100	8070	7223	10
100	4063	2154	20
1000	6065	5266	20

Durante las pruebas de envío simultáneo de 100 mensajes SMS, el sistema ha necesitado las 10 réplicas disponibles del servicio envio_mensaje. El tiempo necesario para que se detecte el aumento del tráfico y se levanten las réplicas necesarias provoca una pequeña acumulación en los mensajes pero se resuelve adecuadamente. Posteriormente, ya que se utilizaron el total de las réplicas, se decidió aumentar el número de réplicas disponibles de 10 a 20 para una nueva prueba. Se analizó si un aumento del número de réplicas se traduce en una reducción del tiempo necesario para el envío de los mensajes o reducción en el *delay* de los mensajes de forma individual. Como se aprecia en la tabla, se consigue una reducción considerable tanto en el tiempo total requerido como en el *delay* de los mensajes. La reducción significativa del máximo *delay* se debe a que el tiempo de *delay* es acumulativo y con 20 réplicas ninguno de los mensajes tiene que esperar a que el módulo envíe previamente otros mensajes, como sí ocurre cuando se tienen menos réplicas disponibles.

En el envío de 1000 mensajes no se produce un aumento en el tiempo total proporcional al número de mensajes enviados respecto al envío de 100 mensajes, esto se debe a que, como hemos dicho, el sistema está orientado al envío masivo de mensajes y hay más que ganar en el procesado de mensajes que en la gestión de los *bind*.

Durante el proceso de pruebas, también se probaron tres clientes de forma simultánea de forma que cada cliente era atendido por una instancia del proxy. Se comprobó que era mejor que un cliente inicie de forma simultánea 3 sesiones y envíe 100 mensajes a través de cada sesión que únicamente abrir una sesión y envíe 300 mensajes a través de ella. Esto es debido a que a pesar de que la carga de trabajo es la misma, el proceso de recibir los mensajes, enviarlos al servicio de envio_mensaje y el envío de los

DLR se realiza concurrentemente entre las 3 réplicas y no de forma secuencial. Esto era de esperar ya que como explicamos anteriormente, los cliente contratan con SONOC un número de sesiones disponibles, por tanto, contratar un número mayor se traduce en una mayor capacidad de envío simultáneo de mensajes.

VI. Conclusiones y líneas futuras

A. *Conclusiones.*

Vamos a resumir lo desarrollado en este trabajo:

En primer lugar, se ha desarrollado un módulo balanceador de carga para el protocolo de SMPP que se divide en dos aplicaciones encargadas de distintas funcionalidades; una de ellas para procesar los mensajes SMS para optimizar su envío y la otra para gestionar las conexiones SMPP de los clientes. Esto ha requerido estudiar y comprender el protocolo SMPP y qué funcionalidades de la biblioteca JSMPP eran necesarias. También ha sido clave entender el escenario previo de partida del trabajo para ser capaces de integrar las nuevas funcionalidades de forma que se obtuviesen los mejores resultados para los clientes.

A continuación, se ha pasado a la implementación del escenario desarrollado en un entorno de Kubernetes. Esto ha afianzado los conocimientos obtenidos de las aplicaciones anteriormente mencionadas, pues hemos sido capaces de configurarlas en distintos entornos para ofrecer la misma funcionalidad con una mejora en la fiabilidad y disponibilidad. Por supuesto, el mayor conocimiento que hemos obtenido ha sido en el despliegue de aplicaciones en Kubernetes, para lo cual hemos realizado un estudio de los diferentes elementos de un *cluster*, los hemos valorado y hemos hecho uso de lo necesario para ofrecer las funcionalidades del escenario anterior. Esto también ha supuesto la comprensión del concepto de contenedor y de su desarrollo en el motor de Docker.

Por último, dados los resultados de las pruebas realizadas se ha observado que es más interesante centrarse en la mejora del procesamiento de mensajes que en la gestión de las sesiones con los clientes.

B. *Líneas futuras.*

La mejora más clara para nuestro sistema es eliminar el componente IRISGW, de forma que sean los módulos de envio_mensaje los que en vez de mandar una petición HTTP contra IRISGW manden el mensaje a Kafka, de esta forma mejoramos el tiempo de envío y procesamiento del mensaje además de hacer el sistema mucho más escalable al eliminar un cuello de botella en la capacidad de procesamiento de IRISGW.

Actualmente los datos que usa la aplicación para la comprobación y gestión de los cliente se insertan estáticamente en el código, una implementación futura y para la que el código ya está preparado sería que la aplicación consumiera la información de una cache Redis que se actualizase de forma dinámica la información y pudiese ser modificada también por el resto de módulos de IRIS.

Otra línea de desarrollo futuro es la relacionada con el despliegue del sistema. Actualmente se está desplegando en un único *cluster* para realizar las pruebas, sin embargo puesto que la empresa SONOC tiene múltiples clientes alrededor del mundo es necesario que se despliegue el sistema en múltiples *cluster* por todo el mundo de forma que todos los usuarios tengan buena calidad de servicio.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos NEWLAN (PID2022-136476OB-100) del gobierno de España, CeNIT (T31_23R) del Gobierno de Aragón y NeWLAN (UZ2022-IAR-08) de la Universidad de Zaragoza.

Referencias

[1] *Vodafone pone fin a los mensajes MMS que mejoraron los SMS con imágenes y vídeo.* https://bandaancha.eu/articulos/vodafone-finiquita-mensajes-chat-rcs-10475. consultado en 29/05/2023. Consultado en 06/06/2023

[2] *Bulk messaging.* https://en.wikipedia.org/wiki/Bulk_messaging. Consultado en 06/06/2023

[3] *¿Qué son los A2P SMS? Beneficios de los mensajes A2P para empresas.* https://www.sopranodesign.com/es/que-son-los-a2p-sms-beneficios-de-los-mensajes-a2p-para-empresas/. Consultado en 06/06/2023

[4] *El SMS A2P: la salvación del mensaje de texto que no beneficia a las operadoras.* https://www.xatakamovil.com/movil-y-sociedad/sms-a2p-salvacion-mensaje-texto-que-no-beneficia-a-operadoras. Consultado en 06/06/2023

[5] *SMS Gateway.* https://en.wikipedia.org/wiki/SMS_gateway. Consultado en 06/06/2023

[6] *SMPP Protocol: API to enable SMS messaging between applications and mobiles.* https://smpp.org/. Consultado en 06/06/2023

[7] Diferencias entre los contenedores y las máquinas virtuales. https://www.redhat.com/es/topics/containers/containers-vs-vms. Consultado en 06/06/2023

[8] ¿Qué es Docker y cómo funciona? https://www.redhat.com/es/topics/containers/what-is-docker. Consultado en 06/06/2023

[9] *¿Qué es Kubernetes?* https://kubernetes.io/es/docs/concepts/overview/what-is-kubernetes/ Consultado en 06/06/2023

[10] *¿Qué es Kong Gateway?* https://es.linkedin.com/pulse/qu%C3%A9-es-kong-gateway-ithreexglobal. Consultado en 06/06/2023

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

NFT para la Gestión de Recetas Médicas

Josep Genovard Oliver, Macià Mut Puigserver, M. Magdalena Payeras Capellà, Jaume Ramis Bibiloni
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
Carretera de Valldemossa, Km. 7,5, 07122, Palma
josep.genovard@uib.cat, macia.mut@uib.cat, mpayeras@uib.cat, jaume.ramis@uib.cat

En los últimos años, la aparición de los Non-Fungible Tokens (NFTs) ha revolucionado el mundo de la blockchain. Cada uno de estos activos digitales es único y puede ser transferido. Sus aplicaciones más conocidas son las que representan la propiedad de un bien físico o digital. En este artículo se plantea un protocolo que utiliza la tecnología blockchain y los NFTs para representar la propiedad de las recetas médicas y gestionar su uso, lo que permite resolver algunos de los problemas actuales como la falsificación o la duplicación y plantea al usuario como su único propietario. El protocolo diseñado e implementado tiene un actor principal (Autoridad médica) que gestiona al resto (hospitales, médicos, usuarios y farmacias), que interactuaran entre sí para cumplir con la funcionalidad esperada.

***Palabras Clave*—Blockchain, NFT, Smart Contract, Recetas, Medicamento, Salud**

I. INTRODUCCIÓN

Según la Organización Mundial de la Salud (OMS), la eHealth consiste en la utilización de las tecnologías de la información y las comunicaciones en el ámbito de la salud y otros relacionados con ella. Su objetivo es pues la mejora de las condiciones sanitarias de los ciudadanos y del desempeño de la labor del personal médico. Ejemplos de las tecnologías utilizadas incluyen la impresión 3D para simular partes del cuerpo, máquinas de monitorización en tiempo real y remoto, así como la tecnología blockchain [1]. La blockchain ofrece la capacidad de crear registros de salud electrónicos en propiedad del paciente, verificar las credenciales del personal médico, mejorar la seguridad IoT para la monitorización remota, y hacer un seguimiento de ensayos clínicos y productos farmacéuticos. En los últimos años se ha empezado a explorar el potencial de la aplicación en el área de salud [2].

Una de las propuestas más pioneras en el entorno sanitario fue la aplicación MedRec [3]. En este trabajo se explora la aplicación de registros médicos electrónicos en un entorno descentralizado.

El ámbito especifico de las recetas médicas ha tardado más en investigarse, pero ya se han publicado algunos protocolos para este fin. Un ejemplo es el trabajo *"Safe Prescription": A decentralized blockchain protocol to manage medical prescriptions* [4]. Este último crea un protocolo de gestión de recetas médicas mediante tokens, aunque la estructura no jerárquica utilizada, a diferencia de la nuestra, limita la posibilidad de crecer.

En el trabajo presentado en [5] se propone utilizar la tecnología de los Non-Fungible Tokens (NFTs) para diseñar un protocolo que permita a los médicos crear recetas a nombre de pacientes, a los pacientes gestionar estas recetas y a las farmacias validar las recetas recibidas para poder entregar el medicamento.

La tecnología de los NFTs ya se ha utilizado en diferentes ámbitos para representar la propiedad de bienes físicos o digitales y como clave de servicio o contenido, entre otros. La propiedad de representación de un bien se puede aplicar al sistema de recetas, ya que una receta no es más que un conjunto de datos a nombre del paciente y firmadas por el médico.

Actualmente, el Sistema Nacional de Salud de España no incluye los centros privados en la gestión de recetas médicas. Además, el uso de su servicio centralizado reduce la seguridad del sistema. El uso de una blockchain distribuida, en cambio, ofrece ventajas significativas para el sistema de gestión de recetas médicas. Al permitir el cambio de propiedad de las recetas al mismo usuario recetado, se eliminan intermediarios innecesarios y se mejora la seguridad de los datos. Además, la tecnología de los NFTs permite la creación de tokens que representan recetas médicas individuales y únicas, lo que garantiza su autenticidad y previene la falsificación o duplicado de las mismas.

El protocolo diseñado para la gestión de recetas médicas basado en blockchain tendrá múltiples beneficios para los pacientes, médicos y farmacias. Por ejemplo, los pacientes podrán acceder a sus recetas médicas desde cualquier lugar del mundo, y las farmacias podrán validar las recetas de cualquier hospital registrado en el sistema.

El objetivo del protocolo es mejorar el sistema de gestión de recetas actual: mejorando la seguridad de los datos, eliminando la posibilidad de fraude por falsificación o duplicado de recetas y estandarizando un sistema de

recetas válido para todos los hospitales. En concreto, se mejorará con el uso de las tecnologías Blockchain y NFT.

Finalmente, este protocolo ya se ha implementado en una red de prueba de Ethereum utilizando el lenguaje de programación Solidity, y se presenta brevemente en el artículo. Además, los mismos autores han creado una interfaz gráfica para interactuar de manera amigable con la blockchain y se ha comprobado su buen funcionamiento. Así mismo, se han analizado los costes de las diferentes posibles acciones.

La estructura del presente trabajo es la siguiente: la Sección II describe los actores involucrados, el formato propuesto para las recetas, el protocolo diseñado, así como su implementación. En la Sección III se detallan las propiedades que aporta el sistema propuesto de gestión de recetas. A continuación se calcula el coste en *gas* para el despliegue de los smart contracts y para sus diferentes funcionalidades. Finalmente, las principales conclusiones se exponen en la Sección V.

II. PROTOCOLO

En la actual Sección se presenta el protocolo diseñado. Este protocolo deberá permitir el registro y la gestión de los diferentes actores involucrados, definidos en el Apartado A de esta misma Sección, y una fácil gestión de todo el recorrido de las recetas. La estructura de la receta creada para este proyecto se explica en el Apartado B. El protocolo, presentado en mayor detalle en el Apartado C, se divide, principalmente, en las siguientes cinco fases y/o conjunto de funciones:

1) Gestión de hospitales, usuarios y farmacias.
2) Gestión de médicos del hospital.
3) Creación de recetas.
4) Entrega de las recetas.
5) Envío de las recetas a la autoridad.

A. Actores involucrados

El protocolo diseñado contempla cinco actores diferentes. Además, todos ellos son indispensables para completar el ciclo de la receta, ya que cada función solo la puede ejecutar un único tipo de actor.

- **Autoridad jerárquica**: entidad encargada de validar la identidad de los hospitales, usuarios y farmacias que quieran acceder al sistema, e introducirlos en él, y de eliminarlos cuando sea necesario. Además, recibirá las recetas utilizadas para tener el comprobante de los medicamentos que ha vendido cada farmacia. Solo habrá un actor de este tipo y será el propietario de la aplicación desplegada. En un ámbito nacional, el Ministerio de Salud podría ocupar este papel, ya que se trata de una entidad reconocida en el ámbito. Si por lo contrario se decide continuar con el sistema actual, donde cada comunidad autónoma está separada, en cada una se tendría que asignar una autoridad jerárquica. Por lo tanto, en el caso de separar por comunidad autónoma, se desplegaría un contrato por cada una de estas.
- **Hospital**: entidad encargada de gestionar su propia plantilla de médicos. Cada hospital tendrá un solo actor de este tipo, la función del cual se limita a poder añadir y eliminar actores *médicos* a su propia plantilla. Los centros de atención primaria o consultas privadas también podrán registrarse como entidades de este tipo.
- **Médico**: trabajador de un hospital registrado que solamente podrá crear recetas para usuarios registrados.
- **Usuario**: usuario, o paciente, del sistema de salud que podrá recibir las recetas, de los médicos registrados, a través del sistema. Los usuarios serán los dueños de sus recetas por lo que podrán gestionarlas una vez recibidas, enviándolas a un farmacia o rechazándolas.
- **Farmacia**: entidad que podrá recibir las recetas de los usuarios y deberá comprobar su validez antes de aceptarlas y entregar el medicamento. También podrá enviar las recetas recibidas a la autoridad jerárquica para finalizar el ciclo de la receta.

B. Formato de la receta

El sistema diseñado se basa en la comunicación entre médicos, usuarios y farmacias para una correcta operación de las recetas de medicamentos. Esta receta se ha creado a partir del estándar de NFTs ERC-721 [6], adaptando las recetas actuales.

La Tabla I muestra que las recetas tendrán un identificador y registrarán: el usuario para el cual se crean, el médico que las ha creado, la fecha de caducidad de esta e información sobre la receta, como el Identificador Único del Medicamento (IUM), el período entre tomas, o la duración de esta. Además, las recetas tendrán un estado que irá cambiando según se vaya interactuando con ella (*creada*, *enviada*, *tramitada*, *finalizada* y *repudiada*).

Actualmente, algunas recetas están parcial o totalmente subvencionadas por el estado. Entonces, el último campo de la Tabla I, *subvención*, indica el porcentaje de subvención del estado sobre la receta. Así, cuando un usuario recoja un medicamento, solo deberá abonar $precioProducto \cdot \left(\frac{100-subvencion}{100}\right)$.

De todos modos, se tiene que entender que este trabajo se centra en la implementación del sistema y se prevén cambios en la estructura de la receta para un mejor encaje con las actuales.

Tabla I
TABLA DE DATOS DE UNA RECETA.

Nombre	Tipo	Explicación
id	uint	Identificador único de la receta.
estado	enum	Estado actual de la receta.
caducidad	uint	Fecha de caducidad de la receta.
idUsuario	address	Identificador del usuario al que pertenece la receta.
idMedico	address	Identificador del médico que creo la receta.
ium	string	Identificador único del medicamento.
intervalo	uint8	Intervalo entre tomas (en horas).
duracion	uint8	Días de tratamiento.
subvencion	uint8	Porcentaje de subvención por el estado/aseguradora.

C. Descripción del protocolo

El protocolo se divide en cinco operaciones principales, presentes en la Fig. 1 y explicadas a continuación. Además de las operaciones principales, se han diseñado funciones extra, por ejemplo de visualización de los datos, para dotar al protocolo con una interfaz más amigable.

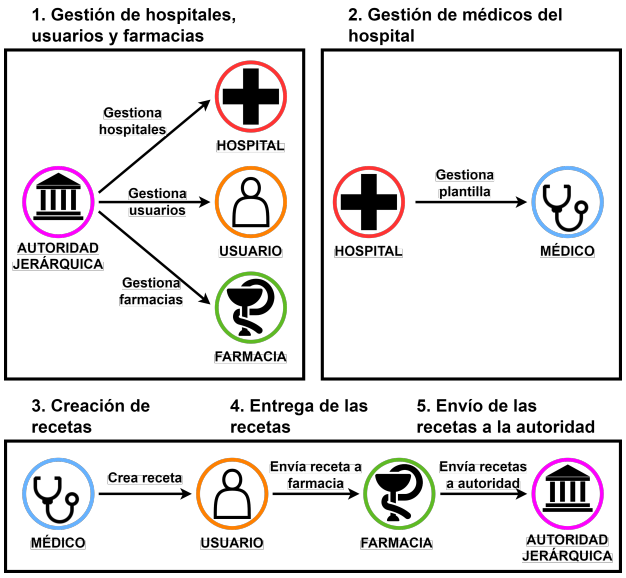


Fig. 1. Funciones generales del protocolo.

1) **Gestión de hospitales, usuarios y farmacias**. Al desplegar el sistema, el único actor con acceso a él será la autoridad jerárquica. Esta se encargará de registrar a los hospitales, usuarios y farmacias que quieran acceder al sistema. Los diferentes actores tendrán que solicitar el acceso por un canal alternativo facilitando los datos necesarios para una correcta identificación y, cuando la autoridad decida registrarlos, solo deberá ejecutar la función correspondiente, introduciendo ciertos datos como la dirección de la cartera y otros datos identificativos, que guardará el actor en una lista de actores registrados. De la misma forma en la que la autoridad puede registrar nuevos actores, también tendrá funciones para eliminar los diferentes actores, por lo que un actor podrá ser eliminado por petición propia o sin ella.
2) **Gestión de médicos del hospital**. Los hospitales serán registrados por la autoridad, pero estos no pueden crear nuevas recetas; la única función del hospital será la de gestionar su propia plantilla de médicos. En otras palabras, los hospitales podrán crear médicos, que automáticamente se incorporaran a una lista propia de médicos, y eliminar los médicos de esta, previniendo que un hospital pueda eliminar médicos externos. La gestión de médicos es igual a la de los otros actores; el sistema incluirá una función para registrar médicos, que en este caso también se encargará de incorporarlos en la plantilla del hospital, y otra para eliminarlos. Además, si un hospital es eliminado del sistema, todos los médicos de su plantilla también los serán, no dejando así la posibilidad de que existan médicos sin control, ya que el hospital que los registró es el único actor que los podrá eliminar.
3) **Creación de recetas**. Las recetas solo pueden ser creadas por médicos registrados. Los médicos solo podrán ejecutar esta función, en la que deberán aportar todos los datos de la receta, detallados en el anterior Apartado B de esta misma Sección, aparte del *id*, *estado* e *idMédico*, que se asignan automáticamente. La función encargada de crear recetas solo podrá ser ejecutada por un médico registrado y, en primer lugar, comprobará el correcto registro del médico correspondiente. A continuación se creará la receta, en forma de NFT, para el usuario en cuestión, asignándole los datos. Una vez creada, la receta será propiedad del usuario indicado y solo este podrá gestionarla.
4) **Entrega de las recetas**. Los usuarios, propietarios de sus propias recetas, serán los únicos que podrán gestionarlas. Eso se traduce en visualizarlas, repudiarlas y enviarlas a una farmacia. Al ejecutar la función de enviar una receta, se tendrá que indicar la receta y la dirección de la farmacia a la cual se desea entregar. La función validará que la receta existe, es propiedad del usuario que ejecutó la función y que la dirección es propiedad de una farmacia registrada en el sistema. Una vez recibida, la farmacia ejecutará una función de *validar receta*, que comprobará la fecha de caducidad de esta, el médico que la creó y el usuario para el cual estaba creada. Si el resultado es correcto, la receta se moverá a una tabla de recetas tramitadas y se podrá entregar el medicamento al usuario, mientras que en caso contrario se eliminará.
5) **Envío de las recetas a la autoridad**. Después de haber sido tramitadas, las recetas seguirán a nombre de la farmacia. Estas recetas deben entregarse a la autoridad jerárquica para cerrar el ciclo de las recetas (simulando el posible cobro de la subvención por parte del Gobierno o una aseguradora privada). El envío de las recetas a la autoridad está formada por una única función que podría automatizarse y ejecutarse cada cierto tiempo. Entonces, al ejecutar esta función se enviarán todas las recetas tramitadas y la autoridad deberá abonar el importe subvencionado. El método de pago queda fuera del ámbito de este documento.

Aparte de las funciones incluidas en estos 5 grupos, también son necesarios un conjunto de funciones de visualización de diferentes parámetros, como la visualización de la plantilla por parte de un hospital o de las recetas en propiedad por parte del usuario.

D. Implementación

Una vez descrito el protocolo, se presentará la implementación que se ha hecho de este. Es importante mencionar que cada actor que desee utilizar el sistema

deberá tener una cuenta de la blockchain seleccionada. Cada tipo de actor contará con un smart contract diferente, con un conjunto de operaciones disponibles. En la Fig. 2 se observa la distribución de actores y contratos, con las interacciones que hay a través de ellos.

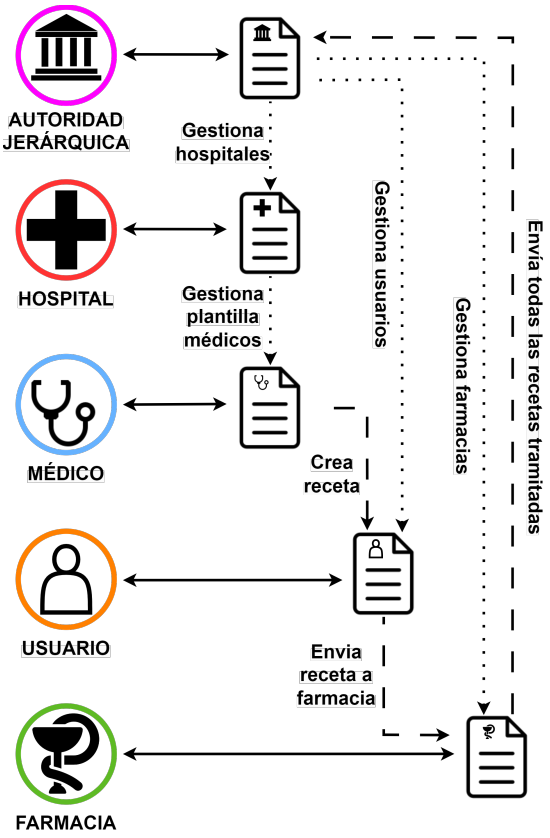


Fig. 2. Distribución e interacción de los smart contracts.

Las líneas continuas indican la interacción entre actor y smart contract; las discontinuas de puntos hacen referencia a la gestión de usuarios, por lo que no se interacciona entre contratos pero se modifican los permisos de acceso a estos; y las discontinuas de rayas hacen referencia a una acción realizada en un contrato para el propietario del siguiente. Estas interacciones descritas en la Fig.2 definen el ciclo de la receta.

Además de estos smart contracts, se ha creado un contrato para el token de receta. Este define la estructura de las recetas, con los datos de la Tabla I, e implementa las funciones necesarias para poder interactuar con ella (como crear, repudiar o enviar). Estas funciones no podrán ser accedidas directamente por un actor sino desde los contratos de los respectivos actores.

Para que los contratos puedan interactuar entre ellos es necesario que conozcan sus respectivas direcciones. Por eso se debe seguir un cierto orden para desplegarlos:

1) Se despliega el contrato del token de receta.
2) Se despliega el contrato de la autoridad jerárquica, indicando la dirección del contrato de recetas.
3) La autoridad ejecuta la función *despliegaTodosSC* de su contrato, que desplegará los contratos restantes. La misma función se encarga de compartir las direcciones entre los contratos que deben interactuar entre ellos.

El primer elemento del Apartado C de esta misma Sección (*Gestión de hospitales, usuarios y farmacias*) se implementa en el contrato de la autoridad jerárquica. El Listado 1 enseña el código utilizado para dar de alta y de baja a un actor de tipo *Hospital*, utilizando una estructura simplificada en que solo se requiere almacenar el nombre y la dirección de la cartera de este.

```
1  function creaHospital(address _aHospital,
     string memory _nombre) public
     onlyByAutoridad(msg.sender){
2    require(bytes(_nombre).length > 2, "Indicar
     nombre completo");
3    Hospital memory h;
4    h.estado = estadoActor.alta;
5    h.nombre = _nombre;
6    HospitalesMap[_aHospital] = h;
7  }
8
9  function bajaHospital(address _aHospital)
     public onlyByAutoridad(msg.sender){
10   require(HospitalesMap[_aHospital].estado ==
     estadoActor.alta, "Hospital no esta de alta"
     );
11   HospitalesMap[_aHospital].estado =
     estadoActor.baja;
12 }
```

Código 1. Funciones de gestión de hospitales del contrato *autoridad.sol*.

Anteriormente se ha comentado que los actores no interactuarán directamente con el contrato de recetas. En cambio, lo harán a través de las funciones del contrato que les corresponde. Por ejemplo, la función *creación de recetas* (tercer elemento del Apartado C dentro de la anterior Sección), que se detalla en el Listado 2, se encuentra en el contrato de recetas, pero el médico ejecutará la función del Listado 3, que se encuentra en el smart contrat de los médicos.

En el resto de funciones que modifican o consultan los datos de la receta también se tendrá que interactuar indirectamente con el contrato *tokeReceta.sol* a través del contrato correspondiente a cada actor. Por ejemplo, para el envío de las recetas de usuario a las farmacias, los usuarios deberán ejecutar una función de su propio contrato que llamará a la enseñada en el Listado 4.

El resto de funciones que interactúan con la receta tienen una logística similar a la enseñada en los dos ejemplos, correspondientes a la creación y al envío de una receta. Además, en cada función se restringe el acceso y la funcionalidad para evitar un mal uso. Por ejemplo, la función de envío, previamente enseñada en el Listado 4, comprueba que el usuario es el verdadero propietario de la receta que quiere enviar y que la dirección de envío pertenece a una farmacia registrada, evitando el intercambio de recetas entre usuarios.

El despliegue de la aplicación en la red Ethereum, con su interfaz gráfica hecha con ReactJS, se puede realizar mediante nodeJS o Docker. Para acceder a la DApp se tendrá que disponer de una cartera en el navegador, como

```
function crearReceta(address _addressUsuario,
 address _addressMedico, string memory _ium,
 uint16 _anoCaducidad, uint8 _mesCaducidad,
 uint8 _diaCaducidad, uint8 _intervalo, uint8
  _duracion, uint8 _subvencion) public
 onlyByContractos(msg.sender) {
 require(stringLength(_ium) == 13, "El IUM
 debe tener 13 digitos");

  _tokenIds.increment();                    //
 Incremento del id
  uint auxIdReceta = _tokenIds.current();   //
 Se guarda el id

  DatosReceta memory dr;
  dr.estado = estadoReceta.creada;
  if(_anoCaducidad > 0 && _mesCaducidad > 0 &&
  _diaCaducidad > 0) {
      dr.caducidad = DateTime.toTimestamp(
 _anoCaducidad, _mesCaducidad, _diaCaducidad
 +1);
  } else {
      dr.caducidad = 2147483647;            //
 Ultimo numero unixtime
  }
  dr.idUsuario = _addressUsuario;
  dr.idMedico = _addressMedico;
  dr.ium = _ium;
  dr.intervalo = _intervalo;
  dr.duracion = _duracion;
  dr.subvencion = _subvencion;
  DatosRecetaMap[auxIdReceta] = dr;

  // Minado del token
  _mint(dr.idUsuario, auxIdReceta);

  PosicionArrayaMap[auxIdReceta] =
 recetasPropietarioMap[dr.idUsuari].ids.
 length; // Guardar posicion del array en la
 que se guardara el id
  recetasPropietarioMap[dr.idUsuario].ids.push
 (auxIdReceta);
}
```

Código 2. Función del contrato *tokenReceta.sol* que crea una receta nueva.

```
function creaReceta(address _addressUsuario,
 string memory _ium, uint16 _anoCaducidad,
 uint8 _mesCaducidad, uint8 _diaCaducidad,
 uint8 _intervalo, uint8 _duracion, uint8
 _subvencion) public onlyByMedicos(msg.sender
 ){
 tokenReceta.crearReceta(_addressUsuario, msg.
 sender, _ium, _anoCaducidad, _mesCaducidad,
 _diaCaducidad, _intervalo, _duracion,
 _subvencion);
}
```

Código 3. Función del contrato *tokenReceta.sol* que envía una receta de usuario a una farmacia registrada.

MetaMask [7]. Al acceder a la aplicación, esta utilizará la dirección de la cartera para determinar el tipo de actor y dirigir al usuario a una página específica donde solo se verán las funciones que tenga disponibles el actor. Finalmente, cuando un usuario quiera realizar una función (como crear un actor, crear una receta o enviar una receta a la farmacia) tendrá que aceptar la transacción de MetaMask con su costo.

```
function enviaReceta(uint _idReceta, address
 _adUsuario, address _adFarmacia) public
 onlyByContractos(msg.sender){
 require(ownerOf(_idReceta) == _adUsuari, "
 Solo el propietario");
 require(datosRecetaMap[_idReceta].estado ==
 estadoReceta.creada, "Solo recetas con
 estado \"Creada\"");
 require(autoridad.estadoFarmacia(_adFarmacia)
 , "Solo farmacias registradas");

 // Transferencia de la propiedad del token
 transferFrom(_adUsuario, _adFarmacia,
 _idReceta);

 datosRecetaMap[_idReceta].estado =
 estadoReceta.enviada;

 // Se elimina la receta del historial del
 usuario
 uint posicionReceta = PosicionArrayaMap[
 _idReceta];
 delete recetasPropietarioMap[_adUsuario].ids[
 posicioReceta];

 //Se guarda en el historial de farmacia
 PosicionArrayaMap[_idReceta] =
 recetasPropietarioMap[_adFarmacia].ids.
 length;
 recetasPropietarioMap[_adFarmacia].ids.push(
 _idReceta);
}
```

Código 4. Función del contrato *token.sol* que llama a la función del contrato *tokenReceta.sol* para que cree una receta nueva.

III. PROPIEDADES

En esta Sección se describen un total de 8 propiedades que aporta el sistema propuesto de gestión de recetas.

A. Disponibilidad

En el protocolo propuesto, las recetas médicas se almacenan en la blockchain, distribuida en diferentes bases de datos. Al estar distribuida, la caída de un nodo no afecta a la disponibilidad de los datos y del sistema, permitiendo así continuar realizando las diferentes operaciones disponibles.

B. Persistencia/Integridad

En el protocolo propuesto, se garantiza la persistencia e integridad de las recetas médicas, y de los otros datos, gracias a la utilización de la tecnología blockchain. Cada transacción queda registrada de forma permanente y transparente en la blockchain, lo que evita la modificación o eliminación de datos sin dejar un rastro de la manipulación.

C. Interoperabilidad

El protocolo propuesto busca establecer un sistema de gestión de recetas médicas estándar y válido para todos los hospitales. Al utilizar tecnologías como blockchain y NFTs, se proporciona una infraestructura común que puede ser adoptada por diferentes instituciones y sistemas de salud. Esto promueve la interoperabilidad entre los distintos actores del sector de la salud, permitiendo la

comunicación y el intercambio seguro de recetas médicas entre diferentes entidades y sistemas.

D. Propiedad de las recetas y Autogestión

El uso de la tecnología de los Non-Fungible Tokens (NFTs) en el protocolo garantiza la propiedad, por parte del usuario, de las recetas médicas. Cada receta es representada por un token único e indivisible, lo que permite que el paciente sea el propietario legítimo de su receta. Al tener la propiedad de su receta, el usuario puede visualizarla sin la dependencia de ningún tercero y puede repudiarla o elegir dónde dispensarla.

E. No repudio

El protocolo basado en blockchain y NFTs proporciona un mecanismo de no repudio de las acciones realizadas en el sistema de gestión de recetas. Cada transacción y modificación realizada en la blockchain queda registrada de forma permanente y transparente. Esto significa que los actores involucrados, como médicos, pacientes y farmacias, no pueden negar su participación o responsabilidad en las acciones realizadas. La información registrada en la blockchain se convierte en una evidencia inmutable y verificable, lo que refuerza la confianza y la integridad del sistema.

F. Inmutabilidad

La tecnología blockchain proporciona inmutabilidad a los datos almacenados en el sistema de gestión de recetas médicas. Una vez que una receta se registra en la blockchain, no se puede modificar ni eliminar sin dejar un rastro de la modificación. Esto garantiza la integridad de las recetas y evita la manipulación o alteración de los datos a lo largo del tiempo. La inmutabilidad de los datos en la blockchain proporciona confianza en la información almacenada y previene la posibilidad de fraudes o alteraciones maliciosas.

G. Trazabilidad

El protocolo propuesto garantiza la trazabilidad completa de las recetas médicas. Cada receta y todas las transacciones relacionadas, como la creación, modificación o validación de una receta, quedan registradas en la blockchain. Esto permite un seguimiento detallado de cada etapa del proceso de gestión de las recetas, desde la creación hasta el envío a la autoridad jerárquica.

H. Privacidad

Aunque el sistema implementado no cifra las diferentes recetas, se contempla la opción de poder aplicar cifrados que permitan visualizar los datos de una receta solamente al propietario de esta.

IV. COSTES

Para la Sección actual se ha calculado el costo en *gas* para el despliegue de los smart contracts y para sus diferentes funcionalidades. En el contexto de los smart contracts, el *gas* se refiere a una medida de costo computacional necesaria para ejecutar operaciones en la plataforma blockchain. Cada instrucción u operación que se realiza en un contrato inteligente consume una cantidad específica de *gas*. Las unidades del *gas* son *Gwei* y, en la plataforma de Ethereum, se paga con *ether* ($1ether = 10^9 Gwei$). El precio de *ether* a pagar se calcula de la siguiente forma: $ether = gas \cdot tasaGas$, donde la *tasaGas* influirá en el precio pero también a la velocidad de minado (a mayor tasa, más rápido será el minado).

Aunque las funciones más costosas son el despliegue de los diferentes contratos, solo se realizarán una sola vez. Además de los 3 despliegues (*TokenReceta.sol*, *Autoridad.sol* y el resto de contratos, que se despliegan ejecutando una función de la autoridad), se han elegido 8 funciones representativas para estimar el precio del funcionamiento de este sistema en una plataforma pública, como Ethereum. Para ello se ha utilizado el precio del *ether* de día 28 de mayo de 2023, 1719 EUR/ETH, y se han calculado en los casos de utilizar la tasa de *Gwei* media (25 *Gwei* = 3 min) y máxima (30 *Gwei* = 30 seg) de este mismo día. La Tabla II muestra el gas y el precio, en euros, que supondría el despliegue de cada una de la funciones en la plataforma Ethereum.

Se puede apreciar que los precios del funcionamiento del sistema propuesto en la red Ethereum son inaceptables. La intención del proyecto no es funcionar sobre la mainnet de Ethereum, red principal; sino sobre una red permisionada o una sidechain [9] con menor coste transaccional, como Polygon [8]. En las dos últimas columnas de la Tabla II también se han plasmado los costes de las mismas funciones del sistema utilizando esta última plataforma. Para ello se ha utilizado el precio de la moneda de la plataforma (MATIC) del mismo día 28 de mayo de 2023, 0,8598 EUR/MATIC, y la tasa media (150 *Gwei* = 10-30 seg) y máxima (155 *Gwei* = 5-10 seg) de este mismo día.

Los precios de las diferentes funciones son perfectamente asumibles utilizando la plataforma Polygon ya que, aun utilizando la mayor tasa de gas, el coste del despliegue de todos los contratos costaría poco mas de 1,50 €. Además, el ciclo completo de una receta (creación, envío a farmacia, tramitación y envío a la autoridad) costaría, en el peor de los casos, poco más de 0,10 €.

Otro detalle a tener en cuenta es que el sistema no debería suponer un coste económico para los actores al utilizarla, principalmente para los usuarios. Por eso, una posible solución se basa en desplegar la aplicación en una red blockchain privada o permisionada, como Alastria, sobre la que las transacciones no tienen un coste económico directo, o incluso crear una blockchain específica para su uso.

V. CONCLUSIONES

En el presente artículo se ha utilizado la tecnología blockchain y los NFTs para plantear una mejora, en distintos aspectos, en el sistema sanitario actual. En la propuesta, la descentralización de los datos permite que distintos centros hospitalarios y farmacias, con bases de datos separadas, puedan acceder a los mismos datos. Estos datos estarían blindados frente a ataques de integridad ya

Tabla II
TABLA DEL COSTE DE EJECUCIÓN EN ETHEREUM Y POLYGON.

		Ethereum		**Polyigon**	
Función	Gas (weis)	EUR (25 Gwei)	EUR (30 Gwei)	EUR (150 Gwei)	EUR (155 Gwei)
TokenReceta.sol	3 098 953	133,178	159,813	0,400	0,413
Autoridad.sol	3 202 286	137,618	165,142	0,413	0,427
Resto contratos	2 409 976	103,569	124,282	0,311	0,321
Alta Hospital	47 949	2,061	2,473	0,006	0,006
Alta Médico	156 178	6,712	8,054	0,020	0,021
Baja hospital	46 376	1,993	2,392	0,006	0,006
Creación receta	303 571	13,046	15,655	0,039	0,040
Rechaza receta	77 636	3,336	4,004	0,010	0,010
Envío de receta a farmacia	141 661	6,088	7,305	0,018	0,019
Tramitación de receta	64 858	2,787	3,345	0,008	0,009
Envío recetas a la autoridad	261 933	11,257	13,508	0,034	0,035

que se duplicarían en diferentes bases de datos. Además, los usuarios pasan a ser los propietarios de sus recetas y pueden gestionarlas y visualizarlas. Los requisitos para el protocolo, todos satisfechos, eran:

- Permitir la interacción entre diferentes actores para poder cumplir el ciclo completo de la receta; del médico a la autoridad, pasando por el usuario y una farmacia.
- Evitar los sistemas centralizados actuales para poder utilizarse en diferentes sistemas hospitalarios y para otorgar la propiedad de la receta al propio usuario.
- Evitar acciones maliciosas como: enviar recetas entre usuarios, utilizar dos veces una misma receta o modificar los datos de una receta.

También se ha observado que el coste económico de este sistema se podría asumir perfectamente.

La interfaz actual aporta una completa funcionalidad al sistema, aunque para la implementación definitiva se podría facilitar el uso de esta, por ejemplo utilizando un QR en las farmacias para que los usuarios pudieran escanearlo para enviar la receta.

El problema real es que todos los usuarios del sistema deberían tener acceso a una cartera virtual de la blockchain correspondiente y tener algún dispositivo con el que se pueda gestionar la receta (móvil, tablet o ordenador), lo que a día de hoy no se cumple.

AGRADECIMIENTOS

Esta publicación es parte del proyecto de I+D+i PID2021-122394OB-I00 (Blobsec), financiado/a por MCIN/ AEI/10.13039/501100011033/ y por “FEDER Una manera de hacer Europa”.

REFERENCIAS

[1] Priti Tagde, Sandeep Tagde, Tanima Bhattacharya, Pooja Tagde, Hitesh Chopra, Rokeya Akter, Deepak Kaushik y Md. Habibur Rahman ”Blockchain and artificial intelligence technology in e-Health”, Environ Sci Pollut Res 28, 52810–52831, 2021.

[2] T. Kumar, V. Ramani, I. Ahmad, A. Braeken, E. Harjula and M. Ylianttila, ”Blockchain Utilization in Healthcare: Key Requirements and Challenges,” 2018 IEEE 20th International Conference on e-Health Networking, Applications y Services (Healthcom), Ostrava, Czech Republic, 2018, pp. 1-7, doi: 10.1109/HealthCom.2018.8531136.

[3] A. Azaria, A. Ekblaw, T. Vieira y A. Lippman, ”MedRec: Using Blockchain for Medical Data Access and Permission Management,” 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 2016, pp. 25-30, doi: 10.1109/OBD.2016.11.

[4] Claudio Cilli, Giulio Magnanini, Marco Silipigni y Fabrizio Venettoni. (2021). ”Safe Prescription”: A decentralized blockchain protocol to manage medical prescriptions.

[5] Ethereum, ”Non-fungible tokens (NFT)”, https://ethereum.org/en/nft/.

[6] William Entriken, Dieter Shirley, Jacob Evans y Nastassia Sachs, ”ERC-721: Non-Fungible Token Standard”, https://eips.ethereum.org/EIPS/eip-721.

[7] MetaMask, ”The crypto wallet for Defi, Web3 Dapps and NFTs”, https://metamask.io/.

[8] Polygon Technology, ”Blockchains for mass adoption”, https://polygon.technology/.

[9] Amritraj Singh, Kelly Click, Reza M. Parizi, Qi Zhang, Ali Dehghantanha y Kim-Kwang Raymond Choo ”Sidechain technologies in blockchain networks: An examination and state-of-the-art review”, Journal of Network and Computer Applications, 149, 2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

SoulBound Tokens Rechazables para la Asignación de Credenciales y Aceptación de Términos

Rosa Pericàs Gornals, M. Magdalena Payeras Capellà, Macià Mut Puigserver, Miquel À. Cabot Nadal, Llorenç Huguet Rotger
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
Carretera de Valldemossa, Km. 7,5, 07122, Palma
rosa.pericas@uib.cat, mpayeras@uib.cat, macia.mut@uib.cat, miquel.cabot@uib.cat, l.huguet@uib.cat

Las credenciales digitales son emitidas por entidades autorizadas para facilitar la identificación de los usuarios. Estas pueden ser clasificadas en credenciales de identidad, credenciales académicas o profesionales, y credenciales de acceso. La tecnología blockchain ofrece algunas características inherentes que son muy útiles para la gestión de credenciales. Los NFTs parecen encajar perfectamente en la implementación de credenciales digitales. Pero, algunos requisitos cruciales de las credenciales son la no transferibilidad y el hecho que la entidad receptora reciba una aceptación explícita del usuario que va a ser propietario de la credencial, los cuales no se pueden conseguir con la definición actual de los diferentes estándares de NFTs. Sin embargo, actualmente ya existe un nuevo tipo de token, llamado SBT, el cual fue diseñado para asegurar la no transferibilidad de los activos. Este artículo presenta un protocolo enfocado a la emisión de credenciales digitales de acceso, con la habilidad adicional de permitir la asociación de términos y condiciones que los usuarios deben aceptar al recibir las credenciales. Para dicho protocolo hemos utilizado una mejora de los SBTs, conocida por RejSBTs, para representar las credenciales y los términos asociados, proporcionando las pruebas de no repudio en origen y en recepción.

***Palabras Clave*—Credenciales digitales, Identidad, Blockchain, Soulbound tokens**

I. INTRODUCCIÓN

Desde hace años, la digitalización de los activos es cada vez más frecuente, y el concepto de identidad digital también ha seguido esta tendencia. Hoy en día, las entidades expiden multitud de credenciales digitales (CDs) para facilitar la identificación digital de sus usuarios. En concreto, una credencial digital puede definirse como "el equivalente digital de documentos en papel, fichas de plástico (como por ejemplo las fichas intercambiables por dinero en un casino) y otros objetos tangibles emitidos por terceras partes de confianza" [1]. Las credenciales digitales pueden clasificarse en tres categorías principales:

- Credenciales de identidad: representan la identidad de un individuo, como pasaportes digitales o permisos de conducir.
- Credenciales académicas y profesionales: validan los logros académicos y/o profesionales del sujeto.
- Credenciales de acceso: proporcionan autorización de acceso en sistemas o plataformas digitales.

La adopción de la tecnología blockchain es una solución común para implementar aplicaciones seguras con activos digitales. La blockchain ofrece algunas funcionalidades inherentes que son muy ventajosas para este tipo de aplicaciones. Estas características son la inmutabilidad de los datos, la transparencia, la seguridad proporcionada a través del uso de técnicas criptográficas, y la trazabilidad. Con estas características inherentes, la tecnología blockchain proporciona una base importante para asegurar la integridad y seguridad de los activos digitales. Actualmente en la red Ethereum, o en redes blockchain compatibles, ya se han definido numerosos estándares relevantes que encajan muy bien con la implementación de este tipo de aplicaciones. Algunos ejemplos son el estándar ERC-721, el cual define tokens no fungibles (NFTs) o el estándard ERC-20, que define tokens fungibles. Los NFTs son únicos y en general carecen de la posibilidad de ser intercambiados, específicamente los NFTs sirven como representaciones digitales de activos reales, los cuales representan el valor de los bienes o servicios, o bien ofrecen una utilidad basada en el valor intrínseco del propio token, lo cual permite encontrar diferentes tipos de objetos como son los bienes inmuebles, entradas a eventos o certificados digitales. A

primera instancia, puede parecer que los NFTs encajen a la perfección con la implementación de credenciales digitales, debido a su habilidad de representar digitalmente activos digitales. Además, actualmente existen multitud de implementaciones de credenciales digitales basadas en el uso de NFTs. Pero, en el caso de credenciales digitales de acceso, un requisito crucial es la no transferibilidad de la credencial y la aceptación explícita del usuario que va a ser propietario de la nueva credencial. Este tipo de credenciales una vez emitidas necesitan restringir su transferibilidad, con el objetivo de prevenir accesos no autorizados, y además deben proporcionar al usuario receptor la opción de decidir si desea aceptar o denegar su recepción.

Para solucionar este requisito, Buterin et al. [2] recientemente introdujo una variante de los NFTs, conocida como Soulbound tokens (SBTs). El objetivo primordial de los SBTs es el hecho de asegurar la no transferibilidad de los activos. Los SBTs específicamente son diseñados para estar enlazados a la cartera del usuario propietario, que los autores de [2] se refieren como "Souls". Consecuentemente, este nuevo tipo de tokens encaja mejor con los requisitos fundamentales de las credenciales digitales de acceso, proporcionando las limitaciones necesarias de transferibilidad.

Sin embargo, actualmente los SBTs tienen una limitación, ya que no permiten a los usuarios aceptar o rechazar la recepción de tokens, además de aún no tener un estándar específico que implemente la propuesta de Buterin et al. En nuestro trabajo anterior [3], introducimos los NFT rechazables (Rejectable NFTs, RejNFTs), que se corresponden a una mejora del estándar ERC-721, dónde implementamos la funcionalidad de rechazo selectivo de NFTs.

Este artículo presenta un protocolo focalizado en la emisión de credenciales de acceso digitales, con la habilidad adicional de aceptar unos términos y/o condiciones asociados, que los usuarios deben aceptar con la recepción de las credenciales. Para representar las CDs y los términos asociados se hace uso de RejSBTs, proporcionando el no repudio en origen y en destino.

El artículo se organiza en las siguientes secciones. La Sección II introduce las propiedades esenciales de las credenciales digitales de acceso y el estado del arte. La Sección III indica la contribución principal del protocolo diseñado. Seguidamente, en la Sección IV, se presenta el protocolo de emisión de credenciales digitales de acceso. En la Sección V proporcionamos la implementación de smart contracts. La Sección VI evalúa los costes asociados de la implementación. Y finalmente, la Sección VII, presenta las conclusiones del protocolo.

II. PROPIEDADES Y ESTADO DEL ARTE

Las credenciales digitales (CD) fueron por primera vez propuestas por Brands en el año 1993 [4] como un modo seguro de representar objetos del mundo real en formato digital. El National Institute of Standards and Technology (NIST) define la identidad digital como: "La representación única de un sujeto que participa en una transacción en línea. Una identidad digital es siempre única en el contexto de un servicio digital, pero no tiene por qué identificar al sujeto de forma única en todos los contextos. En otras palabras, acceder a un servicio digital puede no significar que la identidad del sujeto en la vida real sea conocida" [5].

Tradicionalmente, las identidades digitales de acceso o credenciales han sido gestionadas de forma centralizada, con un control limitado de los titulares de las identidades sobre estas [6]. Sin embargo, con la introducción de la tecnología blockchain y otras tecnologías descentralizadas, las CDs y las identidades han obtenido la característica de identidad autosoberana (SSI, self-sovereign identity) [7]. La SSI capacita a los titulares de las CDs con el control completo de sus datos, proporcionando valor propio a las CDs, sin la intervención del emisor.

Actualmente, existen multitud de protocolos de gestión de CDs. Por ejemplo, Herbke et al. en [8] presenta un protocolo que aplica el paradigma de identidad autosoberana a las CDs de estudiantes.

Un tipo de protocolo similar de gestión de credenciales digitales está enfocado en la gestión de certificados digitales. Reza et al. [9] propone un sistema basado en el uso de la tecnología blockchain para la gestión tradicional de certificados basados en papel o credenciales en un registro distribuido. El sistema propuesto incorpora un sistema de almacenaje seguro que restringe el acceso únicamente a las partes autorizadas, eliminando la necesidad de intervención de una tercera parte en la certificación de los certificados. En [10], Eltuhami et al. proponen un nuevo enfoque para los sistemas de gestión de certificados que utilizan NFTs no transferibles. Sin embargo, una limitación de su protocolo es la ausencia de un mecanismo que permita al receptor rechazar la transferencia del token. Consecuentemente, podría suceder que un usuario malicioso asocie CDs o certificados con individuos comprometiendo su identidad.

Las propiedades deseadas a proporcionar por un sistema de credenciales digitales de acceso son:

- **Integridad e inmutabilidad**. Los datos almacenados dentro de la CD deben permanecer intactos e inalterados, asegurando que los datos no se pueden modificar o manipular sin ser detectados. Además, los datos no deben poder ser eliminados, y se debe mantener un registro histórico de los diferentes estados.
- **Disponibilidad**. Las CDs y el sistema de gestión asociado debe ser siempre accesible y utilizable.
- **Transferibilidad de evidencia**. La evidencia asociada con la CD emitida se debe poder presentar o transmitir a las partes relevantes, durante los procesos de autenticación o autorización.
- **Interoperabilidad**. El formato de la CD debe poder ser aceptado en diferentes sistemas y plataformas.
- **No repudio**. La entidad emisora de la CD no debe poder denegar el origen y el propietario de la CD no debe poder denegar la recepción de esta.

- **Identidad auto-soberana**. El sistema de CDs debe permitir al propietario gestionarlas y controlarlas, sin la necesidad de intervención de autoridades centralizadas.
- **Recepción selectiva**. El sistema de emisión de CDs debe permitir al usuario receptor decidir si quiere aceptar o rechazar la recepción de una CD.

III. CONTRIBUCIÓN

En la Sección II se han presentado múltiples sistemas de gestión de credenciales digitales basados en el uso de la blockchain. Nuestro artículo se focaliza en dos aspectos muy importantes, los cuales pretenden ayudar a mejorar las propuestas anteriores. Por una parte, los NFTs y los SBTs, tal y como son definidos en sus estándares, no permiten al receptor rechazar la transferencia del token. Entonces, si los NFTs o SBTs son utilizados para proporcionar CDs, el receptor no puede rechazar la recepción. De este modo, todo tipo de CD se podría asociar con la identidad del usuario, incluso sin su consentimiento. Con la propuesta de un nuevo tipo de token, RejSBT, cada CD deberá ser explícitamente aceptada por el receptor (proporcionando la evidencia de no repudio de recepción) y no podrá ser transferida a terceras partes.

Por otra parte, en multitud de aplicaciones, la entrega de CDs se asocia con la aceptación de términos y condiciones de utilización de credenciales. Por este motivo, nuestra propuesta de protocolo incluye la aceptación de términos y condiciones al aceptar la entrega de las credenciales.

IV. DESCRIPCIÓN DEL PROTOCOLO

En este artículo presentamos un nuevo protocolo para enviar credenciales junto a términos de uso, siguiendo el formato de RejSBT, permitiendo a los usuarios poder decidir si desean aceptar o rechazarlos. En esta Sección presentamos una visión general del protocolo propuesto, describiendo los actores involucrados y sus interacciones.

En el protocolo intervienen los siguientes actores:

- Emisor (S): usuario o entidad responsable de emitir las CDs y los términos. Es la entidad encargada de identificar al receptor (R) del token.
- Receptor (R): usuario que recibe la transferencia de la CD y los términos asociados. Su acción principal se basa en la decisión de aceptar o rechazar la recepción de la CD juntamente con los términos.
- Verificador (V): usuario, entidad o sistema que necesita validar la CD proporcionada por R. Solo necesita tener acceso al contenido de la CD y comprobar los datos almacenados en esta.

La CD utilizada está compuesta por diferentes componentes clave. Primero, ésta incluye las credenciales y los correspondientes términos, los cuales encapsulan la información relacionada con la CD que se comparte. Adicionalmente, la CD incorpora una fecha límite (*deadline*), que define el tiempo que tiene R para decidir si aceptar la recepción del token. Esta fecha límite proporciona al receptor un período de tiempo adecuado para tomar la decisión sobre aceptar o rechazar la recepción de la CD. Cabe tener en cuenta que en caso de excederse el *deadline* no se va a realizar la emisión del token en la blockchain. Por último, opcionalmente se puede incluir una fecha de caducidad (*expiry*), la cual actúa como límite, indicando el momento en que la CD va a expirar. Una vez llegada la fecha de caducidad, el token deja de ser válido, enfatizando el aspecto temporal del RejSBT y que la CD que contiene tiene una vida útil limitada.

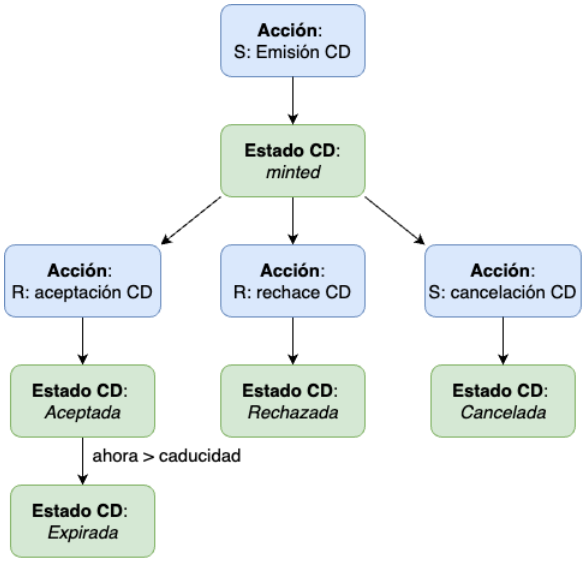


Fig. 1. Diagrama de estados CD RejSBT.

El protocolo presentado para el envío de CDs se compone de los siguientes pasos, los cuales son presentados en la Figura 1:

1) **Emisión de la CD**: el emisor (S) crea el contenido de la CD y define los términos a asociar con la aceptación de la CD. Si se considera necesario o importante para el sistema, S puede cargar tanto las credenciales como los términos en un sistema de almacenamiento descentralizado, como es el IPFS, Interplanetary File System por sus siglas en inglés. Sin embargo, los detalles específicos del proceso de carga no se encuentran en el alcance del protocolo descrito en este artículo. Seguidamente, S realiza la emisión ($mint$) del nuevo token, que contiene la CD y los términos (siendo estos últimos opcionales). Durante el proceso de emisión, S también define dos parámetros cruciales: la fecha límite ($deadline$) y la fecha de caducidad ($expiry$).
2) **Aceptación de la CD**: en caso que R desee aceptar la CD recibida, debe ejecutar la función de aceptación antes de la fecha límite definida. Ejecutando la función de aceptación, R formalmente indica su consentimiento y acuerdo de la CD recibida.
3) **Rechace de la CD**: si R no desea aceptar la CD recibida, tiene la opción de activamente rechazar la transferencia, o simplemente no realizar ninguna acción. La CD pasará a no ser válida una vez alcanzada la fecha límite.

4) **Cancelación de la CD**: si *R* no ha aceptado ni rechazado la CD, *S* tiene la opción de cancelar la transferencia. En esta situación, *S* tiene la autoridad de cancelar el token y anular la transferencia de la CD.
5) **Verificación de la CD**: una vez *R* ha aceptado la recepción de la CD únicamente comprobando la dirección de blockchain de *R*, cualquier verificador *V* puede verificar la propiedad de la CD, todos los metadatos asociados y por lo tanto, proporcionar el acceso adecuado a *R* de acuerdo con el resultado de la verificación de la CD.

V. IMPLEMENTACIÓN

Con la intención de implementar, testear, y evaluar el funcionamiento del protocolo descrito en la Sección anterior, hemos desarrollado un conjunto de smart contracts utilizando el lenguaje de programación Solidity. Estos smart contracts, juntamente con los tests correspondientes, se encuentran en un repositorio de Github dedicado[1].

Para la implementación del protocolo de emisión y gestión de credenciales digitales de acceso, se ha utilizado el smart contract RejSBT para almacenar los datos de la CD. Siguiendo nuestra propuesta publicada en [3], tanto emisor como receptor deben tener el control para transferir una CD. Específicamente, el emisor va a poder cancelar la transferencia de la CD, mientras el receptor no haya decidido si aceptar o rechazar la transferencia del token. Estas funcionalidades van a estar predeterminadas por un deadline definido por el emisor de la CD.

Tal y como se ha mencionado anteriormente, debido a no haber ninguna implementación de SBT estándar, para implementar nuestro RejSBT se ha seguido el código estándar de implementación de NFTs (ERC-721), y lo hemos modificado únicamente manteniendo las funciones que no involucran o están relacionadas con la funcionalidad de transferencia de los NFTs, exceptuando las funciones `Mint` y `Burn`, ya que estas son esenciales para la creación y la destrucción de los SBTs. De acuerdo con el protocolo propuesto en [3], para los RejSBT también se ha implementado la funcionalidad de rechace.

```
// Mapping from token ID to transferable owner
mapping(uint256 => address) private
    _transferableOwners;
```

Listing 1. Mapping de propietarios transferibles

Para poder permitir el repudio de la recepción de un nuevo SBT, por parte de *R*, es necesario modificar la función `mint()` actual. Con esta modificación se elimina la ejecución directa de la transferencia del SBT y se pospone hasta que *R* ejecuta la aceptación de la recepción. Por esta razón, se introduce un nuevo mapping llamado `_transferableOwners`, presentado en el Listing 1. Este mapping es responsable de almacenar el propietario a quien *S* desea transferir el nuevo token. Utilizando este mapping, durante el proceso de minteo sólo se establece la transferencia de propiedad deseada sin transferir directamente el token. La transferencia de propiedad ocurrirá una vez *R* acepte la solicitud de transferencia.

```
function mint(
    address to,
    uint256 deadline,
    string memory credentials_,
    string memory terms_,
    uint256 expiry_
) public returns (uint256) {
    require(
        (keccak256(abi.encodePacked((
            credentials_))) !=
            keccak256(abi.encodePacked(("")))),
        "CredentialsRejectableSBT: credentials
            and terms are empty"
    );

    require(
        expiry_ != 0 ? deadline < expiry_ :
            deadline > expirity,
        "CredentialsRejectableSBT: incorrect
            expirity date value"
    );

    uint256 tokenId = _tokenIdCounter.current()
        ;
    _tokenIdCounter.increment();
    _mint(to, tokenId, deadline);

    credentialsData[tokenId] = CredentialsData
        ({
        credentials: credentials_,
        terms: terms_,
        expirity: expiry_
    });

    return tokenId;
}
```

Listing 2. Función pública para la emisión de CDs.

Los Listings 2 y 3 presentan la implementación de la funcionalidad de emisión de CDs. Dicha implementación consiste en una función pública (Listing 2) y una función privada (Listing 3). La función privada es heredada del smart contract que define los RejSBT genéricos con un deadline, el cual restringe el período de tiempo en que se permite la aceptación, el rechace o la cancelación de la emisión del SBT.

Con el objetivo de ejecutar la emisión de la CD, el emisor *S* ejecuta la función pública `mint()` (Listing 2). Esta función realiza algunas verificaciones para asegurar la correcta definición de la CD. Específicamente, verifica que se proporcionen credenciales y/o términos, y valida la fecha de caducidad ($expiry$) y la fecha límite ($deadline$).

Adicionalmente, si la CD incluye una fecha de caducidad, esta debe ser posterior a la fecha límite definida. Dicha condición es crucial para prevenir la aceptación de CDs expiradas. Por lo contrario, si la CD no necesita una fecha de caducidad, el valor del parámetro $expiry$ debe definirse a 0. Aplicando estas validaciones, el proceso de emisión asegura que la CD sea definida correctamente, y mantiene la integridad y la validez del protocolo.

Dentro de la función pública $mint$, se llama a la función

[1]https://github.com/secomuib/credentialssoulboundtokenmain

```
function _mint(
    address to,
    uint256 tokenId,
    uint256 deadline
) internal virtual {
    require(
        to != address(0),
        "RejectableSBTDeadline: mint to the
            zero address"
    );
    require(
        !_exists(tokenId),
        "RejectableSBTDeadline: token already
            minted"
    );
    require(
        deadline > block.timestamp,
        "RejectableSBTDeadline: deadline
            expired"
    );

    _minters[tokenId] = _msgSender();
    _transferableOwners[tokenId] = to;
    _deadlines[tokenId] = deadline;
    _states[tokenId] = State.Minted;

    emit TransferRequest(_msgSender(), to,
        tokenId);
}
```

Listing 3. Función privada para la emisión de CDs.

privada *mint*, y finalmente, los metadatos de la CD se rellenan. Los metadatos incluyen información sobre las credenciales, los términos y la fecha de caducidad definida.

La función *mint* privada es responsable de realizar las siguientes tres verificaciones:

1) La dirección del receptor (*to*), representando la dirección de *R*, no debe ser la dirección zero. Asegurando que se proporciona un receptor válido para la emisión de la CD.
2) El *tokenId* dado, no puede haber sido emitido previamente. Garantizando la no reutilización del identificador del token (*tokenId*) en múltiples CDs, asegurando la unicidad y la prevención de conflictos.
3) El valor asignado al *deadline* debe ser posterior a la fecha y hora actual. Asegurando que el *deadline* se encuentre en el futuro, permitiendo el tiempo suficiente para la aceptación, el repudio o la cancelación de la emisión de la CD.

Una vez ejecutadas exitosamente estas verificaciones, los mappings utilizados se rellenan:

- $\_minters$ contiene la dirección del emisor de la CD, (*S*).
- $\_transferableOwners$ almacena la dirección del receptor de la CD, (*R*).
- $\_deadlines$ contiene la fecha límite (*deadline*) de la CD.
- $\_states$ almacena el valor del estado de la CD, que ha sido emitida.

Es importante tener en cuenta que la ejecución de la función `_mint()` no activa la transferencia del RejSBT. La transferencia se lleva a cabo en la ejecución de la función `acceptTransfer()`, tal como se presenta en el Listing 4.

```
function acceptTransfer(uint256 tokenId) public
    virtual override {
        require(
            _transferableOwners[tokenId] ==
                _msgSender(),
            "RejectableSBTDeadline: accept
                transfer caller is not the
                receiver of the token"
        );
        require(
            _deadlines[tokenId] > block.
                timestamp,
            "RejectableSBTDeadline: deadline
                expired"
        );
        require(
            _states[tokenId] == State.Minted,
            "RejectableSBTDeadline: token is
                not in minted state"
        );

        address from = minterOf(tokenId);
        address to = _msgSender();

        _balances[to] += 1;
        _owners[tokenId] = to;
        _states[tokenId] = State.Accepted;
        // remove the transferable owner from
            the mapping
        _transferableOwners[tokenId] = address
            (0);

        emit AcceptTransfer(from, to, tokenId);
    }
```

Listing 4. Función para aceptar la transferencia de CDs.

Para la aceptación de la transferencia de la CD, hemos introducido una nueva función llamada `acceptTransfer()` (Listing 4), y para el rechazo hemos introducido una función llamada `rejectTransfer()` (Listing 5). Ambas funciones comparten las mismas verificaciones:

1) La dirección de *R* debe encontrarse en el mapping `_transferableOwners`. Asegurando que *R* ha sido propuesto como propietario transferible del $tokenId$ dado.
2) El momento de ejecución debe ser previo al *deadline* definido. Asegurando que la aceptación o el rechace de la transferencia se ejecute dentro del período de tiempo permitido.
3) El $tokenId$ dado no debe haber sido previamente utilizado para la emisión de otra CD. Asegurando que la aceptación o el rechace aplique a una CD valida.

Implementando estas verificaciones, ambas funciones `acceptTransfer()` y `rejectTransfer()` aseguran que se cumplen las condiciones apropiadas para la aceptación o el rechace de la CD.

Una vez verificado el cumplimiento de estas condiciones, la función `acceptTransfer()` procede con la transferencia del $tokenId$ al nuevo propietario cambiando el atributo de propiedad a la nueva dirección, y eliminando los datos almacenados en el mapping `_transferableOwners` del $tokenId$.

Por otra parte, la función `_rejectTransfer()` únicamente elimina los datos almacenados en el mapping `_transferableOwners` sobre el *tokenId* dado. En este caso, no hay transferencia de propiedad, ya que la transferencia se rechaza.

```
function rejectTransfer(uint256 tokenId) public
     virtual override {
    require(
        _transferableOwners[tokenId] ==
            _msgSender(),
        "RejectableSBTDeadline: reject transfer
             caller is not the receiver of the
             token"
    );
    require(
        _deadlines[tokenId] > block.timestamp,
        "RejectableSBTDeadline: deadline
            expired"
    );
    require(
        _states[tokenId] == State.Minted,
        "IBERejectableSBT: token is not in
            minted state"
    );

    address from = minterOf(tokenId);
    address to = _msgSender();

    _states[tokenId] = State.Rejected;
    _transferableOwners[tokenId] = address(0);

    emit RejectTransfer(from, to, tokenId);
}
```

Listing 5. Función para rechazar la transferencia de CDs.

Finalmente, el Listing 6 proporciona la implementación para permitir a *S* cancelar la propuesta de emisión de la CD abierta, siempre que *R* no haya ejecutado ninguna de las funciones `acceptTransfer()` o `rejectTransfer()`. En estos casos, *S* puede ejecutar la función `cancelTransfer()`. Esta función incluye las siguientes verificaciones:

1) El actor que ejecuta esta función debe ser *S*, el usuario que inició la emisión de la CD. Asegurando que el emisor se corresponde con el mismo usuario que ejecuta la cancelación de esta.
2) El momento de ejecución debe ser anterior al *deadline* definido. Asegurando que la cancelación se ejecuta dentro del período de tiempo permitido.
3) El *tokenId* dado no puede haber sido emitido. Asegurando que la cancelación se realiza sobre una CD no emitida.
4) El *tokenId* dado debe tener algún propietario transferible (`_transferableOwners`) propuesto. Asegurando que hay una propuesta de emisión del *tokenId* abierta.

En caso que las verificaciones se ejecuten exitosamente, la función `cancelTransfer()` elimina los datos del mapping `_transferableOwners` asociados al *tokenId*. Entonces, la implementación presentada en esta Sección consigue las funcionalidades principales definidas en el protocolo.

```
function cancelTransfer(uint256 tokenId) public
     virtual override {
    require(
        minterOf(tokenId) == _msgSender(),
        "RejectableSBTDeadline: cancel transfer
             caller is not the minter of the
            token"
    );
    require(
        _deadlines[tokenId] > block.timestamp,
        "RejectableSBTDeadline: deadline
            expired"
    );
    require(
        _states[tokenId] == State.Minted,
        "IBERejectableSBT: token is not in
            minted state"
    );

    address from = minterOf(tokenId);
    address to = _transferableOwners[tokenId];

    require(
        to != address(0),
        "RejectableSBTDeadline: token is not
            transferable"
    );

    _states[tokenId] = State.Cancelled;
    _transferableOwners[tokenId] = address(0);

    emit CancelTransfer(from, to, tokenId);
}
```

Listing 6. Fucnión para cancelar la transerencia de CDs.

VI. ANÁLISIS DE COSTES

Dada la implementación del protocolo, hemos testeado su rendimiento mediante el uso del entorno de desarrollo Hardhat[2]. Hardhat facilita el cálculo de uso de gas de la implementación, proporcionando métricas de los despliegues y las funciones ejecutadas en tests unitarios.

Para asegurar la precisión del protocolo, ejecutamos varios tests unitarios y evaluamos la eficiencia del protocolo en coste de gas. Específicamente, configuramos Hardhat para ejecutar en local un fork de la red blockchain Polygon. La Figura 2 presenta el informe obtenido del entorno de desarrollo Hardhat. Es importante tener en cuenta, tal como se presenta en la Figura 2, durante la ejecución de los tests (Mayo 2023), el precio del token nativo MATIC era de 0.87 USD, y el precio de gas correspondiente era de 188 gwei.

Los datos proporcionados por Hardhat demuestran que hemos ejecutado exitosamente todas las funciones implementadas para el protocolo de emisión y gestión de credenciales digitales de acceso.

Analizando los resultados del coste de gas, el despliegue de los smart contracts implementados, se corresponde con la transacción más cara que requiere el protocolo, con un coste de 0.25 USD. Todas las otras transacciones, comparadas con esta son muy baratas, con un precio máximo de 0.03 USD para la función `mint()`, que emite la nueva CD. La función `acceptTransfer()`, por el

[2]https://hardhat.org/

Solc version: 0.8.7	Optimizer enabled: true	Runs: 200	Block limit: 30000000 gas
Methods	188 gwei/gas		0.87 usd/matic

Contract	Method	Min	Max	Avg	# calls	usd (avg)
CredentualsRejectableSBT	acceptTransfer	–	–	95206	1	0.02
CredentualsRejectableSBT	cancelTransfer	–	–	50888	1	0.01
CredentualsRejectableSBT	mint	168211	185311	172486	8	0.03
CredentualsRejectableSBT	rejectTransfer	–	–	50688	1	0.01
Deployments					% of limit	
CredentualsRejectableSBT		–	–	1558987	5.2 %	0.25

Fig. 2. Informe de costes emitido por Hardhat.

hecho de involucrar el cambio de propietario, es ligeramente más cara que las funciones `cancelTransfer()` y `rejectTransfer()`, que únicamente eliminan el propietario transferible propuesto.

A partir de los resultados obtenidos, comprobamos como los costes asociados al sistema de emisión de credenciales digitales utilizando SBT rechazables seria muy reducido comparado con sistemas actuales.

VII. CONCLUSIONES

Este artículo presenta una propuesta innovadora que combina el uso de tokens digitales con sistemas de gestión de credenciales. El resultado es una solución poderosa para la gestión de credenciales y que al mismo momento introduce la aceptación de términos y condiciones relacionados con el uso de estas credenciales. Además, genera evidencias de no repudio en ambos casos, la recepción de las credenciales y la aceptación de los términos.

Debido a que la blockchain ofrece diversas propiedades inherentes que son altamente ventajosas para la gestión de credenciales, los tokens son un elemento interesante para representar credenciales. Pero los estándares existentes de tokens no cumplen las propiedades deseadas para poder ser utilizados adecuadamente para representar credenciales digitales, como es la necesidad de no-transferibilidad juntamente con el rechazo. Hemos presentado un nuevo tipo de token, el RejSBT, para representar las credenciales y los términos asociados, proporcionando las pruebas de no repudio de recepción y origen, consiguiendo las propiedades previamente citadas de no-transferibilidad y rechazo.

El protocolo propuesto ha sido implementado y evaluado para comprobar su viabilidad, demostrando que es un sistema interesante para la gestión de credenciales, proporcionando múltiples facilidades para la creación y la autogestión de las credenciales por parte de los propios usuarios sin la necesidad de intervención de ninguna tercera parte, además de conseguir un coste muy reducido coste a diferencia de sistemas actuales de credenciales como podría ser la emisión de documentos de identidad.

AGRADECIMIENTOS

Esta publicación es parte del proyecto de I+D+i PID2021-122394OB-I00 (Blobsec), financiado por MCIN/ AEI/10.13039/501100011033/ y por ”FEDER Una manera de hacer Europa”.

REFERENCIAS

[1] Brands, S.: A Technical Overview of Digital Credentials. (2002)
[2] Weyl, Eric Glen and Ohlhaver, Puja and Buterin, Vitalik, Decentralized Society: Finding Web3’s Soul (May 10, 2022). Available at SSRN: https://ssrn.com/abstract=4105763 or http://dx.doi.org/10.2139/ssrn.4105763
[3] M. À. Cabot-Nadal, M. Magdalena Payeras-Capellà, M. Mut-Puigserver and A. Soto-Fernández, ”Improving the Token ERC-721 Implementation for Selective Receipt: Rejectable NFTs,” 2022 6th International Conference on System Reliability and Safety (ICSRS), Venice, Italy, 2022, pp. 243-250, doi: 10.1109/ICSRS56243.2022.10067494.
[4] Stefan Brands. Privacy-protected transfer of electronic information. U.S. Patent ser. no. 5,604,805, February 1997. Filed August 1993
[5] Grassi, P. A., Garcia, M. E., & Fenton, J. L. (2017). Digital Identity Guidelines. NIST special publication, 800, 63-3.
[6] Allen, C. (2016). The path to self-sovereign identity. Life with Alacrity.
[7] R. Mecozzi, G. Perrone, D. Anelli, N. Saitto, E. Paggi and D. Mancini, ”Blockchain-related identity and access management challenges: (de)centralized digital identities regulation,” 2022 IEEE International Conference on Blockchain (Blockchain), Espoo, Finland, 2022, pp. 443-448, doi: 10.1109/Blockchain55522.2022.00068.
[8] P. Herbke and H. Yildiz, ”ELMO2EDS: Transforming Educational Credentials into Self-Sovereign Identity Paradigm,” 2022 20th International Conference on Information Technology Based Higher Education and Training (ITHET), Antalya, Turkey, 2022, pp. 1-7, doi: 10.1109/ITHET56107.2022.10031276.
[9] M. S. Reza, S. Biswas, A. Alghamdi, M. Alrizq, A. K. Bairagi and M. Masud, ”ACC: Blockchain Based Trusted Management of Academic Credentials,” 2021 IEEE International Symposium on Smart Electronic Systems (iSES), Jaipur, India, 2021, pp. 438-443, doi: 10.1109/iSES52644.2021.00104.
[10] M. Eltuhami, M. Abdullah and B. A. Talip, ”Identity Verification and Document Traceability in Digital Identity Systems using Non-Transferable Non-Fungible Tokens,” 2022 International Visualization, Informatics and Technology Conference (IVIT), Kuala Lumpur, Malaysia, 2022, pp. 136-142, doi: 10.1109/IVIT55443.2022.10033362.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Modelado de un gemelo digital para la optimización de un sistema de auto-abastecimiento energético de uso residencial

Laura Rodríguez de Lope*, Víctor M. Maestre†, Luis Diez*, Alfredo Ortiz†, Ramón Agüero*, Inmaculada Ortiz†
*Departamento de Ingeniería de Comunicaciones
†Departamento de Ingeniería Química y Biomolecular
Universidad de Cantabria
*{laura.rdelope, luisfrancisco.diez, ramon.agueroc}@unican.es
†{victormanuel.maestre, alfredo.ortizsainz, inmaculada.ortiz}@unican.es

La preocupante situación climática y la crisis energética, junto con la actual inestabilidad política, ha impulsado en Europa una serie de políticas que favorezcan la instalación de fuentes de energía renovables. Para combatir la intermitencia y las fluctuaciones asociadas a su funcionamiento, el hidrógeno renovable se presenta como una solución de interés para descarbonizar diferentes sectores económicos. Así, el diseño e implementación de un sistema híbrido de energía renovable-hidrógeno ha dado como resultado la primera vivienda social eléctricamente autosuficiente de España, situada en la localidad de Novales (Cantabria). Por otro lado, la digitalización de este tipo de sistemas permitiría una adaptación automática a situaciones cambiantes, incrementando la eficiencia energética. En este contexto, el proyecto HY2RES propone una arquitectura de gemelo digital que, utilizando técnicas de aprendizaje automático e inteligencia artificial, facilite la optimización del rendimiento del sistema físico, mediante la actuación sobre sus elementos de control. Para ello se plantea el uso de soluciones de telemetría que permitan la captación y almacenamiento de datos del propio sistema físico y del entorno (por ejemplo meteorológicos), cuando sea necesario. Este trabajo muestra algunos resultados iniciales del gemelo digital propuesto, que incorpora modelos de los componentes eléctricos del sistema físico.

***Palabras Clave*—Gemelo digital, energías renovables, hidrógeno.**

I. INTRODUCCIÓN

La situación actual de crisis climática y energética ha derivado en políticas que fomentan el uso de fuentes de energía renovable (FER) que favorezcan la independencia energética a través de soluciones sostenibles.

Desde la Conferencia de las Partes (COP) 21, celebrada en París en 2015 [1], se han promovido una serie de hojas de ruta y estrategias para limitar los efectos nocivos causados por el cambio climático. Así, la dependencia de la sociedad actual de los combustibles fósiles es el principal factor responsable de la situación climática mundial. En concreto, las actividades relacionadas con la energía contribuyen a más de tres cuartas partes de las emisiones totales de dióxido de carbono equivalente (CO2eq) [2], [3]. En este contexto, la Unión Europea ha aprobado el plan "Fit for 55", que incluye limitar las emisiones de gases de efecto invernadero (GEI) en un 55% para 2030.

Por otro lado, la inestabilidad política actual, derivada principalmente de la guerra entre Ucrania y Rusia, ha provocado una alta inflación y la escasez de combustibles fósiles (principalmente gas natural y petróleo) importados de Rusia [4].

Ante esta situación, el despliegue a gran escala de fuentes de energía renovables se hace esencial para garantizar un sistema energético descarbonizado, que además proporcione cierta independencia energética a través de soluciones eficientes y sostenibles. Sin embargo, es fundamental encontrar soluciones tecnológicas eficientes para el almacenamiento de energía, que respondan de forma rápida, segura y flexible al comportamiento intermitente y fluctuante de las FER. Por este motivo, la Comisión Europea aprobó el plan REPowerEU para 2022; con él, la Unión Europea (UE) pretende promover la independencia energética de Europa a través de las FER, el aumento de la eficiencia energética, y la economía del hidrógeno. Así, el uso del hidrógeno como vector energético y materia prima es una solución eficiente y sostenible para el almacenamiento de energía a gran escala y estacional. Aparece como una alternativa idónea para impulsar la presencia de las FER en el ámbito energético, y favorecer además la descarbonización de diferentes sectores relacionados con la energía [5].

En este contexto, destaca el sector residencial, al ser un consumidor masivo de energía en la UE, contribuyendo al 40% del consumo final. Además, es un sector altamente ineficiente, debido a su envejecimiento, lo que contribuye negativamente a la huella de carbono causada por este sector. Por otra parte, la inflación sin precedentes de la economía ha agravado la situación de los ciudadanos más vulnerables, llevándolos en muchos casos a la pobreza energética. Por ello, es fundamental mejorar el rendimiento energético del sector residencial para así disminuir su contribución al cambio climático y abaratar las facturas eléctricas que están deteriorando los estándares de calidad de la población [6].

El proyecto europeo SUDOE ENERGY PUSH[1] [7] dio como resultado el un sistema híbrido de energía renovable-hidrógeno instalado en la localidad de Novales (Cantabria), siendo la primera vivienda social eléctricamente autosuficiente de España.

Para optimizar el funcionamiento del sistema basado en hidrógeno renovable (SHR), se ha propuesto el desarrollo de un gemelo digital (GD) de la planta piloto, en el marco del proyecto HY2RES, que permita mejorar del rendimiento del SHR mediante algoritmos desarrollados y validados sobre la réplica digital. Este trabajo presenta la arquitectura del GD, compuesta por el modelo digital de la planta piloto, el interfaz de comunicaciones para la recogida de datos y el envío de señales de control, y un módulo para favorecer la compartición de datos con terceros.

El artículo se estructura como sigue: la Sección II describe el estado del arte en lo que se refiere a la utilización de gemelos digitales en el sector de la energía. La Sección III describe la arquitectura, tanto del sistema físico como del gemelo digital que se ha diseñado. Posteriormente, la Sección IV se centra en el desarrollo de los modelos que conforman el GD, para mostrar, en la Sección V, los resultados obtenidos en las primeras implementaciones que se han integrado. Por último, la Sección VI concluye el documento, resumiendo el trabajo e indicando los aspectos en los que se profundizará en el futuro.

II. ESTADO DEL ARTE

El concepto de GD se propuso originalmente a principios de siglo [8] para entornos industriales, aunque más recientemente su uso se ha extendido a diferentes sectores [9], [10], aprovechando los avances en la digitalización y las mayores capacidades de los sistemas de comunicación y computación. Como se explica en [11], la creciente complejidad en los procesos solo puede ser replicada a través de datos masivos, que adquieren una utilidad evidente cuando se aplican en los GD.

En el caso del sector químico, un escenario de especial interés es el de las fuentes de energía renovables. En este caso, tal y como manifiestan los autores en [12] y [10], existen pocos trabajos donde se haya tratado de aplicar el concepto de GD a este tipo de sistemas. Es más, en [13] se confirma que no existen estudios profundos acerca del uso de los GD en este sector.

[1]https://www.sudoe-energypush.eu/

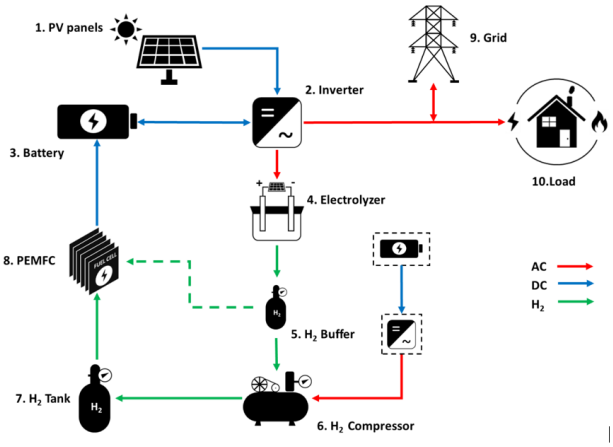


Fig. 1. Arquitectura del Sistema basado en Hidrógeno Renovable

Uno de los pocos trabajos en los que se aplica el concepto de gemelo digital al ámbito energético es el de Nguyen *et al.* [14]. En él los autores proponen el uso de un GD para mejorar el comportamiento de sistemas de distribución de energía, destacando su capacidad para tomar decisiones óptimas de control, basadas en análisis que se llevan a cabo en tiempo real. De forma similar, en [15] Agostinelli *et al.* analizan los posibles beneficios del GD en la gestión de la distribución y el consumo de energía en edificios, destacando el papel que las técnicas de inteligencia artificial podrían jugar.

Centrándonos en el proceso de producción de hidrógeno, en [16] se propone el uso de un GD para abordar las incertidumbres asociadas a los costes de inversión y explotación del sistema, analizando la influencia de distintos factores en los indicadores financieros.

Como se puede observar, a pesar de los potenciales beneficios que pueden presentar, el uso de gemelos digitales en el ámbito energético en general, y en el de sistemas basados en hidrógeno renovable en particular, es en la actualidad muy limitado. Es precisamente en este punto en el que se sitúa la principal contribución del proyecto HY2RES, tal y como se describe en este trabajo.

III. METODOLOGÍA

En esta sección se describe el sistema de hidrógeno renovable instalado (sistema físico), indicando los elementos que lo constituyen y el funcionamiento general del mismo. Posteriormente se describe el diseño global del gemelo digital, describiendo su funcionamiento.

A. Planta Piloto

El SHR diseñado y desplegado en el marco de la propuesta SUDOE ENERGY PUSH combina tanto energías renovables como novedosas tecnologías basadas en el hidrógeno para garantizar la completa autosuficiencia eléctrica de una vivienda social a lo largo del año. La Fig. 1 presenta un esquema de la planta piloto desplegada, así como los flujos eléctricos y de hidrógeno dentro del sistema.

Así, los paneles fotovoltaicos (punto 1 de la Fig. 1) instalados en el tejado del edificio recogen la energía solar para abastecer a la vivienda (punto 10 de la figura) como fuente primaria. Si existe un excedente de energía después de abastecer la vivienda, ésta se acumula en primer lugar en un conjunto de baterías de iones de litio (punto 3), que almacenan energía para el consumo a corto plazo. En caso de que el excedente sea alto, éste se emplea para la generación, compresión y almacenamiento del hidrógeno, para el ahorro energético estacional. El hidrógeno se genera a través de un electrolizador (punto 4), que genera hidrógeno por electrolisis, por lo que está alimentado con energía eléctrica. Este hidrógeno es almacenado, en primer lugar, en un buffer (punto 5). Al llenarse el buffer, el hidrógeno generado se comprime para ser almacenado en un tanque de alta presión (puntos 6 y 7). En el caso de que siguiera habiendo excedente de energía eléctrica procedente de los paneles solares después de abastecer estos procesos, ésta es vertida a la red (punto 9). Durante los periodos de déficit de energía fotovoltaica, las baterías suministran electricidad a la vivienda y, cuando alcanzan un determinado umbral de descarga, se cargan mediante una pila de combustible (punto 8) que, finalmente, cubre la demanda del hogar. Esta pila de combustible genera energía eléctrica a partir del hidrógeno almacenado en el buffer o el tanque de alta presión.

El funcionamiento de la planta piloto se ha automatizado por completo y se controla remotamente gracias a la ayuda de un controlador lógico programable (PLC). Además, el SHR funciona con una estrategia de gestión energética basada en el estado de la autonomía almacenada, y se monitoriza a través de un sistema de control y adquisición de datos (SCADA).

B. Arquitectura del gemelo digital

Sobre la planta piloto descrita, en el proyecto HY2RES se propone el diseño y desarrollo de un GD en el que se caractericen los componentes de la planta y se desarrollen soluciones de control para afrontar la mejora automática de los parámetros de control. Para el modelado de los componentes del sistema se aplicarán técnicas de aprendizaje automático (ML) e inteligencia artificial (IA) a los datos recopilados cuando el comportamiento de la algorítmica programada no se adapte al funcionamiento real del SHR.

Para desarrollar el GD, el proyecto HY2RES propone una arquitectura de tres etapas principales. La primera, centrada en la interacción físico-virtual, se encarga de la recopilación de información del sistema real, así como de la aplicación de las políticas de decisión. Es más, también se contempla un módulo para la gestión e integración de datos procedentes de fuentes externas, como puede ser la previsión meteorológica o el precio de la energía. En una segunda etapa, y aprovechando los módulos mencionados, se implementa el GD, mediante un conjunto de librerías de software que replican el comportamiento del sistema real. En este sentido, una vez identificadas las variables de entrada/salida y de control de los principales componentes del SHR, se aborda su modelado. Este se basa en soluciones algorítmicas, cuando el comportamiento subyacente es bien conocido, y técnicas de ML en caso contrario. Finalmente, el proyecto HY2RES utilizará el GD para evaluar el rendimiento de diversas políticas de control sobre la réplica digital, incluidas las basadas en la previsión meteorológica.

La Fig. 2 representa la arquitectura completa del GD a alto nivel, incluyendo el flujo lógico básico de la operación del GD. Como se ha mencionado, la planta piloto utiliza un sistema SCADA para la supervisión del funcionamiento de los dispositivos. Así, el GD interactuará con dicho sistema a través del PLC, para recopilar datos y aplicar las acciones de control adecuadas (puntos 1 y 4 de la figura).

El sistema digital está formado por un componente principal, que es el modelo del GD, entrenado por los datos recopilados del sistema físico a través del sistema de control, que captura la lógica y funcionamiento del piloto. El GD seguirá un esquema basado en bucles para garantizar que se refleja de manera precisa el comportamiento del sistema real (puntos 5 y 6 de la Fig. 2): (i) análisis de las estrategias de control sobre el modelo GD para optimizar el rendimiento del sistema físico; (ii) aplicación de la estrategia sobre el piloto real, mediante el envío de comandos de control que interactúan con el sistema SCADA desplegado; (iii) el sistema seguirá recibiendo, a su vez, retroalimentación del piloto físico (monitorización continua) para seguir entrenando los modelos del GD en los casos en que se adopten soluciones de ML.

Por último, aunque se pretende que el GD funcione de forma cerrada, los datos generados en su operación, que se consideren más relevantes, serán accesibles para ser explotados por terceros, tal y como se muestra en la Fig. 2. Así, los datos generados por el GD se pondrán a disposición en un mercado de datos FIWARE y en repositorios de acceso abierto como Zenodo. En este sentido, es necesario adoptar modelos de datos abiertos, como es el caso de los Smart Data Models (SDM), que faciliten la interoperabilidad y reutilización de la información por parte de terceros. En caso de no existir modelos para las necesidades específicas del GD, se propondrán nuevas definiciones para ampliar el repositorio de los SDM, facilitando que los conjuntos de datos generados sean abiertos y accesibles, incluso en tiempo real.

Una vez expuesta la arquitectura general del sistema, las siguientes secciones presentan el modelo de GD y los primeros resultados obtenidos.

IV. MODELADO DEL GEMELO DIGITAL

El GD se implementa como un conjunto de módulos software independientes e interconectados entre sí, cada uno de los cuales modela uno o varios componentes físicos del SHR. Esta solución desagregada facilita la implementación, y la validación independiente, de cada uno de dichos elementos. Por otro lado, permite reemplazar los modelos aplicados a diferentes componentes concretos, sin afectar al sistema en su conjunto.

Para cada uno de los módulos se ha identificado el conjunto de variables que influye en su operación, catalogadas como de control y de entrada, así como las

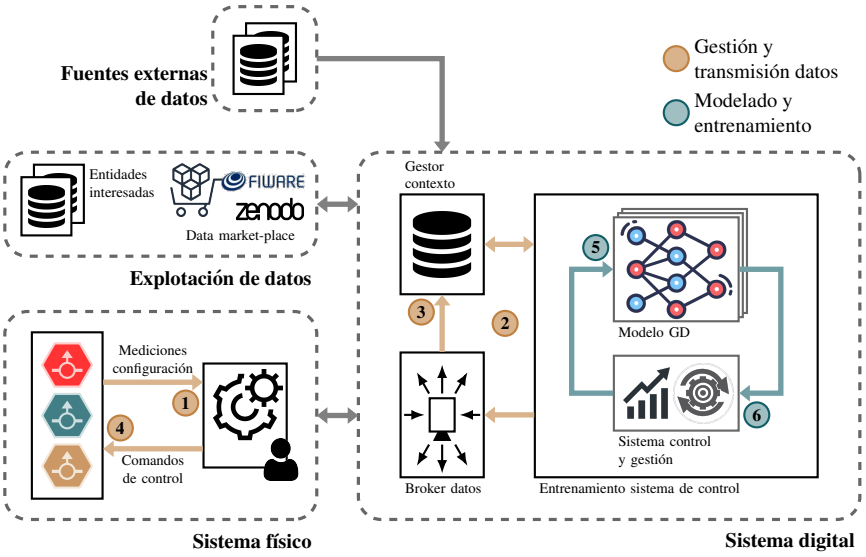


Fig. 2. Arquitectura a alto nivel del gemelo digital

salidas que genera. Las variables de entrada y salida corresponden a magnitudes físicas del sistema, mientras que las de control imitan señales de decisión, como es el caso de las generadas por el PLC. De este modo, al aplicar unas determinadas variables de entrada y de control sobre un módulo, éste genera las correspondientes salidas, replicando el comportamiento de su homólogo físico. Los distintos módulos están conectados de tal manera que las salidas de un módulo pueden actuar como variables de entrada para otros. De hecho, algunas señales realimentan al sistema, como es el caso del estado de carga (State of Charge - SOC) de la batería y las presiones del bloque de almacenamiento de hidrógeno, que realimentan al PLC para la toma de decisiones.

En la Fig. 3 se muestran los módulos identificados, junto con las variables consideradas, diferenciando entre variables externas (color gris), del sistema (color azul) y de control (color naranja). Las primeras son independientes del sistema, mientras que las segundas son modeladas por el GD, y su estado depende de las variables externas, el estado previo del sistema y/o la política de decisiones programada. Como puede observarse en la figura, el GD únicamente requiere de la potencia demandada por la vivienda y la potencia fotovoltaica generada como variables externas, mientras que el resto de variables son estimadas por el propio GD. En próximas fases de la implementación se integrarán otras fuentes de datos externas, como la previsión meteorológica o la tarificación eléctrica, que podrían tener una influencia directa en la optimización del rendimiento del sistema.

A pesar de que en la fase inicial de HY2RES la actividad se ha centrado en la implementación de los módulos de los componentes eléctricos, como son el PLC, el inversor y la batería, a continuación de describen todos los módulos identificados. Como paso inicial, se implementan mediante algoritmos que replican el comportamiento lógico de los componentes físicos. Posteriormente, en los casos en los que soluciones algorítmicas no sean precisas, se adoptarán modelos basados en ML, entrenados a partir de los datos recogidos. A modo de resumen, en la Tabla I se indican todos los parámetros de entrada y salida de cada componente.

A. PLC

Se trata del elemento principal, ya que integra la lógica de gestión del sistema. Este módulo toma como variables de entrada la demanda de potencia de la vivienda y la generación de energía fotovoltaica de los paneles solares. Estas variables, junto con otras internas que determinan el estado del sistema, como es el caso del estado de carga de la batería (SOC) y la presión de los sistemas de almacenamiento de hidrógeno, son utilizadas para estimar las variables de salida, que en este caso se corresponden con las señales de control para el encendido/apagado del resto de módulos. Inicialmente la lógica implementada en el PLC sigue el comportamiento actualmente configurado, tal como está descrito en la Sección III, para la validación del resto de modelos. Posteriormente, la implementación de diferentes algoritmos de control conllevará la modificación de esta lógica.

B. Inversor

El módulo inversor tiene, como variables de entrada, la potencia demandada por la vivienda, así como la que generan los paneles, además de la señal de carga y descarga de la batería, generada por el PLC. El modelo implementado distribuye el excedente/déficit de energía, generado a partir de la diferencia existente entre la demanda de la vivienda y la producción de los paneles solares, hacia/desde los diferentes módulos de almacenamiento de energía y, en su caso, hacia/desde la red eléctrica.

C. Batería

El módulo de batería utiliza la potencia de carga/descarga proporcionada por el módulo inversor para actualizar el SOC. En caso de que la pila de combustible

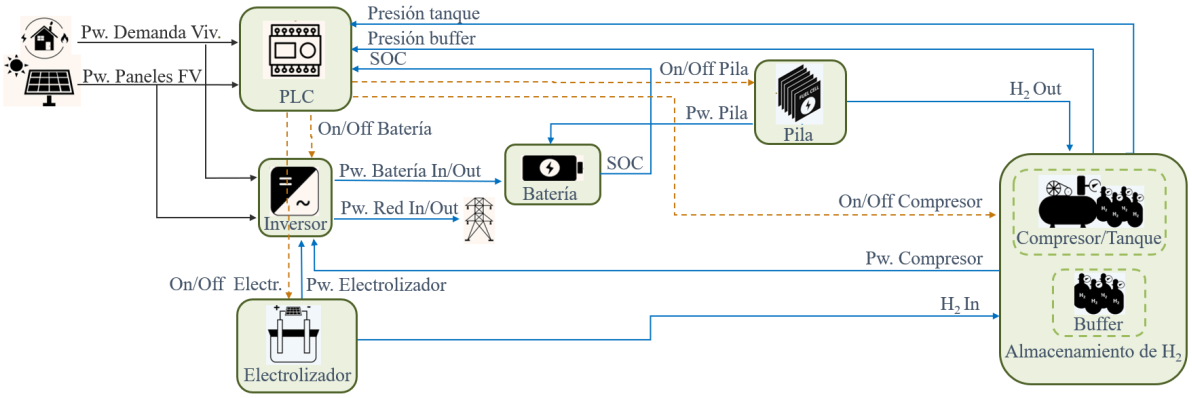


Fig. 3. Módulos y variables del gemelo digital

Tabla I
VARIABLES DE ENTRADA Y SALIDA DE CADA COMPONENTE

Variables de entrada	Variables de salida
PLC	
Pot. paneles solares	Señal on/off electrolizador
Pot. demandada	Señal on/off compresor
SOC	Señal on/off pila de combustible
Presión buffer y tanque	Señal carga/descarga batería
Inversor	
Pot. paneles solares	Pot. carga/desgarga batería
Pot. demandada vivienda	Pot. hacia/des la red
Pot. demandada electrolizador	
Pot. demandada compresor	
Señal carga/descarga batería	
Batería	
Pot. carga/descarga de batería	SOC
Pot. de la pila de combustible	
Electrolizador	
Señal on/off electrolizador	Pot. demandada electrolizador
	Hidrógeno generado
Almacenamiento de hidrógeno	
Señal on/off compresor	Pot. demandada compresor
Hidrógeno gen. electrolizador	Presión del buffer
Hidrógeno demandado pila	Presión del tanque
Pila de combustible	
Señal on/off pila	Pot. generada por la pila
Hidrógeno demandado pila	

esté encendida, la energía generada por ella también será utilizada como variable de entrada para determinar el SOC. La variable de salida de este módulo, SOC, realimenta al sistema, al ser utilizada por el PLC en su toma de decisiones.

D. Electrolizador

El módulo electrolizador modela, a partir de la señal de encendido, el flujo de hidrógeno generado y de potencia consumida por éste. Estas variables actuarán, además, como entrada de los módulos de almacenamiento de hidrógeno e inversor, respectivamente. Otras variables, como la temperatura ambiental, también serán consideradas en próximas fases, para modelar el comportamiento de este bloque.

E. Bloques de almacenamiento de hidrógeno

Este módulo está compuesto por un primer bloque de almacenamiento de hidrógeno en un buffer y un segundo bloque de hidrógeno comprimido, formado por el compresor y el tanque. La variable compresor on/off del PLC determina si el hidrógeno generado por el electrolizador debe ser almacenado en el buffer o en el tanque de hidrógeno comprimido. El flujo de entrada/salida de hidrógeno determinará la presión en estos bloques, que serán realimentadas al PLC, ya que se utilizan en la toma de decisiones de la lógica de control.

F. Pila de combustible

Recibe la señal de encendido por parte del PLC cuando la batería se encuentra por debajo de un umbral establecido. Proporciona como variables de salida el hidrógeno consumido y la potencia generada, que actualizan el estado de los módulos de almacenamiento de hidrógeno y batería, respectivamente.

V. RESULTADOS

Durante la fase inicial del proyecto HY2RES, la actividad se ha centrado en la implementación de los módulos de los componentes eléctricos del sistema: PLC, inversor y batería. El modelado ha seguido un enfoque algorítmico, que utiliza el funcionamiento lógico de cada uno de ellos, así como las ecuaciones que rigen su comportamiento. Posteriormente se adoptarán otros enfoques basados en datos, en función de la precisión de la predicción observada. Para validar que los modelos propuestos proporcionan un comportamiento adecuado, reflejando el de sus homólogos físicos, esta sección presenta un análisis de los resultados de los componentes mencionados, centrándose en la interacción entre el inversor y la batería.

Los resultados que se muestran a continuación se basan en los datos obtenidos tras el muestreo del sistema físico, midiendo el estado de cada una de las variables a lo largo de 7 días con una cadencia entre muestras de 5 segundos, lo que permite una granularidad suficiente como para capturar cambios bruscos en el estado de las variables como, por ejemplo, picos en la demanda de energía por parte de la vivienda. Para cada instante de muestreo, el GD genera una estimación de las variables correspondientes.

En el caso del inversor, el análisis compara los valores, reales y estimados por el modelo, de potencia suministrada/consumida a/desde la batería, siendo las señales

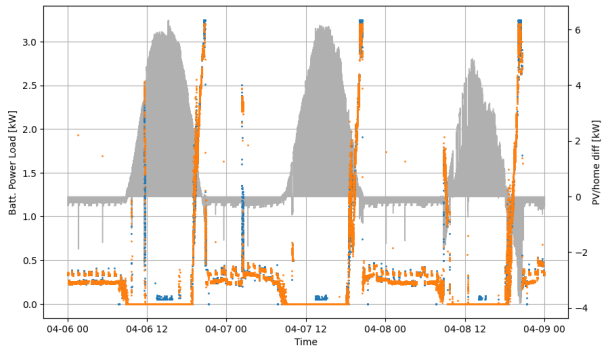


Fig. 4. Potencia de descarga instantánea de la batería. Real (azul) y estimada (naranja).

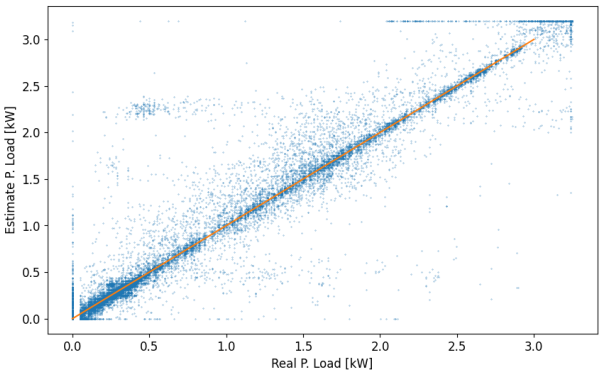


Fig. 5. Potencia de descarga de la batería. Real vs. Estimada

de entrada al modelo la potencia fotovoltaica generada, la demanda de potencia del hogar y las señales estimadas del PLC. En el caso de la batería, la variable de salida SOC se evalúa mediante su potencia de carga/descarga, que se utiliza como señal de entrada para el modelo.

La Fig. 4 muestra la evolución instantánea de la potencia de descarga de la batería (Batt. Power Load). Con fondo sombreado se representa la diferencia entre las dos variables de entrada, la alimentación fotovoltaica y la demanda doméstica, cuya escala se indica en el eje derecho de la figura (PV/home diff.). Los puntos azules y naranjas representan los valores reales y estimados de la descarga de la batería, respectivamente.

La figura muestra que la descarga de la batería se produce principalmente durante el período nocturno, cuando existe un déficit de energía (PV/home diff negativa). Durante el día, la potencia de descarga se mantiene nula, a excepción de ciertos picos de demanda. Se puede apreciar que el modelo sigue la misma tendencia que el comportamiento del sistema físico, proporcionando valores estimados muy próximos a los reales en la mayoría de los casos. Sin embargo, se aprecian diferencias cuando se producen descargas bruscas de potencia, que en la figura se muestran como picos. Esto se debe a que la red eléctrica es más reactiva que la batería, por lo que ante demandas bruscas de energía, es capaz de reaccionar con mayor rapidez. Este fenómeno hace que se produzca una importación de energía de la red, siendo necesaria una descarga de potencia de la batería menor. Este comportamiento del sistema real es impredecible algorítmicamente y, por lo tanto, no está contemplado por el modelo, que siempre realiza una descarga de batería máxima antes de importar potencia de la red.

La Fig. 5 muestra la relación entre los valores de potencia de descarga de la batería medidos y los estimados durante el periodo de muestreo. Cada uno de los puntos azules representa la potencia de descarga estimada por el GD, frente a la medida por el sistema real en un instante temporal. El comportamiento ideal se representa mediante una línea (naranja), donde la estimación y las muestras reales coinciden. Como puede observarse, el modelo es capaz de reflejar, de manera precisa, el comportamiento real en la mayoría de los casos. Sin embargo, en el caso de valores extremos, máximos o mínimos, se producen ciertos desajustes. En la parte izquierda de la figura se puede apreciar que el modelo proporciona valores en todo el rango, mientras que el sistema de medición no detectó ninguna descarga. Este comportamiento se debe a lagunas en el proceso de monitorización en el sistema físico, en concreto por parte del PLC, durante ciertos intervalos temporales, en los que se proporcionan medidas nulas. El fenómeno mencionado anteriormente, en el que se producen demandas bruscas de potencia, se refleja también en la parte superior derecha de la figura, donde se aprecia que la potencia de descarga estimada es máxima, mientras que la real adopta valores intermedios debido a la reacción más rápida de la red eléctrica ante estas demandas de consumo.

Para caracterizar numéricamente la desviación entre los valores proporcionados por el GD y los reales, se ha calculado el error cuadrático medio normalizado (NRMSE) a partir de las estimaciones de la potencia de descarga de la batería, obteniéndose un NRMSE de 0.0389, lo que evidencia la gran precisión del modelo propuesto. No obstante, en próximas fases, se adoptarán técnicas de aprendizaje máquina que permitan replicar la reacción del sistema ante demandas bruscas, tratando de reducir aún más esta desviación.

Para analizar el comportamiento del módulo de batería, se ha comparado la señal de salida SOC frente a su valor real, monitorizado en el sistema físico. La Fig. 6 muestra la evolución temporal del SOC de la batería real y el estimado por el GD, con líneas azules y naranjas, respectivamente. Adicionalmente, para ilustrar mejor el comportamiento observado, se representan las variables de entrada del modelo de batería, potencias de carga/descarga, mediante un fondo gris sombreado y escala en el eje derecho (In/Out power).

Como puede apreciarse, la señal de salida SOC estimada por el modelo del GD y la real siguen la misma tendencia. En un análisis más detallado, se puede observar que el modelo se comporta adecuadamente (con mucha precisión) en valores medios de SOC, mientras que en valores altos y bajos de SOC la diferencia se vuelve más apreciable. Esta circunstancia se debe al comportamiento no lineal de la batería cuando se encuentra en valores

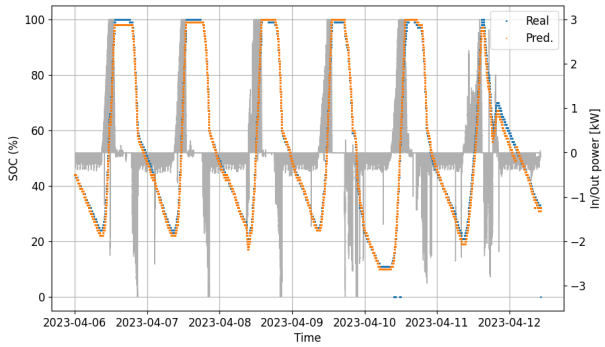


Fig. 6. SOC real y estimada

extremos de carga.

En este caso, el NRMSE obtenido de todo el periodo de medición $1.72e-2$. Como puede verse, a pesar de esas diferencias en valores extremos del SOC, el modelo propuesto en el GD presenta un comportamiento adecuado, capturando de manera bastante precisa el del componente correspondiente en el sistema físico.

VI. CONCLUSIONES

Este trabajo presenta el diseño y los primeros pasos de la implementación de un GD para un sistema híbrido de energía renovable-hidrógeno, instalado en una vivienda de la localidad de Novales (Cantabria).

El modelo GD, estructurado en módulos funcionales que emulan los componentes físicos, utiliza como variables externas la demanda energética de la vivienda y la energía generada por los paneles fotovoltaicos instalados, efectuando la predicción del valor del resto de variables presentes en el sistema.

Los resultados obtenidos para los primeros módulos incorporados al GD, que representan los componentes eléctricos del sistema físico, muestran un buen ajuste entre los valores monitorizados en el sistema real y los obtenidos a partir del GD, con errores, NRMSE, bajos. Sin embargo, también se observan comportamientos no previstos de los componentes físicos.

Los próximos pasos irán dirigidos a la ampliación de las funcionalidades del modelo GD, implementando el resto de componentes del sistema como módulos de software independientes, que se conectarán entre sí mediante las variables apropiadas. En paralelo, se analizarán alternativas basadas en ML que permitan reducir el NRMSE siempre que la algorítmica programada no se ajuste en la predicción a los valores reales, como es el caso de los valores extremos del SOC. Además, se integrarán los elementos adecuados para facilitar el acceso abierto a los datos más relevantes.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Gobierno de España (Ministerio de Ciencia e Innovación), y por la Unión Europea (Next GenerationEU/RTRP) a través de los proyectos *Gemelo digital de un sistema híbrido solar fotovoltaica-hidrógeno para el abastecimiento en el ámbito residencial* (TED2021-129951B-C22) y *Piloto demostrador de un sistema híbrido solar fotovoltaica-hidrógeno para el abastecimiento energético en el ámbito residencial* (TED2021-129951B-C21), 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 "Ayudas a proyectos de investigación con alto potencial industrial de agentes tecnológicos de excelencia para la competitividad industrial TCNIC".

REFERENCIAS

[1] "Paris agreement." [Online]. Available: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=_en

[2] International Energy Agency, "CO_2 emissions from fuel combustion 2020 edition. Database documentation," 2020. [Online]. Available: https://iea.blob.core.windows.net/assets/474cf91a-636b-4fde-b416-56064e0c7042/WorldCO2_Documentation.pdf

[3] H. Ritchie, M. Roser, and P. Rosado, "Co2 and greenhouse gas emissions," *Our World in Data*, 2020, https://ourworldindata.org/co2-and-greenhouse-gas-emissions.

[4] European Council, "Fit for 55 - the eu's plan for a green transition," 2023. [Online]. Available: https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/

[5] ——, "Repowereu: Affordable, secure and sustainable energy for europe," 2022. [Online]. Available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/repowereu-affordable-secure-and-sustainable-energy-europe_en

[6] V. Maestre, A. Ortiz, and I. Ortiz, "Transition to a low-carbon building stock. techno-economic and spatial optimization of renewables-hydrogen strategies in spain," *Journal of Energy Storage*, vol. 56, p. 105889, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X22018771

[7] V. M. Maestre, A. Ortiz, and I. Ortiz, "Implementation and digitalization of a renewable hydrogen-based power system for social housing decarbonization," *Chemical Engineering Transactions*, vol. 96, pp. 223–228, 2022. [Online]. Available: https://www.cetjournal.it/cet/22/96/038.pdf

[8] M. Grieves, "Digital twin: manufacturing excellence through virtual factory replication," *White paper*, vol. 1, no. 2014, pp. 1–7, 2014.

[9] 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.

[10] F. Tao, B. Xiao, Q. Qi, J. Cheng, and P. Ji, "Digital twin modeling," *Journal of Manufacturing Systems*, vol. 64, pp. 372–389, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612522001108

[11] 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.

[12] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," *Journal of Manufacturing Systems*, vol. 58, pp. 346–361, 2021, digital Twin towards Smart Manufacturing and Industry 4.0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612520301072

[13] A. Ebrahimi, "Challenges of developing a digital twin model of renewable energy generators," in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, 2019, pp. 1059–1066.

[14] S. Nguyen, M. Abdelhakim, and R. Kerestes, "Survey paper of digital twins and their integration into electric power systems," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, 2021, pp. 01–05.

[15] S. Agostinelli, F. Cumo, G. Guidi, and C. Tomazzoli, "Cyber-physical systems improving building energy management: Digital twin and artificial intelligence," *Energies*, vol. 14, no. 8, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/8/2338

[16] B. Gerard, E. Carrera, O. Bernard, and D. Lun1, "Smart design of green hydrogen facilities: A digital twin-driven approach," *E3S Web of Conferences*, vol. 334, no. 2022, pp. 1–7, 2022.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Custodia de fondos avanzada con Timelocks y Miniscript

Álvaro López Sánchez, Diego García Muñoz,
Juan Carlos Delgado de la Torre, Miguel Ángel Fortes Santiago, Antonio Tóvar

(Universidad de Málaga, Fortris)

En el ámbito de las criptomonedas la seguridad en la custodia de los activos es esencial. En el caso de Bitcoin, la complejidad técnica en la experiencia de usuario en la gestión y protección de activos han limitado su adopción. Este documento aborda dicho desafío con la presentación de un modelo de custodia avanzada para Bitcoin, basada en los conceptos de Miniscript y Timelocks. Este modelo de custodia avanzada tiene como objetivo facilitar compartir la custodia de fondos de manera segura entre una entidad personal o jurídica a cualquier entidad de su misma jerarquía en una línea temporal predefinida. En circunstancias excepcionales, las entidades heredera podrá liberar fondos sin necesidad de acceder a las claves privadas de la entidad que comparte, minimizando el impacto en experiencia de usuario y seguridad. Esta solución tiene el potencial de eliminar barreras para la adopción más amplia de Bitcoin.

***Palabras Clave*—bitcoin, wallet, security, custody, timelock, miniscript**

I. INTRODUCCIÓN

En este documento la custodia digital avanzada se contextualiza al conjunto de mecanismos de seguridad para la gestión y protección de activos digitales. Este conjunto de mecanismos se vuelve tangible a través de un modelo que establece condiciones específicas que deberán cumplirse para que los fondos sean liberados, garantizando así su seguridad sólo bajo circunstancias predefinidas.

Las billeteras actuales de Bitcoin son conocidas tradicionalmente como 'wallets' que tienen limitaciones de funcionalidades de custodia avanzada para transacciones. Se enfocan solamente en operaciones básicas y no ofrecen opciones como bloqueos temporales o condiciones específicas para acceder a los fondos. Dada su arquitectura y diseño, estas billeteras no son compatibles con formas avanzadas de custodia, lo que las hace menos flexibles y más rudimentarias.

La respuesta a estas limitaciones ha llegado en forma de una billetera de nueva generación. El modelo del que hablaremos pone un fuerte énfasis en la implementación de custodia avanzada, un conjunto de medidas de seguridad adicionales, para el 'control condicional' de los fondos, es decir, las reglas específicas bajo las cuales los fondos pueden ser liberados.

Con el fin de mover fondos, estas billeteras hacen uso de estándares y formatos que van evolucionando a lo largo del tiempo. Para realizar una gestión mas avanzada, actualmente se utilizan transacciones de formato Pay-to-Witness-Script Hash [1] (P2WSH). Este formato se deriva de la evolución en el mundo de las transacciones de Bitcoin para conseguir la capacidad de asignar condiciones más complejas de gestión de fondos.

La estructura y lógica de las transacciones en Bitcoin están diseñadas para especificar cómo, qué y quiénes pueden gastar los fondos. Ésto se logra a través del lenguaje de script incorporado en cada transacción: 'script de bloqueo' o scriptPubKey. Este enfoque de funcionalidad sobre el script de bloqueo ha experimentado una evolución significativa a lo largo del tiempo gracias a las mejoras y/o proposiciones (BIP).

En sus primeras etapas, las transacciones de Bitcoin eran relativamente simples y utilizaban formatos de direcciones como Pay-to-Public-Key (P2PK) y Pay-to-Public-Key-Hash (P2PKH). Las billeteras funcionaban con claves públicas, permitiendo así al emisor recibir criptomonedas. Para acceder a esos fondos, se generaba un script de bloqueo con las claves públicas de las entidades. Los fondos asignados a este bloqueo solo podían ser desbloqueados por el receptor que tuviera la clave privada correspondiente; era un mecanismo bastante sencillo.

Con la introducción de Pay-to-Script-Hash [2] (P2SH), las transacciones se volvieron más flexibles y complejas. Con P2SH se permitió la incorporación de scripts de bloqueo avanzados que podían incluir múltiples firmas o [5] Timelocks (añadidos en BIP-65[10], BIP-68[11], BIP-112[12] y BIP-113[13]) . En lugar de incluir todo el script de bloqueo en la transacción, se utilizaba un hash del script. Esto simplificaba la transacción al reducir su tamaño, y permitía al emisor enviar fondos sin necesidad de conocer los detalles del script, mejorando así la eficiencia y la seguridad del sistema.

La evolución continuó con Pay-to-Witness-Script-Hash [1] (P2WSH) heredando las ventajas de P2SH, con una mejora a nivel de optimización en transacciones y permite scripts más grandes y eficientes en términos de espacio. Esta capacidad extendida de P2WSH es especialmente beneficiosa para Miniscript, un lenguaje de script que facilita la creación de condiciones de bloqueo más complejas y variadas, desde múltiples firmas hasta bloqueos temporales y lógicas condicionales.

II. BITCOIN WALLETS

Bitcoin es una moneda completamente digital y descentralizada. La especificación de Bitcoin fue introducida en 2008 por Satoshi Nakamoto [3] y la primera prueba de concepto se creó en 2009. La red de Bitcoin actúa como una autoridad financiera no centralizada, utilizando la criptografía para controlar la transferencia y representación del dinero digital o criptomonedas.

Una cartera, billetera o wallet en el dominio de Bitcoin es un programa o dispositivo que permite a los usuarios almacenar, enviar y recibir el dinero digital. Funciona generando un par de claves criptográficas: una clave pública, exportable, utilizada para recibir fondos; y una clave privada, protegida, utilizada para acceder y gastar los fondos. La cartera mantiene un registro de las transacciones y firma digitalmente las operaciones salientes con la clave privada correspondiente.

La importancia de una billetera radica en que es el medio para gestionar de forma segura las criptomonedas, proporcionando control sobre los activos digitales y protegiendo contra el acceso no autorizado.

III. SCRIPTS EN TRANSACCIONES

El sistema de transacciones es una de las partes fundamentales del ecosistema de Bitcoin. Las transacciones son estructuras de datos que codifican la transferencia de valor entre los participantes en la red de Bitcoin y existen en diferentes formatos, mencionados anteriormente.

Estos formatos, bajo diferentes tipos de script de bloqueo, representan como y de que manera se pueden gastar los fondos e incorporar datos más complejos. Un **script de bloqueo** en Bitcoin (scriptPubKey) [15] es un conjunto de instrucciones que establecen las condiciones que deben cumplirse para que los fondos puedan ser gestionados. Este script esta incluído en la transacción que envía los fondos y se almacena dentro de la cadena de bloques (Bitcoin).

Para gastar los fondos, se debe proporcionar otro script, de firma o testigo (scriptsig o testigo en el caso de SegWit [1]) que satisface las condiciones del bloqueo ligado a la transacción (pk : public key):

```
(scriptPubKey) <pk> OP_CHECKSIG
(scriptSig)    <signature>
OK ?           <signature> <pk> OP_CHECKSIG
```

La generación de un scriptPubKey depende del tipo de transacción y las condiciones de gasto que desees establecer. Algunos ejemplos básicos de cómo se generan scriptPubKeys para diferentes tipos de transacciones en Bitcoin:

- **Pay-to-Public-Key-Hash (P2PKH)** doble hash sobre la clave pública del destinatario.
  ```
  OP_DUP OP_HASH160 <public_key_hash>
  ... OP_EQUALVERIFY OP_CHECKSIG
  ```
- **Pay-to-Script-Hash (P2SH)** doble hash en script de canjeo.
  ```
  OP_HASH160 <script_hash> OP_EQUAL
  ```
- **Pay-to-Witness-Public-Key-Hash (P2WPKH) y Pay-to-Witness-Script-Hash (P2WSH)** simples y sobre SegWit. Aqui es donde encaja Miniscript, util para incorporar scripts de bloqueo sobre estos scripts.
  ```
  (P2WPKH) 0 <public_key_hash>
  (P2WSH)  0 <script_hash>
  ```

Aunque los scripts de bloqueo tipo Pay-to-Script-Hash [2] (P2SH) siguen siendo comúnmente utilizado en transacciones de condiciones más complejas, otros formatos como Pay-to-Witness-Script-Hash [1] (P2WSH) están ganando popularidad, especialmente tras la la adopción de SegWit, una mejora de eficiencia en las transacciones [6].

Dentro del espectro de transacciones, nos encontramos con el standard de el manejo de transacciones no firmadas: **Transacciones Bitcoin Parcialmente Firmadas** (PSBT, por sus siglas en inglés), un concepto introducido en el BIP-174[16]. PSBT es un formato que permite que múltiples partes contribuyan a crear una transacción, añadir las firmas necesarias y luego transmitirla a la red de Bitcoin. La PSBT mejora la portabilidad de las transacciones no firmadas, permitiendo que múltiples participantes firmen la misma transacción de manera sencilla. Esto resulta especialmente útil para la interacción entre carteras en relaciones de tipo jerárquicas.

Para complementar estas capacidades, existen los **descriptores** (Output script descriptors o descriptores de salida) o 'manuales de instrucciones', una característica introducida en la biblioteca de Bitcoin Core [17] para

simplificar y mejorar la gestión de claves y scripts en transacciones. Fueron introducidos para mejorar la representación de direcciones y resolver el problema de que las claves públicas compartidas en ellas no contenían información sobre el tipo de scripts que utilizaban [4].

Los descriptores son una representación en forma de normas de cómo los scripts y direcciones deberían ser generadas, como si se trataran de un manual de instrucciones [7] para recuperar la parte pública o privada de las billeteras o wallets. Los descriptores proporcionan explícitamente una dirección con instrucciones programadas sobre cómo podrán gastarse, así como todas las claves y scripts necesarios para firmarlos (un manual de instrucciones o guión).

Los descriptores de salida en Bitcoin ofrecen una forma codificada y legible de incluir toda la información necesaria para generar un script de desbloqueo. Esto abarca las claves utilizadas (privadas y/o públicas), las claves públicas requeridas, el tipo de dirección y el camino de derivación asociado, así como el tipo de script empleado (multisig, segwit, etc.).

Este enfoque introduce una capa de abstracción que separa las claves de los scripts de salida (scriptPubKey), eliminando la necesidad de hacer suposiciones sobre cómo se generan estos scripts. En lugar de eso, se utiliza un gestor de scripts que almacena estos elementos. Cuando es necesario firmar una transacción, el gestor de scripts determina, basándose en los descriptores, qué elementos son necesarios para la firma. De esta manera, se garantiza un proceso de firma más preciso y eficiente.

Teniendo en cuenta la información anterior un ejemplo de descriptor que representa una cartera de tipo Pay-to-Witness-Script-Hash (P2WSH) y que sigue el estandard [9] BIP44 de jerarquia de claves:

```
p2wsh([a1b2c3d4/44'/0'/0'/0]xpub6ERApfZ
wUNrhLCkDtcHTcxd75RbzS1ed54G1LkBUHQVHQK
qhMkhgbmJbZRkrgZw4koxb5JaHWkY4ALHY2grBg
d9JU1jR5e3GK4Piz3fqzyzm/0/*)
```

Siendo los componentes de este descriptor los siguientes:

- **wpkh** es el prefijo de la función que genera un script de gasto P2WPKH. P2WPKH es el formato de dirección SegWit nativo que utiliza hashes de claves públicas para las direcciones.
- **[a1b2c3d4/44'/0'/0'/0']** es el prefijo del derivado de la clave. La cadena **a1b2c3d4** es el "fingerprint" o huella digital de la clave maestra. Este fingerprint es especialmente útil en situaciones donde se necesita referenciar la clave maestra sin exponerla. Por ejemplo, en un descriptor de Bitcoin, el fingerprint permite saber cuál es la clave maestra que se utilizó para derivar las claves subsiguientes, sin tener que incluir la clave maestra completa en el mismo descriptor. Por otro lado **/44'/0'/0'/0'** es el camino de derivación [9] BIP44 que identifica la cuenta específica y el tipo de dirección dentro de una wallet de tipo deterministica. Dada esta derivación: 44' indica que se utiliza el standard BIP44, 0' especifica el tipo de moneda, Bitcoin; 0' numero de la cuenta (dentro de la wallet) y 0' para indicar que las direcciones generadas serán direcciones de cambio.
- **xpub...zyzm** es la clave pública extendida (xpub) que se utilizará como semilla para derivar las claves públicas de las direcciones de la cartera. La inclusión de esta xpub en el descriptor permite la generación de múltiples direcciones de Bitcoin sin exponer la clave privada correspondiente.

El uso de éstos simplifica la creación de scripts de transacción al permitir una especificación más clara y concisa de los requisitos de gasto. Asimismo, facilita la construcción de contratos inteligentes y escenarios de gasto condicional, ya que los descriptores encapsularían la lógica compleja bajo una estructura legible.

IV. MINISCRIPT

Miniscript es un lenguaje de scripting utilizado en Bitcoin que simplifica y mejora la legibilidad de los scripts de transacciones [8]. Fue desarrollado para facilitar la creación de scripts complejos al mismo tiempo que mantiene la seguridad y la flexibilidad de Bitcoin. La característica principal de Miniscript es su enfoque en la composición modular, lo que permite a los desarrolladores combinar una serie de operaciones y condiciones comunes para crear scripts de forma más intuitiva. Además, Miniscript se basa en un sistema de políticas o reglas (lenguaje de políticas), que indican qué condiciones deben cumplirse para que una transacción sea válida.

Un ejemplo de Miniscript con una política simple es aquella donde los fondos solo pueden ser gastados si el propietario presenta una firma válida y cumple otra condición basada en el tiempo (después de minar 1000 bloques). Esta politica se reduciría a la siguiente expresión:

```
and_v(v:pk(K),older(1000))
```

- pk(K). Clave pública del propietario
- older(1000). 1000 bloques deben de ser minados para que los fondos puedan ser desbloqueados.

Miniscript simplifica la forma en que se crean y leen las condiciones para mover fondos en Bitcoin. Imagina que es como un "traductor" que convierte las complejas reglas de gasto en un lenguaje más sencillo y comprensible; pueden reducir errores y mejorar la legibilidad de los scripts. Esto no solo facilita la vida de los desarrolladores, sino que también hace que las transacciones sean más seguras y eficientes para todos.

V. CONDICIONALES

Los **Timelocks** son una característica de Bitcoin que permite establecer restricciones temporales en las transacciones. Estas restricciones se utilizan para la gestión avanzada de fondos, ofreciendo un control adicional sobre cuándo se pueden gastar. En términos técnicos, los timelocks se implementan sobre los scripts en las transacciones de Bitcoin.

Los timelocks se pueden implementar utilizando las operaciones *OP_CHECKLOCKTIMEVERIFY* (CLTV) y *OP_CHECKSEQUENCEVERIFY* (CSV). Estas operaciones permiten bloquear una cantidad de fondos hasta que se cumpla una cierta condición de tiempo o de número de bloques.

En un script de bloqueo (scriptPubKey) utilizando CLTV para asegurarse de que los fondos solo puedan ser gastados después de un tiempo específico (representado por *timestamp*) sería representado como:

```
<timestamp> OP_CHECKLOCKTIMEVERIFY
... OP_DROP <public_key> OP_CHECKSIG
```

Para ponerlo en términos más simples, los timelocks actúan como un temporizador que establece cuándo se puede acceder al dinero. Miniscript traduce la función *cltv(time)* a una operación *OP_CHECKLOCKTIMEVERIFY* en Bitcoin. En paralelo, los descriptores se comportan como manual de instrucciones y facilitan la organización y gestión de esos fondos.

En resumen, con la creciente adopción de tecnologías digitales y la descentralización de las finanzas, estas herramientas tienen el potencial de transformar las wallets de criptomonedas. No solo elevan los estándares de seguridad y custodia, sino que también hacen que las soluciones sean más innovadoras y accesibles para el usuario, alejándose de la complejidad técnica hacia un diseño más centrado en el usuario.

A. Wallet de custodia avanzada

Para garantizar la la seguridad en la custodia de los activos por diseño, nuestra proposicion de billetera o wallet se encuentra estructurada en dos componentes (Fig. 1): offline y online.

El componente **online** es la parte de wallet conectada a Internet encargándose de la creación de las transacciones parcialmente firmadas (PSBT) [16] con ayuda de la libreria de Bitcoin Core [17], y la inclusión de la transacción a la cadena de bloques. Ninguna funcionalidad necesita de claves privadas, y este componente no tiene custodia de ellas. Para la firma de cada transacción parcial, el componente online delegaría la función de firma al componente **offline**.

Por otro lado, el componente **offline** se encontraría totalmente aislado de la red. Se asegura la protección de los activos custodiando las claves privadas generadas y asignadas en la wallet, junto con las claves públicas.

Para una operación de custodia condicional compartida de fondos las entidades implicadas a través de sus wallets (offline-online) construyen una condición temporal y firman una parte parcial del contrato (PSBT). En este momento se genera un descriptor que definirá la operación y establecerá el bloque de manera implícita con ayuda de la tecnología mencionada anteriormente.

Para el desbloqueo de los fondos se ha diseñado un script donde una entidad -origen de los fondos- inicia un contrato condicional para asignar gestión de fondos sobre otras entidades en una jerarquía determinada. En este contexto jerárquico pueden caber diferentes casos de uso: herencias familiares, contratos administrativos o jurídicos. A partir de una fecha establecida por un contrato, las entidades -herederos de los fondos- podrán comenzar a gestionar y acceder a los activos. Esta condicion temporal ha sido gracias a la tecnología Timelock descrita de manera implícita a través de un descriptor y generado gracias a la tecnologia de Miniscript.

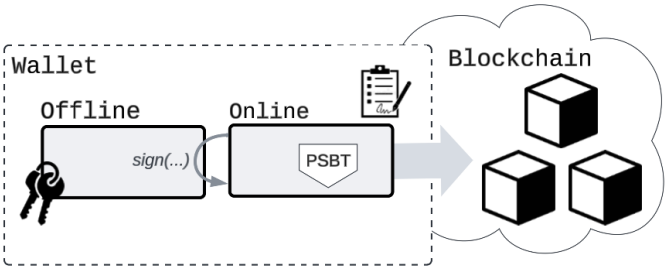


Fig. 1. Modelo Wallet Offline-Online. Responsabilidad de cada componente en la custodia de claves y operatividad con la red Bitcoin.

B. Herencia familiar (padres-hijos)

Procederemos a presentar el caso de uso correspondiente a la custodia avanzada, con particular enfoque en la herencia de padres a hijos (Ver diagrama 2). Un ejemplo con respecto a lo mencionado anteriormente, se daría cuando padres pretenden dejar un legado a partir de una edad determinada o una fecha preestablecida.

1) Lo hijos generan una clave pública y la comparten con sus padres.
2) Los padres reciben esa clave pública con la que generará un descriptor ligado a una fecha determinada (un cumpleaños, por ejemplo).
3) Los padres pasarán el descriptor generado a sus hijos. Este descriptor no contiene ninguna referencia a clave privada, pero sirve de referencia del contrato.

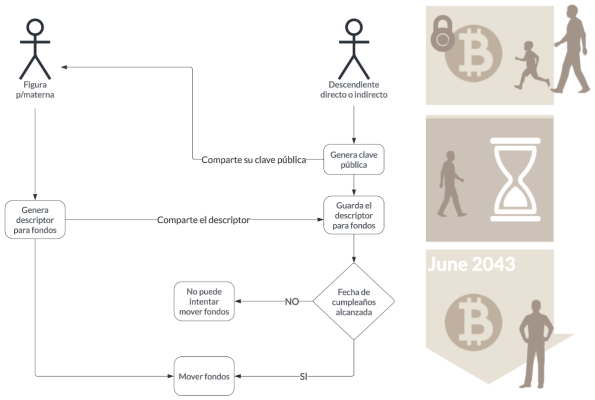


Fig. 2. Diagrama explicativo de intercambio de información necesaria para gestionar la custodia de fondos bloqueados condicionalmente y premisas necesarias para su desbloqueo.

4) Tanto padres como hijos, guardarán este descriptor en común.
5) Con estos datos, los padres pueden mover los fondos sin ningún problema, sin ningun tipo de condición. Pero los hijos aún no tienen acceso a los fondos regidos por el contrato.
6) Una vez llegue la fecha establecida en el contrato (cumpleaños), éstos serán capaces de mover los fondos y gestionarlos sin condición alguna.

Desde un aspecto técnico, el mecanismo que generaría el descriptor necesario para representar estos aspectos condicionales quedaría desglosado en una serie de sencillos pasos:

1. Creación de la política (condición) con lenguaje Miniscript

```
or(pk(@parent),and(pk(@child),after(@timestamp)
```

2. Generación del descriptor (se pueden utilizar librerias, ej. bitcoinerlab [18])

```
wsh(andor(pk(0362. .2d),after(1685972700)
...,pk(125..6d)))
```

3. Generación en una dirección Pay-To-Witness-Script-hash (P2WSH)

```
bc1qlsew8rtcwvuhxzfzg2x3v22h...e2zd8
```

C. Beneficios

Este tipo de billetera ofrece la posibilidad de que entidades distintas compartan fondos en una estructura jerárquica sin tener que compartir ni el mnemónico ni una clave privada. Además, no es necesario que ambas partes utilicen la misma billetera. La incorporación de la función de bloqueo temporal, o Timelock, añade un nivel de sofisticación que supera a las wallets multifirma convencionales.

Este enfoque aborda eficazmente los retos asociados con la herencia de Bitcoin y la recuperación de fondos en caso de pérdida de claves. Un elemento clave es el Timelock aplicado a los fondos, que permite a una segunda entidad rescatar los fondos usando sus propias claves, distintas de las que el custodio original pudo haber perdido. Así, se asegura tanto la seguridad como la accesibilidad de los fondos, incluso cuando las claves originales se pierden.

VI. CONCLUSIONES

En relación a lo expuesto, se presenta como una solución innovadora al desafío actual en la custodia avanzada de fondos. La emergencia de nuevas tecnologías como Miniscript, Timelocks y descriptores ha dado lugar a un nuevo paradigma en el diseño de wallets.

Los beneficios de utilizar esta clase de wallet son significativos. Se destaca, en primer lugar, la capacidad de compartir fondos entre múltiples partes, como en el ejemplo de la herencia entre padre e hijo. Sin embargo, su aplicabilidad es amplia y versátil, transformando a Bitcoin de una mera moneda de intercambio a una plataforma programable e ”inteligente”.

AGRADECIMIENTOS

Este modelo ha sido diseñado, concebido y validado tras una colaboración entre el equipo de Investigación y Desarrollo de la empresa **Fortris** [19], con los formantes Antonio Tóvar y Álvaro López; y los estudiantes de la Universidad de Málaga [20]: Diego García, Juan Carlos Delgado y Miguel Ángel Fortes.

REFERENCIAS

[1] Eric Lombrozo, Johnson Lau, Pieter Wuille, ”Segregated Witness (Consensus layer) BIP-141” https://github.com/bitcoin/bips/blob/master/bip-0141.mediawiki.
[2] Gavin Andresen. ”Pay to Script hash (BIP-16)”.https: //github.com/bitcoin/bips/blob/master/bip-0016.mediawiki.
[3] Nakamoto, S.: ”Bitcoin: A peer-to-peer electronic cash system (2008)”.http://bitcoin.org/bitcoin.pdf (retrieved March 2014)
[4] Pieter Wuille: ”Output script descriptor (proposal)”.https: //gist.github.com/sipa/e3d23d498c430bb601c5bca83523fa82
[5] Bit2Me Academy. Timelock’s explanation.https://academy.bit2me.com/que-es-timelock
[6] Antonopoulos, Andreas M. Mastering Bitcoin: unlocking digital crypto-currencies.
[7] BitcoinDevKit ”Documentation. Descriptors”. https://bitcoindevkit.org/descriptors
[8] Pieter Wuille, Andrew Poelstra, and Sanket Kanjalkar. ”Miniscript”. https://bitcoin.sipa.be/miniscript
[9] Marek Palatinus, Pavol Rusnak. Multi-Account Hierarchy for Deterministic Wallets (BIP-44). https://github.com/bitcoin/bips/blob/master/bip-0044.mediawiki
[10] Peter Todd. OPCHECKLOCKTIMEVERIFY (BIP-65). https://github.com/bitcoin/bips/blob/master/bip-0065.mediawiki
[11] Mark Friedenbach. Relative lock-time using consensus-enforced sequence numbers (BIP-68). https://github.com/bitcoin/bips/blob/master/bip-0068.mediawiki
[12] Mark Friedenbach. CHECKSEQUENCEVERIFY (BIP-112). https://github.com/bitcoin/bips/blob/master/bip-0112.mediawiki
[13] Mark Friedenbach. Median time-past as endpoint for lock-time calculation (BIP-113). https://github.com/bitcoin/bips/blob/master/bip-0113.mediawiki
[14] Pieter Wuille, Andrew Chow. Non-Segwit Output Script Descriptors (BIP-381). https://github.com/bitcoin/bips/blob/master/bip-0381.mediawiki
[15] Johnson Lau. Address Format for Segregated Witness (BIP-142). https://github.com/bitcoin/bips/blob/master/bip-0142.mediawiki
[16] Andrew Chow. Partially Signed Bitcoin Transaction Format (BIP-174). https://github.com/bitcoin/bips/blob/master/bip-0174.mediawiki
[17] BitcoinCore (open source project). https://bitcoincore.org/en/about/

[18] Jose-Luis Landabaso, BitcoinerLab (Bitcoin descriptors library). https://bitcoinerlab.com/modules/descriptors,https://github.com/bitcoinerlab/descriptors
[19] Fortris. Digital asset treasury operations.https://fortris.com/
[20] Universidad de Málaga - NICS (Blockchain Research lines). https://www.nics.uma.es/blockchain/

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Diseño de esquema de carga/descarga para vehículos eléctricos en entornos urbanos

Alberto Bazán Guillén [(1)], Pablo Barbecho Bautista [(2)], Mónica Aguilar Igartua [(1)]

(1) Dpto. de Ingeniería Telemática, Universitat Politècnica de Catalunya, Barcelona.

(2) Dpt. of Electrical Engineering, Electronics and Telecommunications, Universidad de Cuenca, Ecuador.

alberto.bazan@upc.edu, pablo.barbecho@ucuenca.edu.ec, monica.aguilar@upc.edu

Resumen—Con la creciente popularidad de los vehículos eléctricos la necesidad de estrategias de carga eficientes y efectivas se ha vuelto primordial. La programación de la carga de los vehículos eléctricos puede reducir los costos de energía mediante la gestión inteligente de la carga y descarga de acuerdo con las necesidades de los propietarios. La mayoría de las estrategias de carga intentan modelar las incertidumbres del sistema y planificar en consecuencia. No todos los trabajos previos tienen en cuenta la movilidad o la seguridad en la planificación de sus estrategias. En este trabajo se hará una descripción del trabajo en progreso sobre el diseño de servicios para la gestión de la carga de los vehículos eléctricos en entornos urbanos aplicando técnicas de aprendizaje por refuerzo federado e incluyendo redes vehiculares y gestión de la movilidad urbana.

Palabras Clave*— vehículo eléctrico, planificación de carga, Aprendizaje por Refuerzo Federado, *Mobility Hub

I. Introducción

El auge de los vehículos eléctricos ha crecido en los últimos años. Los avances en la tecnología y la preocupación por el medio ambiente propician su inserción como pieza clave en la movilidad urbana. Con el crecimiento del número de vehículos eléctricos crece la preocupación por sus sistemas de carga, específicamente los relacionados con los puntos de carga llamados electrolineras o estaciones de carga.

El consumo de un vehículo eléctrico puede llegar a ser muy grande, incluso en transporte privado de pasajeros. Tanto las capacidades de las baterías de estos vehículos como las estaciones de carga manejan grandes cantidades de potencia en cada acción relacionada con la carga. El conocimiento de la demanda que tendrá cada estación de carga, así como una óptima planificación de carga por parte de los vehículos, son temas de interés de los investigadores en los últimos años [1].

A la par, el desarrollo de las redes vehiculares permite la comunicación entre vehículos en el entorno urbano y, a su vez, la de estos con estaciones terrestres conectadas a la red. La interconexión de los vehículos favorece la creación de servicios vinculados con las comunicaciones y el uso de sistemas inteligentes de transporte.

Existen varias referencias de sistemas de planificación de carga para vehículos eléctricos que buscan maximizar el beneficio tanto de los usuarios de los vehículos como de la red eléctrica [2]–[4]. Estos sistemas pueden utilizar técnicas de aprendizaje de máquina para su solución. Para la creación de servicios de carga usando aprendizaje de máquina se suelen necesitar una gran cantidad de datos de movilidad para extraer de estos los patrones necesarios para el diseño de un plan de carga. Los datos de movilidad son difíciles de obtener de forma libre y abierta, de ahí que el principal método de aprendizaje usado es el Aprendizaje por Refuerzo o *Reinforcement Learning* (RL) [5]–[11]. El RL no necesita datos previos para su entrenamiento, ya que genera una solución a partir del entrenamiento de sus agentes en un entorno con ciertas reglas donde un sistema de penalización y recompensa conforma la estrategia óptima [12].

La congestión, la contaminación del aire y la promoción de modos sostenibles de transporte público son algunos de los desafíos ambientales a los que se enfrentan las ciudades a nivel mundial. Los Centros de Movilidad o *Mobility Hubs*, surgen como un mecanismo para avanzar hacia una red de transporte sostenible y se encuentran en varias etapas de implementación en ciudades de todo el mundo. Un *Mobility Hub* es un espacio con una oferta de modos de transporte diferentes conectados y complementados con instalaciones mejoradas y funciones de información para atraer y beneficiar al/a la viajero/a [13]. Son centros focalizados en viajes dentro de zonas urbanizadas de corta distancia en ciudades o pueblos y sus exteriores. Sus principales objetivos varían dependiendo de las necesidades de cada ciudad, pero son principalmente ambientales y socioeconómicos. Entre los objetivos ambientales están (i) la provisión de transporte sostenible, especialmente el impulsado por energía eléctrica; (ii) facilitar a la ciudadanía el intercambio en viajes multimodales; y (iii) la reducción de emisiones nocivas disminuyendo la cantidad de vehículos de combustión interna, entre otros. En cuanto a objetivos socioeconómicos destacan (i) evitar la congestión del tráfico; (ii) proveer las comodidades y tecnologías para un

trayecto multimodal reduciendo los tiempos de viaje y los accidentes; y (iii) mejorar la calidad de vida de las personas [13].

En este trabajo se hará una descripción del trabajo en progreso correspondiente a la tesis de doctorado de Alberto Bazán sobre el diseño de servicios para la gestión de la carga de los vehículos eléctricos en entornos urbanos aplicando técnicas de aprendizaje de máquina e incluyendo redes vehiculares y gestión de la movilidad urbana.

II. Antecedentes

En esta sección se describirán brevemente aquellas tecnologías y metodologías que serán utilizadas en nuestras propuestas.

A. Redes vehiculares

Dentro de las redes vehiculares el término V2X (vehículo a todo) se introduce recientemente por el grupo 3GPP como la comunicación entre un vehículo y cualquier otro dispositivo. Las redes vehiculares permiten el intercambio de información entre vehículos y diferentes elementos de un sistema inteligente de transporte como otros vehículos, peatones y otras infraestructuras de transporte. Este tipo de redes pueden soportar servicios inteligentes como los de sistema de planificación de carga de vehículos eléctricos. Para que las redes vehiculares funcionen, es necesario disponer de un protocolo de encaminamiento adecuado para este tipo de redes inalámbricas, altamente dinámicas cuyos nodos (vehículos) potencialmente pueden moverse a altas velocidades produciendo frecuentes desconexiones, y bajo diversas densidades de vehículos. Nuestro grupo de investigación SISCOM (*https://siscom.upc.edu*) tiene una larga trayectoria en el diseño de estos protocolos de encaminamiento [14]–[16].

B. Aprendizaje de Máquina y Aprendizaje por Refuerzo

La predicción de la demanda de energía eléctrica en las estaciones de carga es un paso clave en la gestión de la energía. Esta predicción es necesaria para la planificación inteligente de un sistema de carga para vehículos eléctricos. Algunos autores en la literatura proponen métodos de predicción de consumo basados en fuentes de información recolectada previamente como: programación dinámica estocástica [17], modelos de predicción Bayesianos [18], entre otros. Estos métodos consideran el uso de la energía actual y futura como una función del historial de uso de energía pasado. A pesar de la capacidad de estos métodos para predecir patrones de carga a corto plazo, no son fiables en el caso de condiciones no estacionarias o datos no lineales [19].

El uso del Aprendizaje de Máquina, o *Machine Learning* (ML) es ampliamente extendido en la predicción de consumo en las estaciones de carga. En [19] utilizan Redes Neuronales con arquitectura Long-Short Term Memory (LSTM) para desarrollar un sistema óptimo de carga/descarga de vehículos eléctricos para minimizar los picos de electricidad en la demanda. En [20] los autores desarrollan un programa de manejo de reservas para la carga de vehículos eléctricos en una estación de carga usando técnicas de Aprendizaje Profundo o *Deep Learning*. También en [21] se utilizan distintas herramientas de ML para el diseño de estrategias inteligentes de carga.

Para el uso de estrategias de ML es necesario poseer datos recolectados anteriormente de los cuales se extraerán patrones para poder predecir comportamientos futuros. Los datos de consumo, movilidad y otros relacionados con vehículos eléctricos no son fáciles de obtener. Ante la ausencia de datos previos se suelen utilizar estrategias de RL donde solo se necesitan agentes inteligentes y un entorno que les proporciona beneficios o penalizaciones como consecuencia de sus acciones. De esta forma algunos autores [10], [11], [22] han definido sus estrategias de planificación de carga.

C. Aprendizaje Federado y Aprendizaje por Refuerzo Federado

El RL enfrenta varios desafíos en su implementación práctica. Uno de ellos es el problema de exploración en espacios donde es difícil recolectar muestras que abarquen todo el gran espacio de muestreo. Además, la eficiencia de aprendizaje es un problema debido a la baja eficiencia de las muestras. Los enfoques distribuidos y paralelos de RL han surgido para abordar estos problemas, pero generalmente requieren recopilar todos los datos de los agentes en un servidor central, lo que plantea preocupaciones sobre la privacidad y el riesgo de espionaje [12].

El Aprendizaje Federado o *Federated Learning* (FL) surge como una solución prometedora que ha recibido un interés considerable por parte de la academia y la industria. FL permite el uso de datos aislados de múltiples dispositivos sin violar la política de protección de la privacidad. También permite la integración de características parciales observadas por cada agente, lo que proporciona información más completa para la toma de decisiones. Recientemente, un campo emergente llamado Aprendizaje por Refuerzo Federado o *Federated Reinforcement Learning* (FRL) [12] combina las ventajas tanto de FL como de RL. No solo puede proporcionar a los agentes la experiencia para aprender a tomar buenas decisiones en entornos desconocidos y dinámicos, sino también entrenar un modelo global de forma colaborativa sin compartir sus propias experiencias. Algunos autores han utilizado FRL [6], [7], [9] en el diseño de esquemas de carga aprovechando los beneficios que brinda.

III. Trabajos relacionados

Existen en la literatura diferentes trabajos relacionados con el diseño de un esquema óptimo de carga para vehículos eléctricos. Los diseños buscan optimizar el beneficio de los usuarios de vehículos eléctricos con disminuciones de tiempos de espera, tiempo total de viaje o beneficio económico de acuerdo al precio de la electricidad en cada momento. Algunos trabajos pretenden maximizar el beneficio de la comunidad con el aplanamiento de las curvas de consumo, la predicción de energía demandada en las estaciones de carga, entre otros.

La mayoría de los trabajos plantean el problema como un problema de optimización donde en una estación de carga hay una cantidad finita de espacios de carga y los vehículos que pueden llegar o partir en cualquier

momento, mientras estén conectados, podrán cargar o no en cada momento según convenga. La decisión deberá ser óptima y debe tener en cuenta el tiempo de disponibilidad para cargar que será definido previamente. En trabajos como [7], [8], [11], [19], [20], [23]–[25] la cantidad y los horarios de llegada y partida de los vehículos son definidos artificialmente. Aunque algunos son definidos a partir de datos reales ninguno inserta el comportamiento de los vehículos eléctricos dentro de un problema real de movilidad urbana en una ciudad.

La capacidad de almacenamiento y de entrega de energía de vuelta a la red de algunos vehículos eléctricos permite diseños de planificación de carga en ambos sentidos. Con la carga y la descarga de los vehículos en las estaciones se logran beneficios mayores para el usuario del vehículo eléctrico gracias a la compra/venta de energía y también se facilita el aplanamiento de la curva de demanda de la red. Trabajos como [7], [8], [11], [24] han demostrado buenos resultados con esta estrategia.

La movilidad no se suele tener en cuenta en el diseño de planes de carga para vehículos eléctricos. La existencia de *Mobility Hubs* trae consigo la posibilidad de aparcar los vehículos eléctricos personales como parte de un viaje multimodal del usuario. Estos pueden ser concebidos como estaciones de carga y es necesario conocer la movilidad de los usuarios para poder realizar un plan de carga completo que incluya a varios *Mobility Hubs* dentro de un entorno urbano. En trabajos como [5], [10], [26] sí tienen en cuenta la movilidad para la planificación de la carga, pero no brindan los beneficios de optimización mediante carga/descarga o la seguridad y privacidad que aporta el aprendizaje federado.

En este trabajo se propone el diseño de un esquema de carga/descarga para vehículos eléctricos usando FRL para maximizar el beneficio de los usuarios y de la red. El esquema incluirá la movilidad entre varios *Mobility Hubs* dentro del entorno urbano de la ciudad de Barcelona.

IV. Entorno de simulación

El desarrollo de un servicio de carga inteligente conlleva el uso de varios sistemas relacionados con: la movilidad, las comunicaciones y el aprendizaje máquina. La implementación de dichos sistemas en un entorno real sería inviable y, por esa razón, se necesita un entorno de simulación para el diseño del servicio propuesto. Para la simulación de movilidad se utilizará el simulador de movilidad urbana SUMO [27]. SUMO es un *software* libre para la simulación de tráfico microscópico ampliamente usado en la simulación de sistemas inteligentes de transporte y la planificación de rutas, entre otras funciones. Para el mejor manejo de estas herramientas en la confección del escenario de simulación en SUMO se utilizará la herramienta STGT [28].

En SUMO se simularán las comunicaciones utilizando el banco de pruebas de redes modulares objetivas en C++ (OMNET++) [29] específicamente usando VEINs [30] para la comunicación entre vehículos. El control paso a paso del estado de los vehículos será controlado usando la interfaz de comunicación TraCI [31].

El entrenamiento para el aprendizaje federado se hará utilizando Flower [23]. Flower es una infraestructura de FL que ofrece instalaciones para ejecutar experimentos de FL a gran escala y considerar escenarios de dispositivos FL muy heterogéneos. Flower es de código abierto y es adoptado por las principales organizaciones de investigación tanto en la academia como en la industria [23].

En el esquema de carga/descarga que se diseñará, los vehículos eléctricos serán los agentes del sistema de FRL. Los estacionamientos de los distintos *Mobility Hubs* dentro de la ciudad de Barcelona serán las estaciones de carga/descarga de los vehículos. En cada interacción con el entorno los agentes entrenarán sus modelos propios y lo compartirán con el coordinador central del aprendizaje federado para que éste distribuya la agregación de los mismos en la siguiente etapa.

V. Conclusiones y Trabajo futuro

En este trabajo hemos mostrado el estado del avance del trabajo de tesis del 1er autor sobre el diseño de servicios inteligentes para vehículos eléctricos. Una vez finalizado el diseño de la plataforma de simulación, nuestro trabajo futuro se centrará en proponer un esquema de aprendizaje por refuerzo federado para gestionar eficientemente el servicio de carga de la flota de vehículos eléctricos en una ciudad.

Para el desarrollo del esquema se seguirán los siguientes pasos:

(i) Primeramente, es necesario crear un escenario de simulación realística sobre la ciudad de Barcelona. Para esto utilizará SUMO y OpenStreetMap (https://www.openstreetmap.org) para crear la red de calles e intersecciones de la ciudad. Las rutas del escenario se crearán usando las herramientas de SUMO apoyándose en datos reales de tráfico obtenidos de fuentes libres como Open Data Barcelona (https://opendata-ajuntament.barcelona.cat).

(ii) En segundo lugar, será necesario implementar un sistema FRL capaz de simular el entrenamiento de los agentes y la comunicación entre ellos. La infraestructura Flower será la encargada de lograrlo utilizando TraCI como interfaz con SUMO para asignar el rol de agente de entrenamiento a cada vehículo implicado en la simulación y ejecutar las acciones propias del entrenamiento por refuerzo.

(iii) Después se agregará comunicación realista a la simulación con Veins, OMNET++ y un protocolo de encaminamiento para redes vehiculares basado en [14] para agregar veracidad a las comunicaciones del entrenamiento federado.

(iv) Por último, se realizará la optimización de las *recompensas* y *castigos* del entrenamiento por refuerzo utilizando variables relacionadas con el beneficio asociado a la carga y el asociado a la movilidad entre los *Mobility Hubs*.

AGRADECIMIENTOS

Este trabajo está parcialmente soportado por los proyectos de investigación "Anonymization technology for AI-based analytics of mobility data (MOBILYTICS)"; TED2021-129782B-I00 financiado por MCIN/AEI/10.13039/501100011033 y por la European Union NextGenerationEU/ PRTR; por "Enhancing Communication Protocols with Machine Learning while Protecting Sensitive Data (COMPROMISE)" PID2020-113795RB-C3X financiado por MCIN/AEI/10.13039/501100011033; y por la Generalitat de Catalunya con la ayuda AGAUR "2021 SGR 01413". Alberto Bazán dispone de una beca FPI-MICINN de ayudas para contratos predoctorales para la formación de doctores 2021 del Subprograma Estatal de Formación del Programa Estatal para Desarrollar, Atraer y Retener Talento, en el marco del Plan Estatal de Investigación Científica, Técnica y de Innovación 2021-2023.

REFERENCIAS

[1] Y. Cao, Y. Zhang, and C. Gu, *Automated and Electric Vehicle: Design, Informatics and Sustainability*, vol. 3. Springer Nature, 2022.

[2] T. U. Solanke, V. K. Ramachandaramurthy, J. Y. Yong, J. Pasupuleti, P. Kasinathan, and A. Rajagopalan, "A review of strategic charging–discharging control of grid-connected electric vehicles," *J. Energy Storage*, vol. 28, p. 101193, 2020, doi: https://doi.org/10.1016/j.est.2020.101193.

[3] A. S. Al-Ogaili *et al.*, "Review on Scheduling, Clustering, and Forecasting Strategies for Controlling Electric Vehicle Charging: Challenges and Recommendations," *IEEE Access*, vol. 7, pp. 128353–128371, 2019, doi: 10.1109/ACCESS.2019.2939595.

[4] Y. Zheng, S. Niu, Y. Shang, Z. Shao, and L. Jian, "Integrating plug-in electric vehicles into power grids: A comprehensive review on power interaction mode, scheduling methodology and mathematical foundation," *Renew. Sustain. Energy Rev.*, vol. 112, pp. 424–439, 2019, doi: https://doi.org/10.1016/j.rser.2019.05.059.

[5] A. Viziteu *et al.*, "Smart Scheduling of Electric Vehicles Based on Reinforcement Learning," *Sensors*, vol. 22, no. 10, p. 3718, 2022.

[6] M. Moniruzzaman, A. Yassine, and R. Benlamri, "Blockchain and Federated Reinforcement Learning for Vehicle-to-Everything Energy Trading in Smart Grids," *IEEE Trans. Artif. Intell.*, pp. 1–15, 2023, doi: 10.1109/TAI.2023.3262597.

[7] S. Lee and D.-H. Choi, "Dynamic pricing and energy management for profit maximization in multiple smart electric vehicle charging stations: A privacy-preserving deep reinforcement learning approach," *Appl. Energy*, vol. 304, p. 117754, 2021.

[8] Z. Zhang, Y. Jiang, Y. Shi, Y. Shi, and W. Chen, "Federated Reinforcement Learning for Real-Time Electric Vehicle Charging and Discharging Control," in *2022 IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 1717–1722.

[9] Y. Chu, Z. Wei, X. Fang, S. Chen, and Y. Zhou, "A multiagent federated reinforcement learning approach for plug-in electric vehicle fleet charging coordination in a residential community," *IEEE Access*, vol. 10, pp. 98535–98548, 2022.

[10] C. Zhang, Y. Liu, F. Wu, B. Tang, and W. Fan, "Effective charging planning based on deep reinforcement learning for electric vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 1, pp. 542–554, 2020.

[11] N. Mhaisen, N. Fetais, and A. Massoud, "Real-time scheduling for electric vehicles charging/discharging using reinforcement learning," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 2020, pp. 1–6.

[12] J. Qi, Q. Zhou, L. Lei, and K. Zheng, "Federated reinforcement learning: Techniques, applications, and open challenges," *arXiv Prepr. arXiv2108.11887*, 2021.

[13] T. Arnold, M. Frost, A. Timmis, S. Dale, and S. Ison, "Mobility hubs: review and future research direction," *Transp. Res. Rec.*, vol. 2677, no. 2, pp. 858–868, 2023.

[14] L. L. Cárdenas, A. M. Mezher, P. A. Barbecho Bautista, J. P. Astudillo León, and M. A. Igartua, "A Multimetric Predictive ANN-Based Routing Protocol for Vehicular Ad Hoc Networks," *IEEE Access*, vol. 9, pp. 86037–86053, 2021, doi: 10.1109/ACCESS.2021.3088474.

[15] A. M. Mezher and M. A. Igartua, "G-3MRP: A game-theoretical multimedia multimetric map-aware routing protocol for vehicular ad hoc networks," *Comput. Networks*, vol. 213, p. 109086, 2022.

[16] J. Montenegro, C. Iza, and M. Aguilar Igartua, "Detection of position falsification attacks in VANETs applying trust model and machine learning," in *Proceedings of the 17th ACM symposium on performance evaluation of wireless ad hoc, sensor, & ubiquitous networks*, 2020, pp. 9–16.

[17] C. Luo, Y.-F. Huang, and V. Gupta, "Stochastic Dynamic Pricing for EV Charging Stations With Renewable Integration and Energy Storage," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1494–1505, 2018, doi: 10.1109/TSG.2017.2696493.

[18] A. W. Dante, S. Kelouwani, K. Agbossou, N. Henao, J. Bouchard, and S. S. Hosseini, "A Stochastic Approach to Designing Plug-In Electric Vehicle Charging Controller for Residential Applications," *IEEE Access*, vol. 10, pp. 52876–52889, 2022, doi: 10.1109/ACCESS.2022.3175817.

[19] M. Ghafoori, M. Abdallah, and S. Kim, "Electricity peak shaving for commercial buildings using machine learning and vehicle to building (V2B) system," *Appl. Energy*, vol. 340, p. 121052, 2023, doi: https://doi.org/10.1016/j.apenergy.2023.121052.

[20] B. Alshehhi, A. Karapetyan, K. Elbassioni, S. C.-K. Chau, and M. Khonji, "DClEVerNet: Deep Combinatorial Learning for Efficient EV Charging Scheduling in Large-scale Networked Facilities," *arXiv Prepr. arXiv2305.11195*, 2023.

[21] K. L. López, C. Gagné, and M.-A. Gardner, "Demand-side management using deep learning for smart charging of electric vehicles," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2683–2691, 2018.

[22] Q. Dang, D. Wu, and B. Boulet, "A q-learning based charging scheduling scheme for electric vehicles," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019, pp. 1–5.

[23] D. J. Beutel *et al.*, "Flower: A friendly federated learning research framework," *arXiv Prepr. arXiv2007.14390*, 2020.

[24] L. Ren, M. Yuan, and X. Jiao, "Electric vehicle charging and discharging scheduling strategy based on dynamic electricity price," *Eng. Appl. Artif. Intell.*, vol. 123, p. 106320, 2023.

[25] J. Wu and Q.-S. Jia, "On optimal charging scheduling for electric vehicles with wind power generation," *Fundam. Res.*, 2022.

[26] J.-H. Qian, Y.-X. Zhao, and W. Huang, "Model improvement and scheduling optimization for multi-vehicle charging planning in IoV," *Phys. A Stat. Mech. its Appl.*, vol. 621, p. 128826, 2023.

[27] D. Krajzewicz, "Traffic simulation with SUMO--simulation of urban mobility," *Fundam. traffic Simul.*, pp. 269–293, 2010.

[28] P. Barbecho Bautista, L. F. Urquiza-Aguiar, and M. Aguilar Igartua, "STGT: SUMO-Based Traffic Mobility Generation Tool for Evaluation of Vehicular Networks," in *Proceedings of the 18th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, 2021, pp. 17–24, doi: 10.1145/3479240.3488523.

[29] A. Varga, "OMNeT++," in *Modeling and Tools for Network Simulation*, K. Wehrle, M. Güne\cs, and J. Gross, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 35–59.

[30] C. Sommer, R. German, and F. Dressler, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Trans. Mob. Comput.*, vol. 10, no. 1, pp. 3–15, 2011, doi: 10.1109/TMC.2010.133.

[31] A. Wegener, M. Piórkowski Micha\land Raya, H. Hellbrück, S. Fischer, and J.-P. Hubaux, "TraCI: An Interface for Coupling Road Traffic and Network Simulators," in *Proceedings of the 11th Communications and Networking Simulation Symposium*, 2008, pp. 155–163, doi: 10.1145/1400713.1400740.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Plataforma de posicionamiento pasivo en redes WiFi

Israel Martin-Escalona, Enrica Zola, Olga León, Albert Saez Núñez, Nestor Gonzalez Diaz
Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya (UPC)
C. Jordi Girona 1-3, 08034 Barcelona.
(israel.martin, enrica.zola, olga.leon)@upc.edu, albertsigmaeng@gmail.com, nestor.gonzalez.diaz@upc.edu

Resumen

El posicionamiento en redes Wi-Fi vió un impulso notable con la aparición de la norma IEEE 802.11mc, que introduce la posibilidad de calcular de forma precisa el tiempo de ida y vuelta (RTT) de una trama entre dos dispositivos Wi-Fi. Este proceso conlleva la inyección de tráfico de localización, lo que reduce la capacidad disponible en la red Wi-Fi para otros servicios. Existen varias propuestas presentadas para mejorar la escalabilidad del sistema, pero la dificultad de implementar dichas propuestas en dispositivos reales hace que no existan datos de campo que corroboren los resultados analíticos y modelos simulados existentes. Este trabajo establece las bases para la creación de una plataforma que permita la implementación de este tipo de soluciones. Para ello se ha modificado el firmware abierto OpenFWWF y se ha procedido a implementar el algoritmo *passive TDOA* como prueba de concepto. Los resultados iniciales muestran un rendimiento válido para su uso en entornos de interior y abren una vía con la que validar algoritmos de localización en cualquiera de sus vertientes (precisión, escalabilidad, seguridad).

***Palabras Clave*—Localización en interiores, Wi-Fi RTT, FTM, IEEE 802.11mc, OpenFWWF, seguridad.**

I. INTRODUCCIÓN

Los servicios de localización se han revelado en los últimos años como esenciales a la hora de abordar tanto la gestión de las redes de comunicación, cada vez más capaces, complejas y heterogéneas, como en el desarrollo de nuevos servicios de alto valor añadido o el enriquecimiento de los ya existentes [1]. Si bien los sistemas satelitales (GNSS) como GPS se han posicionado como una solución global al posicionamiento en exterior, su rendimiento decae rápidamente en condiciones de falta de visibilidad directa con el cielo. Junto a los ya conocidos cañones urbanos (i.e. conjunto de edificios altos y calles estrechas) han irrumpido un ingente conjunto de servicios basados en localización destinados a ser desplegados en el interior de edificios, donde la degradación de la señal GNSS es extrema y más aún, donde la precisión requerida (1-2 metros) excede las capacidades que frecuentemente ofrecen los receptores GNSS comerciales [2].

Varias tecnologías han sido propuestas como reemplazo de GNSS en esos escenarios [3]: ultrasonidos, Ultra-Wide Band (UWB), Bluetooth, Identificación por Radio Frecuencia (RFID), redes celulares 2G-5G, Wireless Fidelity (Wi-Fi), etc. De todas ellas, Wi-Fi destaca por su ubicuidad, aunque al tratarse de una red de comunicaciones carece de métricas específicas para el posicionamiento, empleando generalmente para ello el indicador del nivel de señal recibida (RSSI) [4].

Pese a que el RSSI se considera un dato pasivo, es decir disponible en cualquier dispositivo Wi-Fi sin necesidad de inyectar tráfico de localización específico, su alta volatibilidad y extrema dependencia del entorno, desaconseja su uso en sistemas de posicionamiento. Los sistemas basados en huella de señal (fingerprinting) fueron inicialmente propuestos para afrontar dichos handicaps, si bien introdujeron otras limitaciones de igual consideración, como la necesidad de construir una base de datos de RSSI previamente al despliegue del sistema, así como el mantenimiento constante de los datos recogidos en dicha base de datos [5].

El estándar 802.11mc [6] introdujo la posibilidad de estimar con una alta precisión el tiempo de ida y vuelta (RTT) desde una estación (i.e. dispositivo de usuario (UE)) hasta un punto de acceso (AP), tal y como ilustra la Fig. 1. El proceso se inicia con el usuario enviando al AP una trama de petición Fine Timing Measurement (FTM) con la que estimar el RTT entre ambos. Una vez aceptada la petición, el AP inicia el proceso de estimación, enviando una trama FTM, que es respondida por el UE con una trama ACK. Durante este proceso, el AP guarda las marcas temporales correspondientes al instante de salida de la trama FTM y la recepción del ACK (t_1 y t_4 en la Fig. 1); mientras que el UE hace lo propio con el instante de recepción de la trama FTM y la salida del ACK (t_2 y t_3 en

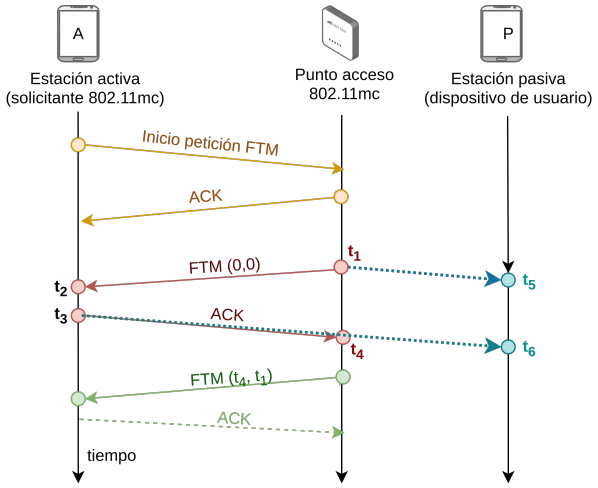


Figura 1. Algoritmo passive-TDOA bajo IEEE 802.11mc.

la Fig. 1). Finalmente, el AP envía una nueva trama FTM con el único propósito de proporcionar al UE las marcas temporales t_1 y t_4 y permitirle el cálculo del RTT como:

$$rtt_k = (t_4 - t_1) - (t_3 - t_2). \quad (1)$$

Si bien el estándar IEEE 802.11mc permite la estimación de distancias con precisiones medias entre 1 y 2 metros, su obtención requiere inyectar tráfico de localización, lo que va en detrimento de la capacidad de la red para atender otros servicios. Se requiere por tanto de una solución pasiva que, aún empleando tramas FTM para estimar el RTT, sea capaz de reducir la cantidad de tráfico de localización requerido y por ende, mejorar la escalabilidad del sistema de posicionamiento. Una posible propuesta en este sentido es el algoritmo passive TDOA [7], donde una única estación ejecutando el procedimiento descrito en IEEE 802.11mc (estación activa) permite al resto de usuarios de la red, obtener su propia posición. La Fig. 1 muestra este concepto, introduciendo una estación pasiva, que monitoriza las tramas transmitidas por el resto de usuarios. De esta forma, tanto las trama FTM enviada por el AP como el consiguiente ACK remitido por la estación activa, son recibidas en la estación pasiva y capturado el instante de recepción de ambas (i.e. t_5 y t_6 en la Fig. 1). La diferencia entre ambos tiempos supone un Observed Time Difference of Arrival (TDOA) y por tanto genera una hipérbola de posibles posiciones. La intersección entre múltiples hipérbolas permitiría a un algoritmo de multilateración el cálculo de la posición de la estación pasiva sin que esta tuviera que inyectar tráfico de localización.

El algoritmo passive TDOA, como tantos otros, ha sido analizado mediante estudios analíticos y modelos simulados. Sin embargo, su correcta implementación requiere del acceso al hardware del dispositivo de red o del firmware que lo gobierna. Lamentablemente, los fabricantes de dispositivos de red no ofrecen tal acceso, lo que hace inviable tal implementación. Este trabajo aborda esta problemática y toma la única implementación conocida de libre acceso a un firmware IEEE 802.11, denominada OpenFWWF [8], para extenderla y permitir la toma de medidas precisas con las que implementar la técnica descrita en [7]. Aunque OpenFWWF está dirigido únicamente a dispositivos Broadcom, supone una base consistente sobre la que evaluar las prestaciones del algoritmo passive TDOA, así como un punto fundacional para posteriores desarrollos relacionados con este u otros algoritmos de posicionamiento basados en medidas temporales.

II. PLATAFORMA DE LOCALIZACIÓN

A. *Arquitectura del modelo software*

La arquitectura software considerada, ilustrada en la Fig. 2, está basada en la empleada en los sistemas operativos Linux, donde los procesos de usuario se ejecutan en un entorno diferente al del núcleo del sistema. Los procesos de usuario interactúan con el hardware a través de llamadas al sistema, que formalizan una vía de acceso, restringida y bajo supervisión del núcleo del sistema, a los recursos hardware del mismo. Ya en el espacio de kernel, se establece una cadena de componentes software, con roles diferenciados, que realizan las acciones solicitadas bien por los procesos de usuario, bien por el propio núcleo del sistema. En el caso del intercambio de tramas Wi-Fi, dicha cadena podría resumirse en: *softmac*, driver de dispositivo y firmware. El *softmac* intregra la lógica común a cualquier dispositivo Wi-Fi; el driver del dispositivo (*b43* para el caso de equipos Broadcom) se encarga de particularizar las acciones solicitadas por el *softmac* al dispositivo concreto. Finalmente, el firmware del dispositivo integra la lógica más esencial y específica y por ende destinada a dialogar de forma directa con el hardware concreto del dispositivo. Para facilitar el acceso por parte del usuario a estos datos, se ha desarrollado un módulo denominado *ptdk*, que actúa de middleware almacenando los datos reportados por el firmware en una cola circular y ofreciéndolos a los distintos procesos que los deseen consumir. Para ello se emplea un enlace vía *NetLink* [9].

La implementación del algoritmo passive TDOA pasa por la modificación de la lógica definida en el firmware, de forma que se puedan capturar las marcas temporales que permitirán el cálculo del TDOA, en el punto más cercano al instante de recepción de la cabecera PLCP. OpenFWWF sin embargo, no incluye en su máquina de estados la gestión de las tramas FTM. Por ello, previamente al desarrollo de toda la lógica necesaria para su tratamiento, se ha optado por una aproximación equivalente, sustituyendo las tramas FTM y ACK, por el binomio de tramas de control RTS y CTS. Pese a que el procedimiento de reserva RTS/CTS tampoco se encuentra implementado en OpenFWWF, sí se identifica su recepción, lo que facilita el desarrollo sin impactar ni limitar el posterior estudio del rendimiento del passive TDOA (e.g. precisión de las medidas) en condiciones de tráfico real.

El principal reto de la implementación de un algoritmo basado en tiempos de llegada o partida, es tomar las medidas temporales con suficiente precisión. OpenFWWF únicamente ofrece la posibilidad de tomar medidas temporales a partir de un oscilador de 8 MHz, lo que se

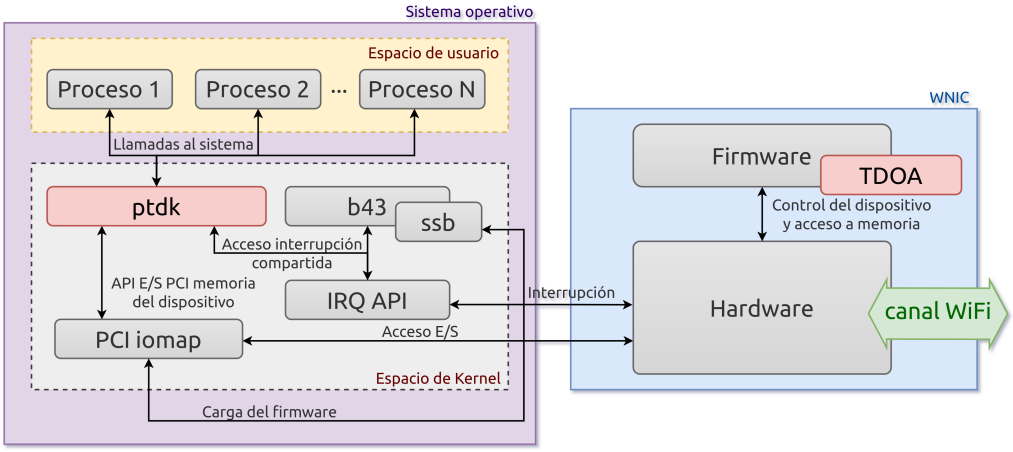


Figura 2. Architectura del sistema.

traduce en un error de cuantificación mínimo de 125 ns. Dicho error temporal equivale a 37.5m de error, valor inasumible para aplicaciones en interiores. Es por ello que se ha optado por emplear otra estrategia que consiste en emplear un timer interno del firmware, el cual permite el acceso a un oscilador de 88 MHz, reduciendo el potencial error de cuantificación a 11.36 ns (i.e. 3.4m de error). Estas cotas son más cercanas a los requerimientos de las aplicaciones en interiores, potencialmente alcanzables de aplicar etapas de filtrado que permitan el refinamiento de los datos obtenidos.

La lógica de las modificaciones concretas realizadas en el código ensamblador de OpenFWWF quedan detalladas en la Fig. 3. Se han añadido dos puntos de control en la cadena que evalúa las tramas de control, con el objetivo de detectar las tramas RTS/CTS. Cuando se recibe una trama RTS se inicia un temporizador vinculado al *transmitter address* de la trama RTS y se queda a la espera de recibir la pertinente trama CTS. Al recibirse una trama CTS, se compara el *destination address* con el transmitter address del RTS que dio inicio al temporizador. Superada esa verificación, se procede a detener el temporizador y calcular el tiempo transcurrido entre la recepción de la cabecera PLCP de las trams RTS y CTS. Este valor es remitido al módulo *ptdk* en el kernel, que finalmente ofrecerá las medidas recogidas a los distintos procesos de usuario.

III. DISCUSIÓN Y LÍNEAS DE INVESTIGACIÓN

Los primeros resultados obtenidos empleando el modelo de distancia teórico $d = v_p * RTT/2$, con v_p igual a la velocidad de propagación de la luz, certifican que el sistema, una vez calibrado, ofrece un error de estimación de distancia entorno a los 3 metros. Se hace necesario en primer lugar, proseguir con el estudio del rendimiento de la implementación. En concreto, caracterizar los distintos focos de error y cómo impactan en la distancia calculada para posteriormente, proponer modelos de distancia que permitan un refinamiento de las medidas obtenidas y permitan la reducción del error en los TDOA calculados.

Demostrada la viabilidad de la implementación actual, se procederá a modificar el código OpenFWWF para rastrear tramas FTM y poder así emplear las medidas precisas calculadas por la estación activa para el refinamiento de las estimadas en la estación pasiva, de acuerdo a lo indicado en [7]. Otra posible línea de investigación consiste en aplicar los parches generados en OpenFWWF a una implementación mediante NexMon [10]. Esta vía, pese a ser más compleja, tiene varios alicientes. Por una parte amplía el número de dispositivos potenciales (ej. dispositivos Android) y por otro, permitiría el uso de información complementaria (e.g. información del estado del canal (CSI)) que permitirían desarrollar sistemas de posicionamiento de mayor precisión.

Otra línea de investigación que actualmente se está abriendo camino es el uso de medidas RTT en soluciones *fingerprinting*. La combinación de una solución pasiva como *fingerprinting* con observables también pasivos como los estimados mediante el *passive TDOA* abre una nueva vía con la que impulsar la escalabilidad y precisión de este tipo de sistemas.

Paralelamente al desarrollo de un sistema cada vez más capaz en términos de precisión y escalabilidad, una línea de investigación que permanece abierta es el análisis de las amenazas de seguridad y el diseño de contramedidas en el estándar 802.11mc, en particular en el proceso de localización basado en FTM. Por motivos de eficiencia, p.ej., con el objetivo de evitar un incremento en el retardo y en el ancho de banda consumido, en su diseño inicial no se consideraron aspectos de seguridad. Como consecuencia, el mecanismo FTM carece de los servicios de seguridad básicos como son la confidencialidad, la autenticación y la integridad de los datos, y por ende es vulnerable a numerosos ataques [11], [12]. En primer lugar, dicho mecanismo no requiere que los APs se autentiquen, lo cual posibilita que un atacante fuerce a una estación a conectarse a un *rogue* AP y le proporcione datos de localización falsos. Por otro lado, todos los mensajes intercambiados entre la estación y el AP se transmiten en claro, hecho que abre una brecha de privacidad puesto que cualquier dispositivo que se encuentre en el rango

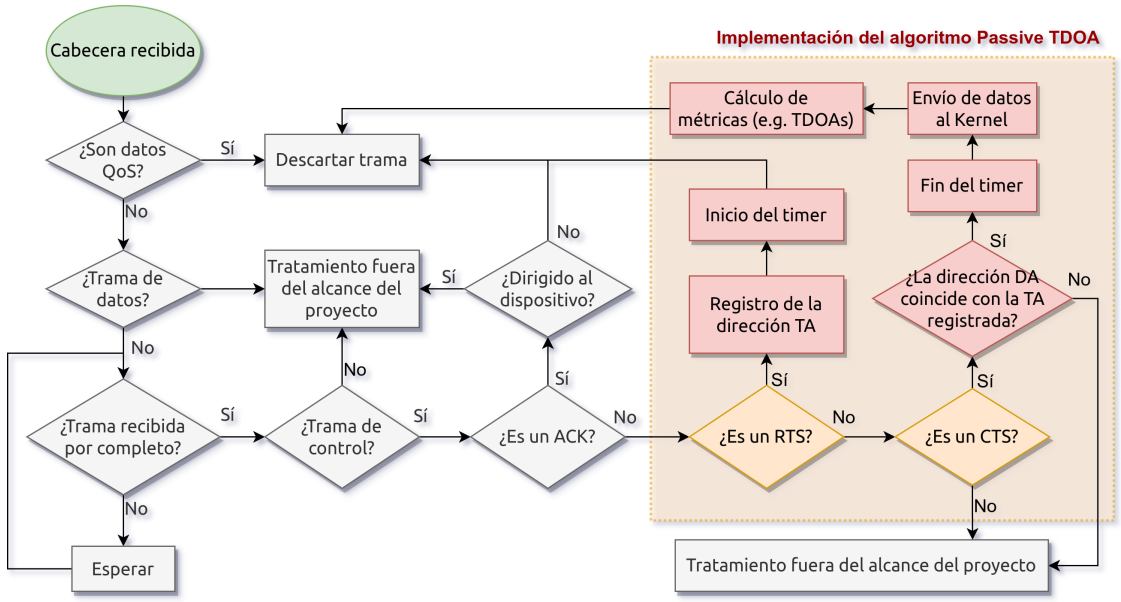


Figura 3. Implementación del algoritmo passive TDOA en la cadena de recepción de tramas de control

de cobertura puede escuchar la conversación y obtener información sensible sobre la estación.

El IEEE 802.11 Amendment for Enhancements for Positioning [13], publicado recientemente, define el denominado Pre-Association Security Negotiation (PASN), un mecanismo que permite que una estación inicie una sesión segura con un AP, protegiendo así el proceso de localización frente a usuarios/dispositivos no autorizados. Sin embargo, la viabilidad de su aplicación está todavía por determinar. A modo de ejemplo, la autenticación de los APs se basa en el uso de una clave compartida (pre-shared key o PSK) y este hecho dificulta su aplicación en redes públicas [12]. El desarrollo de la plataforma presentada aquí permitirá realizar pruebas con dispositivos reales de los aspectos mencionados, para completar los estudios teóricos o simulados presentados hasta ahora.

La adopción de técnicas pasivas como la presentada en [7] permite reducir el tráfico generado durante el proceso de localización y a la vez puede representar una contramedida para evitar los ataques mencionados previamente [14], pero también abre puertas a nuevas amenazas de seguridad que requieren mayor estudio por parte de la comunidad científica.

AGRADECIMIENTOS

Este trabajo ha sido co-financiado por el Ministerio de Ciencia y Educación con el proyecto TCO-RISEBLOCK (PID2019-110224RB-I00), y por la Generalitat de Catalunya con la subvención 2021-SGR-00594.

REFERENCIAS

[1] P. S. Farahsari, A. Farahzadi, J. Rezazadeh, and A. Bagheri, "A survey on indoor positioning systems for iot-based applications," IEEE Internet of Things Journal, vol. 9, no. 10, pp. 7680–7699, 2022.

[2] B. K. P. Horn, "Indoor localization using uncooperative wi-fi access points," Sensors, vol. 22, no. 8, p. 3091, Apr 2022. [Online]. Available: http://dx.doi.org/10.3390/s22083091

[3] S. Hayward, K. van Lopik, C. Hinde, and A. West, "A survey of indoor location technologies, techniques and applications in industry," Internet of Things, vol. 20, p. 100608, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660522000907

[4] P. Roy and C. Chowdhury, "A survey on ubiquitous wifi-based indoor localization system for smartphone users from implementation perspectives," CCF Transactions on Pervasive Computing and Interaction, vol. 4, no. 3, pp. 298–318, Sep 2022. [Online]. Available: https://doi.org/10.1007/s42486-022-00089-3

[5] I. Martin-Escalona and E. Zola, "Improving fingerprint-based positioning by using ieee 802.11 mc ftm/rtt observables," Sensors, vol. 23, no. 1, p. 267, 2023.

[6] 80211standard2016, "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 802.11-2016 (Revision of IEEE Std 802.11-2012), pp. 1–3534, 2016.

[7] I. Martin-Escalona and E. Zola, "Passive round-trip-time positioning in dense ieee 802.11 networks," Electronics, vol. 9, no. 8, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/8/1193

[8] F. Gringoli and L. Nava. (2018) Openfwwf website.

[9] K. Kaichuan He. (2005) Kernel korner - why and how to use netlink socket.

[10] H. Lowe, M. Lamahewage, and K. Gunasekera, "Towards a low-cost wifi based real-time human activity recognition system," in 2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS), 2022, pp. 1–6.

[11] D. Schepers, M. Singh, and A. Ranganathan, "Here, there, and everywhere: Security analysis of wi-fi fine timing measurement," in Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks, ser. WiSec '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 78–89. [Online]. Available: https://doi.org/10.1145/3448300.3467828

[12] D. Schepers and A. Ranganathan, "Privacy-preserving positioning in wi-fi fine timing measurement," Proceedings on Privacy Enhancing Technologies, vol. 2022, no. 2, pp. 325–343, 2022. [Online]. Available: https://doi.org/10.2478/popets-2022-0048

[13] "Ieee draft 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 - amendment 4: Enhancements for positioning," IEEE P802.11az/D4.0, August 2021, pp. 1–282, 2021.

[14] M. Mohsen, H. Rizk, and M. Youssef, "Privacy-preserving by design: Indoor positioning system using wi-fi passive tdoa," 2023.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Detección de fracaso académico en docencia de redes de computadores usando IA

Carneiro V., Cacheda F., Fernández D., López-Vizcaíno M.
Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC),
Universidade da Coruña
Campus de Elviña s/n 15071 A Coruña.
victor.carneiro@udc.es, fidel.cacheda@udc.es, diego.fernandez@udc.es, manuel.fernandezl@udc.es

En este trabajo nos centramos en el desarrollo de técnicas de identificación temprana del fracaso académico en titulaciones de grado universitario, como medio para facilitar a los docentes el desarrollo de acciones proactivas. Partimos de un conjunto de datos de más de mil alumnos, con sus respectivas notas de una materia a lo largo de cuatro años. Se extraen diversas características acerca del esfuerzo y logro alcanzado por cada estudiante correspondientes a las tareas de laboratorio y cuestionarios propuestos, con el objetivo de predecir si un estudiante aprobará el examen final, lo suspenderá o no se presentará. Definimos una tarea de clasificación multiclase de minería de datos siguiendo un enfoque de aprendizaje supervisado donde se evalúa una selección de algoritmos de aprendizaje automático con un esquema de validación cruzada. Nuestros resultados muestran que usando Random Forest podemos predecir con una precisión de más del 86% si un estudiante aprobará el examen final, obteniendo un valor F1 de 0.881. Un análisis de la importancia de las características destaca cómo las prácticas de laboratorio tienen una mayor contribución al modelo de aprendizaje que los cuestionarios.

***Palabras Clave*—Detección temprana, machine learning, validación cruzada, evaluación del aprendizaje**

I. INTRODUCCIÓN

El éxito de los estudiantes se considera una métrica clave en instituciones de educación superior para evaluar su calidad [1] y está ligado estrechamente a la reputación, la posición de las instituciones en los rankings universitarios y consecuentemente la financiación obtenida tanto pública como privada, así como también es clave en los procesos de acreditación y revisión de títulos. El éxito de los estudiantes ha sido definido por York et al. [2] como "rendimiento académico, satisfacción, adquisición de habilidades y competencias, persistencia, consecución de los objetivos de aprendizaje y éxito profesional".

Las distintas actividades de aprendizaje que desarrolla un centro de educación superior genera una ingente cantidad de datos, que en muchos casos no se aprovechan para obtener información y generar conocimiento, en gran medida, por su volumen y heterogeneidad. La aparición de las técnicas de inteligencia artificial y minería de datos, han supuesto en los últimos años una revolución y evolución en el tratamiento de estos datos. De hecho, la Minería de Datos Educativos (EDM) ha surgido como un campo de investigación que involucra estadísticas, minería de datos y aprendizaje automático y otros campos para analizar la gran cantidad de datos generados por la actividad educativa de manera efectiva [3], [4].

En este trabajo nos centramos en la identificación temprana del fracaso académico en la educación superior, que permita a los educadores disponer de herramientas para una intervención proactiva que ayude a los estudiantes en una posición de riesgo a lograr éxito académico. Para lograr este objetivo hacemos uso de un conjunto de datos con más de mil calificaciones de estudiantes en una materia de un grado universitario. Se extraen varias características correspondientes a las tareas de laboratorio y cuestionarios y se propone una tarea de clasificación de minería de datos para predecir si un alumno aprobará o no el examen final.

El resto del artículo está organizado de la siguiente manera. En primer lugar se presentan los trabajos más relevantes del estado del arte. En la sección III se realiza una descripción detallada del curso sobre el que se basa el estudio para, a continuación, presentar el conjunto de datos y las características que se han extraído del mismo, en la sección IV. La sección V describe los modelos de aprendizaje automático que se utilizarán en los experimentos, así como las medidas de evaluación y, en la siguiente sección, se presentan y discuten los resultados experimentales. Finalmente, en la sección VII se presentan las principales conclusiones y los trabajos futuros.

II. TRABAJOS RELACIONADOS

Dentro del ámbito de la Minería de Datos Educativos (EDM) podemos citar múltiples ámbitos de investigación que involucran el descubrimiento de patrones de conocimiento sobre distintos aspectos del proceso de aprendizaje [5], como el rendimiento [6], éxito [7], satisfacción [8] o tasa de abandono [9], entre otros.

Nuestro trabajo está más relacionado con la predicción del rendimiento académico de los estudiantes. En este sentido, los autores de [10] prueban tres algoritmos de clasificación (*Naïve Bayes*, *Neural Network* y *Decision Trees*) para predecir el desempeño de los estudiantes en dos cursos de pregrado, consiguiendo una precisión del 86% con el primero de ellos. El trabajo de Sivasakti [11] aplica diversos algoritmos de clasificación (*Multilayer Perception*, *Naïve Bayes*, *SMO*, *J48* y *REPTree*) para predecir el rendimiento en un curso de programación de computadoras. Asímismo se define un modelo de flujo de conocimiento para los cinco clasificadores y se muestra la importancia de los algoritmos de minería de datos basados en predicción y clasificación en el campo de la educación en programación. En [12] el desempeño de los estudiantes se predice teniendo en cuenta sus antecedentes personales, incluido el género, la beca otorgada, la formación académica previa, tipo de ingreso, talento y provincia de bachillerato, aunque se logró una precisión moderada. En [13] se presenta una comparación simple de diferentes algoritmos de clasificación (*Naïve Bayes*, *Bayesian Network*, *ID3*, *J48* y *Neural Network*) utilizando un conjunto de datos de 225 estudiantes mediante la herramienta WEKA, siendo el primero de ellos el que obtiene mejores resultados de precisión. Otros trabajos, como Yassein et al. [14] hace uso de un conjunto de datos de 150 estudiantes con el objetivo de buscar patrones para predecir rendimiento de los alumnos y descubrir el factor o factores que más afecta en el rendimiento. En este trabajo, se concluye que el factor más importante es la asistencia a clase. Más recientemente, [15] analiza algunos trabajos de investigación publicados entre 2015 y 2021; en concreto se analizan 39 estudios, resaltando seis técnicas de Machine Learning (*Naïve Bayes*, *Decision Tree*, , *Linear Regression*, *Support Vector Machine*, *K-Nearest Neighbor* y *Neural Network*) como las más usadas y concluyendo que el aprendizaje automático puede ser beneficioso para identificar varias áreas de rendimiento académico.

Nuestro trabajo está relacionado con estos trabajos previos en el sentido de que el objetivo es predecir el rendimiento académico, aunque, en nuestro caso, la detección precoz de un potencial bajo rendimiento también es relevante.

III. DESCRIPCIÓN DEL CURSO

Este trabajo se ha realizado recogiendo los datos de una asignatura de Redes impartida en el grado de Ingeniería Informática. Esta materia, de segundo año, se imparte en el segundo semestre y tiene una carga lectiva de 6 *European Credit Transfer System* (ECTS), que corresponden a 60 horas de docencia presencial más 90 horas de trabajo personal. El curso se centra en los aspectos principales de la creación de redes, incluidas las características principales, funcionalidades y estructura de las redes informáticas e Internet. La materia constituye el primer acercamiento a las redes informáticas para la mayoría de los estudiantes y el objetivo principal es que los estudiantes comprendan los diferentes protocolos y capas que entran en acción cuando dos dispositivos se comunican usando TCP/IP. La asignatura tiene asignadas cuatro sesiones por semana (una hora por sesión): dos sesiones teóricas en días diferentes y dos sesiones consecutivas para el laboratorio. El temario del curso es el siguiente:

- Tema I – Introducción a las redes informáticas, Internet y TCP/IP
- Tema II – Capa de aplicación: Web, correo y DNS
- Tema III – Capa de Transporte: UDP y TCP
- Tema IV – Capa de red: IP, división en subredes y enrutamiento
- Tema V – Capa de enlace: ARP, Ethernet y WiFi

A lo largo del curso, cada alumno deberá desarrollar y presentar de forma individual las siguientes prácticas de laboratorio y proyectos, que no son obligatorios:

- Proyecto I: Introducción a la programación de sockets en Java
- Proyecto II: Servidor web Java básico
- Proyecto III: Introducción a la simulación de Redes con Cisco Packet Tracer
- Proyecto IV: Simulación de red – División en subredes y enrutamiento

Además de estos contenidos, los estudiantes resuelven dos cuestionarios en línea a lo largo del curso. Estos cuestionarios se componen de preguntas de las lecciones teóricas y están destinados a reforzar los conocimientos de los estudiantes por aprendizaje continuo. La primera prueba cubre los temas I y II, mientras que la segunda prueba cubre los temas III y IV. La evaluación de la asignatura incluye un examen teórico (hay dos convocatorias disponibles para los alumnos, una al final del semestre y otro aproximadamente un mes después) que corresponde al 70% de la calificación final. Los estudiantes deben alcanzar, en este examen, al menos una calificación de 4 (sobre 10) para el cálculo de la calificación final, que también incluye las notas de laboratorio y de cuestionarios, con un 25% y un 5%, respectivamente (sin que se requiera nota mínima en este caso). Para que un alumno apruebe la asignatura se requiere una nota final mayor o igual a 5.

IV. DATASET Y CARACTERÍSTICAS

En esta sección presentamos el contenido y características del dataset utilizado para el modelado y experimentación, con algoritmos de machine learning, en la detección temprana de estudiantes en riesgo de fracaso académico.

A. DATASET

Hemos creado un conjunto de datos que recopila las calificaciones de la asignatura Redes, presentada anteri-

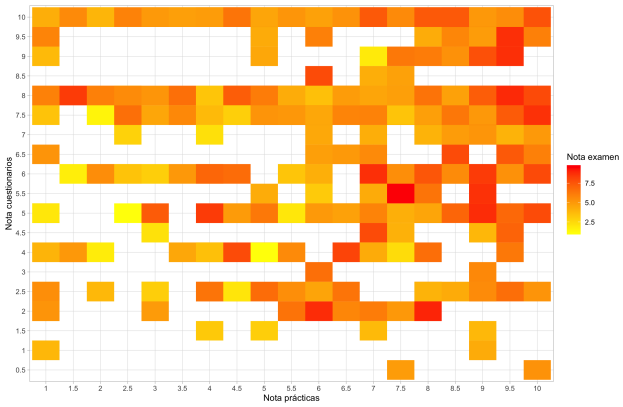


Fig. 1. Mapa de calor de notas de examen con respecto a las notas de cuestionarios y laboratorio.

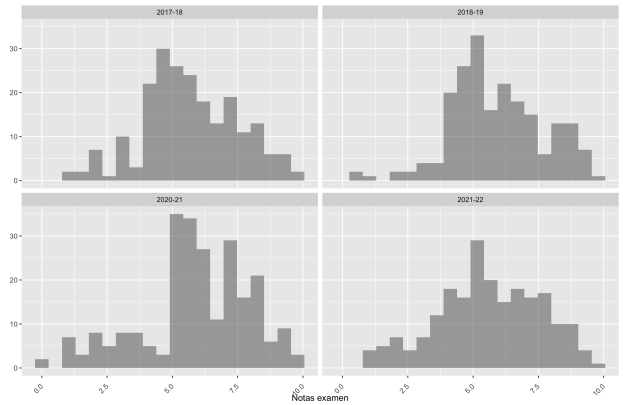


Fig. 2. Histogramas de las notas de examen para los cursos 2017-18, 2018-19, 2020-21 y 2021-22.

ormente, a lo largo de cuatro cursos académicos: 2017-18, 2018-19, 2020-21 y 2021-22. No incluimos el curso 2019-20 debido a que la pandemia SARS-COVID19 produjo cambios muy significativos en la evaluación de los estudiantes y que, sin duda, alterarían artificialmente los resultados del estudio.

La tabla I presenta un resumen de las principales características del conjunto de datos. Está compuesto por más de mil alumnos con sus respectivas notas. Resumimos el número de alumnos que aprobaron y suspendieron cada una de las partes principales de la evaluación (examen, laboratorio y cuestionarios). Para este estudio, consideramos que un estudiante aprobó el examen si en cualquiera de las dos convocatorias obtuvo una calificación mayor o igual a 4. Asimismo, el número de alumnos que suspendieron el examen incluye a los alumnos que no se presentaron. A priori, podemos suponer que los estudiantes que han superado las pruebas durante el curso (laboratorio y cuestionarios), también superarán el examen final. Otro dato, que se puede intuir, es que el peso de la nota de laboratorio va a tener bastante importancia, ya que el número de suspensos en esta tarea es superior al del propio examen.

La Figura 1 presenta un mapa de calor de las calificaciones del examen con respecto a las puntuaciones del laboratorio (eje X) y del cuestionario (eje Y). De la figura podemos observar cómo los tonos más oscuros (correspondientes a calificaciones más altas en los exámenes) se ubican en la mitad derecha de la figura y, especialmente, en la esquina superior correspondiente a calificaciones más altas tanto en tareas de laboratorio como de cuestionarios, confirmando nuestra hipótesis.

La Figura 2 muestra un histograma para las notas de los exámenes obtenidas en cada uno de los cursos académicos del conjunto de datos. En general, el comportamiento de los alumnos en los cuatro cursos académicos es similar, apareciendo un pico en la parte central del histograma, que se corresponde con una nota de 5.

Por otra parte, también es interesante analizar el comportamiento de los alumnos que no se presentan al examen. Como se puede observar en la Figura 3, los alumnos no presentados tienden a obtener puntuaciones inferiores a lo largo del curso en las tareas de laboratorio y en los

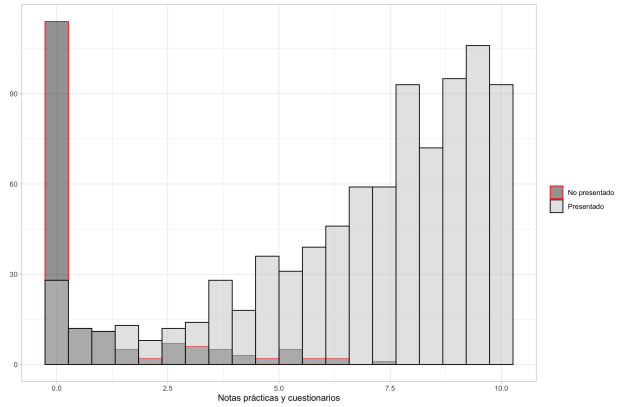


Fig. 3. Histogramas de las notas de examen para los cursos 2017-18, 2018-19, 2020-21 y 2021-22.

cuestionarios. Este comportamiento es previsible, debido al modelo de evaluación que proporciona un peso relevante al trabajo del alumno a lo largo de todo el curso, siguiendo el Plan Bolonia, tanto en el laboratorio como en los cuestionarios. De esta manera, a un alumno que no haya trabajado durante el curso, le resultará complicado superar la materia únicamente realizando el examen final y, por lo tanto, probablemente optará por no presentarse.

B. CARACTERÍSTICAS

Las características extraídas del conjunto de datos se dividen en dos grupos, dependiendo de si se corresponden con calificaciones de laboratorio o de examen. Todas las notas están normalizadas entre 0 y 1. Del conjunto de datos, extraemos varias características que se dividen en dos grupos, dependiendo de si corresponden a calificaciones de laboratorio o cuestionarios. Todas las puntuaciones están normalizadas para operar entre 0 y 1.

Las características del laboratorio son las siguientes:

- Puntuaciones de tareas de laboratorio (indicadas como Lab1, Lab2, Lab3 y Lab4).
- Laboratorio aprobado (Lab_passed), valor booleano que indica si el alumno aprobó las tareas de laboratorio (es decir, puntuación del laboratorio superior o igual a 5).
- Esfuerzo de laboratorio (Lab_effort), porcentaje de trabajos entregados.

	Estudiantes	Laboratorio		Cuestionarios		Exámenes		
		Apto	No Apto	Apto	No Apto	Apto	No Apto	NP
2021-22	256	134	122	177	79	169	44	43
2020-21	289	170	119	225	64	197	43	49
2018-19	244	169	75	159	85	188	17	39
2017-18	261	181	80	175	86	189	26	46
Total	**1050**	**654**	**396**	**736**	**314**	**743**	**130**	**177**

Tabla I
Resumen del dataset. La columna NP hace referencia al número de estudiantes que no se han presentado al examen.

- Promedio, desviación típica y mediana de las puntuaciones de los cuestionarios (Lab_avg, Lab_std y Lab_median).
- Tareas de laboratorio aprobadas (N_Lab_passed).
- Tareas de laboratorio enviadas (N_Lab_tried).

Las características seleccionadas para los cuestionarios son:

- Puntuaciones de los cuestionarios (indicadas como Quiz1 y Quiz2).
- Cuestionario aprobado (Quiz_passed), valor booleano que indica si el alumno aprobó los cuestionarios (es decir, puntuación de los cuestionarios superior o igual a 5).
- Esfuerzo de cuestionarios (Quiz_effort), porcentaje de cuestionarios entregados.
- Promedio, desviación típica y mediana de las puntuaciones de las tareas de laboratorio (Quiz_avg, Quiz_std y Quiz_median).
- Tareas de laboratorio aprobadas (N_quiz_passed).
- Tareas de laboratorio enviadas (N_quiz_tried).

Además, se calculó el promedio de la agregación de las puntuaciones del laboratorio y cuestionarios (denotado como LabQuiz_score).

V. MODELADO

Como hemos indicado anteriormente, el objetivo de este estudio es predecir si un estudiante aprobará el examen final, tomando en consideración el trabajo realizado a lo largo del curso, en concreto, su actividad y puntuaciones en los proyectos de laboratorio y en los cuestionarios realizados. Para llevar a cabo esta labor, hemos definido una tarea de clasificación multiclase de minería de datos siguiendo técnicas de aprendizaje supervisado. El objetivo es predecir, usando las características proporcionadas por las prácticas de laboratorio y cuestionarios, si un alumno aprobará el examen o, por el contrario, no lo aprobará o no se presentará.

En base a los resultados previos observados en la literatura, hemos seleccionado los siguientes algoritmos de aprendizaje automático: *J48*, *JRip*, *LibLinear*, *Logistic Regression (LR)*, *Multilayer Perceptron (MLP)*, *Naïve Bayes (NB)*, *Random Forest (RF)*, *Random Tree (RT)*, y *Support Vector Machine (SVM)*. Estos algoritmos cubren las principales técnicas utilizadas en Minería de Datos Educativos (EDP).

La evaluación se realiza siguiendo un esquema de validación cruzada de 10 iteraciones para validar el rendimiento y robustez de los modelos. Para abordar el desequilibrio de clase (743 estudiantes aprobaron el examen frente a los 130 que suspendieron y los 177 que no se presentaron) sobremuestreamos las clases minoritarias utilizando la técnica de sobremuestreo sintético de minorías (SMOTE). Esta técnica evita el sobreajuste del modelo, al crear nuevos ejemplos sintéticos mediante la elección aleatoria de las k clases minoritarias en lugar de utilizar sobremuestreo por sustitución.

Como métricas de evaluación, usamos las siguientes:

- Precision: porcentaje de aciertos en la predicción frente a todo el conjunto de predicciones.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

- F-score (F1): nos permite combinar la precisión y la exhaustividad (recall)

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{3}$$

- Raiz del Error cuadrático medio (RMSE): Raíz del error cuadrado promedio de nuestras predicciones.

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(\frac{d_i - f_i}{\sigma_i}\right)^2} \tag{4}$$

- Área bajo la curva ROC: El algoritmo AUC permite calcular la curva de característica operativa del receptor (ROC) para representar la tasa de verdaderos positivos frente a la de falsos positivos en diferentes umbrales de clasificación. Indica la probabilidad de que el modelo clasifique un ejemplo positivo aleatorio más alto que un ejemplo negativo aleatorio.
- Área bajo la Precision Recall Curve (PRC): muestra el compromiso entre precisión y exhaustividad (recall) para diferentes umbrales.

VI. RESULTADOS EXPERIMENTALES

En la tabla II, presentamos los resultados de los modelos entrenados con el conjunto de datos propuesto. Random Forest (RF) es el modelo con mejor desempeño, siendo capaz de predecir con precisión más del 86% de los casos y el F-score es 0,881. Además, para todas las métricas restantes, Random Forest (RF) es consistentemente el modelo con mejor rendimiento.

Realizamos un estudio de población, repitiendo la evaluación considerando únicamente las características de laboratorio y cuestionarios. En ambos casos, los resultados no

Modelo	Correctos	F1	ROC AUC	PRC AUC	RMSE
J48	80.78%	0.830	0.900	0.805	0.333
JRip	78.91%	0.828	0.914	0.848	0.337
LibLinear	74.01%	0.796	0.846	0.704	0.416
LR	74.05%	0.793	0.921	0.866	0.353
MLP	78.33%	0.813	0.913	0.853	0.346
NB	65.71%	*0.753*	0.907	0.842	0.467
RF	**86.45%**	**0.881**	**0.962**	**0.943**	**0.272**
RT	79.67%	0.809	*0.835*	0.729	0.362
SVM	*72.27%*	0.786	0.844	*0.678*	*0.430*

Tabla II
Resultados usando todas las características.

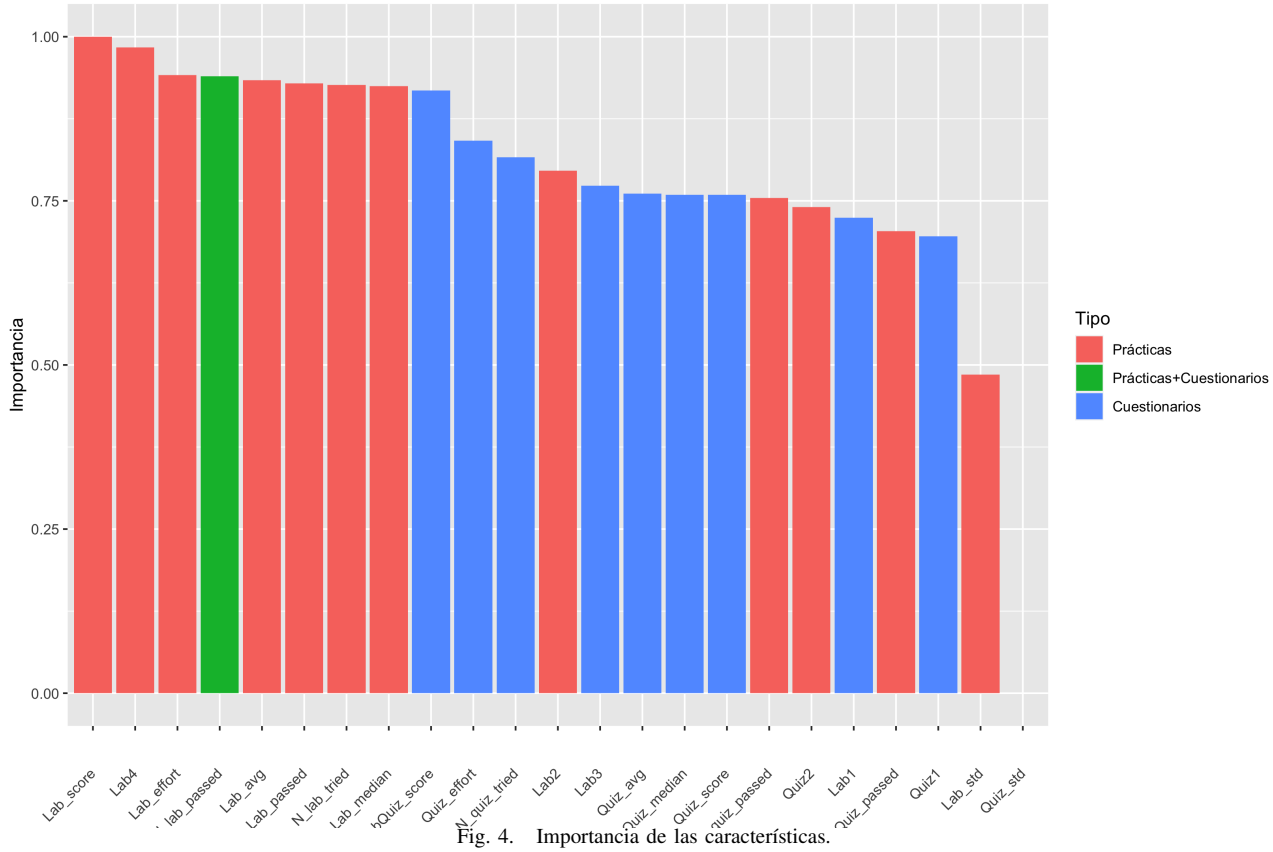


Fig. 4. Importancia de las características.

mejoraron el modelo de mejor desempeño de la tabla II. En términos generales, usar solo características de laboratorio logró mejores resultados que usar solo características de cuestionarios. Este resultado es el esperado, ya que los trabajos de laboratorio deben ser desarrollados individualmente por cada estudiante, mientras las tareas de los cuestionarios se pueden responder de manera colaborativa y, por lo tanto, es posible que no reflejen con precisión el comportamiento del estudiante, así como el esfuerzo y conocimiento en el tema.

Finalmente, analizamos la importancia de las características en la Figura 4 midiendo la correlación de Pearson entre cada característica y la clase (laboratorio, cuestionario o examen). Se aplicó la normalización min-max a los valores de correlación de Pearson obtenidos. Las funciones de laboratorio se representan en rojo, mientras que las funciones de cuestionarios se muestran en azul. La característica de agregación de ambos valores se muestra en verde. Podemos observar cómo las características del laboratorio son más importantes para la tarea de clasificación que las características del cuestionario. También es interesante la alta posición en el ranking de la función de agregación LabQuiz_score.

VII. CONCLUSIONES

Los indicadores de tasa de éxito tienen un fuerte impacto en los procesos de acreditación de títulos, acreditaciones de figuras de profesorado, evaluación de la actividad docente, de la reputación institucional y principalmente, de la calidad docente. Uno de las principales objetivos de los docentes consiste en mejorar estas tasas y, debido al gran volumen de datos generados por la actividad docente, el uso de técnicas y herramientas automatizadas para detectar colectivos en riesgo de fracaso académico son fundamentales actualmente. En este sentido, el uso de técnicas de inteligencia artificial y tratamiento de grandes

volúmenes de datos están cobrando gran importancia en los últimos años, como es el caso de la Minería de Datos Educativos (EDM).

En este trabajo hemos mostrado cómo siguiendo un enfoque de aprendizaje supervisado y usando sólo información extraída de las calificaciones obtenidas en tareas de laboratorio y cuestionarios, somos capaces de predecir correctamente si un alumno aprobará o suspenderá el examen final o no se presentará en más del 86% de los casos. Además, nuestro análisis muestra cómo las características de pruebas de laboratorio tienen una mayor contribución al modelo de aprendizaje que las tareas de cuestionario.

Este trabajo nos ha permitido, en un primer momento, identificar qué características son las que más influyen en la supervisión del fracaso académico. Esto nos ha llevado a modificar las pruebas prácticas, para obtener de forma más fiable estos indicadores y, de este modo, mejorar la precisión del método de predicción. Con los resultados obtenidos, en las primeras pruebas del presente curso académico, se han establecido acciones formativas de refuerzo y tutorías grupales orientadas a este colectivo que, de forma opcional, pueden llevar a cabo los estudiantes con riesgo de fracaso académico. Al finalizar el presente curso, se analizarán los resultados para determinar si se cumple el objetivo de mejorar la tasa de éxito de la materia.

AGRADECIMIENTOS

Esta investigación ha sido financiada por el Ministerio de Economía y Competitividad de España y fondos FEDER de la Unión Europea (Proyecto PID2019-111388GB-I00) y por el Centro de Investigación de Galicia "CITIC", financiado por la Xunta de Galicia y la Unión Europea (European Fondo de Desarrollo Regional - Programa Galicia 2014-2020), mediante subvención ED431G 2019/01.

REFERENCIAS

[1] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: literature review and best practices," *International Journal of Educational Technology in Higher Education*, vol. 17, pp. 1–21, 2020.

[2] T. T. York, C. Gibson, and S. Rankin, "Defining and measuring academic success," *Practical assessment, research, and evaluation*, vol. 20, no. 1, p. 5, 2015.

[3] W. Xiao, P. Ji, and J. Hu, "A survey on educational data mining methods used for predicting students' performance," *Engineering Reports*, vol. 4, no. 5, p. e12482, 2022.

[4] S. Batool, J. Rashid, M. W. Nisar, J. Kim, H.-Y. Kwon, and A. Hussain, "Educational data mining to predict students' academic performance: A survey study," *Education and Information Technologies*, pp. 1–67, 2022.

[5] M. Anoopkumar and A. M. Z. Rahman, "A review on data mining techniques and factors used in educational data mining to predict student amelioration," in *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*. IEEE, 2016, pp. 122–133.

[6] A. A. Saa, "Educational data mining & students' performance prediction," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 5, 2016.

[7] M. P. Martins, V. L. Miguéis, D. Fonseca, and A. Alves, "A data mining approach for predicting academic success–a case study," in *Information Technology and Systems: Proceedings of ICITS 2019*. Springer, 2019, pp. 45–56.

[8] E. Alqurashi, "Predicting student satisfaction and perceived learning within online learning environments," *Distance Education*, vol. 40, no. 1, pp. 133 – 148, 2019.

[9] B. Pérez, C. Castellanos, and D. Correal, "Predicting student drop-out rates using data mining techniques: A case study," in *Applications of Computational Intelligence: First IEEE Colombian Conference, ColCACI 2018, Medellín, Colombia, May 16-18, 2018, Revised Selected Papers 1*. Springer, 2018, pp. 111–125.

[10] A. Mueen, B. Zafar, and U. Manzoor, "Modeling and predicting students' academic performance using data mining techniques," *International Journal of Modern Education and Computer Science*, vol. 8, no. 11, p. 36, 2016.

[11] M. Sivasakthi, "Classification and prediction based data mining algorithms to predict students' introductory programming performance," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*. IEEE, 2017, pp. 346–350.

[12] N. Putpuek, N. Rojanaprasert, K. Atchariyachanvanich, and T. Thamrongthanyawong, "Comparative study of prediction models for final gpa score: a case study of rajabhat rajanagarindra university," in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*. IEEE, 2018, pp. 92–97.

[13] H. Almarabeh, "Analysis of students' performance by using different data mining classifiers," *International Journal of Modern Education and Computer Science*, vol. 9, no. 8, p. 9, 2017.

[14] N. A. Yassein, R. G. M. Helali, S. B. Mohomad *et al.*, "Predicting student academic performance in ksa using data mining techniques," *Journal of Information Technology & Software Engineering*, vol. 7, no. 5, pp. 1–5, 2017.

[15] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and evaluation of different machine learning algorithms for predicting student performance," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Plataforma escalable para mejorar la adquisición de competencias y el proceso de evaluación en el ámbito de la Ciberseguridad

Julia Sánchez, Lluís Camino, Jaume Campeny, Guiomar Corral.
Departamento de Ingeniería,
La Salle Campus Barcelona – Universidad Ramon Llull (URL)
Sant Joan de la Salle 42, 08022, Barcelona.
j.sanchez@salle.url.edu, lluis.camino@students.salle.url.edu, jaume@campeny.net, guiomar.corral@salle.url.edu.

La escasez de expertos en ciberseguridad plantea la necesidad de ofrecer formación práctica en este campo. Los Cyber Ranges se presentan como una solución prometedora para brindar un entorno de aprendizaje realista y seguro. Sin embargo, dada la existencia de numerosas plataformas de Cyber Ranges, es importante establecer una taxonomía específica para evaluar su idoneidad en el ámbito de la formación en ciberseguridad. De entre centenares de Cyber Ranges, *KYPO Cyber Range Platform (KYPO CRP)* se posiciona como una opción viable según la taxonomía propuesta. No obstante, KYPO se puede implementar sobre diferentes arquitecturas de *OpenStack*, lo que requiere pruebas específicas, análisis y comparaciones para determinar cuál es más escalable y estable. El artículo también presenta la evaluación de dos arquitecturas de *OpenStack* y recomienda la más adecuada para aprovechar al máximo las capacidades de KYPO CRP.

***Palabras Clave*- jitel, telemática, ciberseguridad, cyber ranges, virtualización, cloud, formación**

I. Introducción

En el mundo digital actual, la ciberseguridad se ha vuelto de vital importancia debido a la creciente dependencia de las empresas y las personas en la tecnología para llevar a cabo sus operaciones diarias y comunicaciones. No obstante, la demanda de medidas de ciberseguridad suele superar la disponibilidad de expertos capacitados para implementar y mantener dichas medidas. Esta situación puede generar vulnerabilidades y brechas que tienen graves consecuencias tanto a nivel individual como para las empresas y las naciones.

La participación de La Salle Campus BCN - URL en el proyecto REWIRE [1] ha permitido observar que Europa se enfrenta a una escasez significativa de expertos en ciberseguridad, ya que la demanda de estas habilidades ha experimentado un rápido crecimiento en los últimos años. Esta situación ha dado lugar a una brecha de habilidades en el campo, agravada por el hecho de que la ciberseguridad es un área compleja, multidisciplinaria y en constante evolución que requiere una formación y educación continuas. Descubrir, abordar y actualizar las habilidades de seguridad necesarias no es una tarea sencilla. Para satisfacer estas necesidades, se requiere un entrenamiento completo y competente que abarque tanto a los recién llegados al sector como a aquellos perfiles seniors que necesitan actualizar sus conocimientos. Existen diversas formas de transmitir conocimientos y diferentes enfoques de aprendizaje, desde los métodos tradicionales basados en conferencias hasta la aplicación de metodologías activas en las que los estudiantes participan activamente en su proceso de aprendizaje. Recientemente, ha habido un notable incremento en la utilización de los Cyber Ranges, entre los diversos métodos disponibles.

Durante muchos años, La Salle Campus BCN, comprometida con ayudar a disminuir la escasez de profesionales en ciberseguridad, ha brindado formación en este campo siguiendo una metodología centrada en la adquisición de competencias [2]. Sin embargo, la evaluación y la analítica del aprendizaje son áreas que se pueden mejorar, y gracias a la incorporación de los Cyber Ranges junto con otras metodologías activas aplicadas, se facilita abordar este desafío de manera más efectiva.

El artículo presenta una introducción a las metodologías para la formación en ciberseguridad con el objetivo de demostrar cuál se adapta mejor a la enseñanza práctica de las competencias de ciberseguridad. Posteriormente, define las características y funcionalidades a nivel general de los Cyber Ranges que darán solución a esta enseñanza práctica y plantea una taxonomía de Cyber Range aplicable al ámbito de la formación en ciberseguridad. Se utiliza esta taxonomía para buscar, entre los cientos de Cyber Ranges existentes en el mercado, cuál se adapta mejor a la misma y, finalmente, se explican las características, tipos de

instalación y posibilidades de KYPO CRP como propuesta de CR para complementar la formación de expertos en ciberseguridad en la universidad. Se concluye el artículo resumiendo los puntos clave discutidos y exponiendo posibles líneas futuras de investigación.

II. Metodologías para la formación en Ciberseguridad

Existen varias metodologías efectivas para enseñar ciberseguridad, dependiendo de la audiencia, el nivel de experiencia técnica requerida y los objetivos de formación. A continuación, se presentan algunos ejemplos:

- **Entrenamiento práctico** (*Hands-on*). Basado en ejercicios y simulaciones para que los estudiantes apliquen sus conocimientos en situaciones del mundo real. Útil para enseñar habilidades técnicas, como seguridad de redes, evaluación de vulnerabilidades y pruebas de penetración.
- **Aprendizaje basado en escenarios**. Implica presentar situaciones reales que requieren la aplicación de conceptos y técnicas de ciberseguridad. Ayuda a desarrollar habilidades de pensamiento crítico y prepara a los estudiantes para manejar incidentes de seguridad reales.
- **Gamificación**. Convierte la formación en ciberseguridad en un juego, lo que hace que el aprendizaje sea más atractivo e interactivo. Especialmente efectiva para enseñar ciberseguridad a empleados no técnicos, quienes pueden encontrar los métodos de formación tradicionales aburridos o intimidantes [3]-[5].
- **Aprendizaje mixto**. Combina diferentes tipos de formación, como cursos en línea, clases en el aula y ejercicios prácticos, para ofrecer una experiencia de aprendizaje completa. Esta metodología se adapta a diferentes estilos de aprendizaje y brinda flexibilidad a los estudiantes con horarios ocupados.
- **Programas de certificación**. Preparan a los estudiantes para exámenes estándar de la industria, como *CompTIA Security+* o *Certified Information Systems Security Professional* (CISSP). Ofrecen una forma concreta para que los estudiantes demuestren su experiencia y mejoren sus oportunidades laborales.

La metodología más efectiva para enseñar ciberseguridad dependerá de las necesidades y objetivos específicos de los estudiantes y de la organización que brinda la formación. El enfoque en La Salle Campus BCN para formar a los estudiantes en ciberseguridad se asemeja al método de *Aprendizaje Mixto*, incorporando: (1) clases en formato *face-to-face* que crean un vínculo esencial con el estudiante, (2) ejercicios prácticos presenciales en modo guiado y tipo *challenge* para poner a prueba los conocimientos adquiridos y, (3) recursos en línea accesibles a través de *Virtual Learning Environments* (VLE), como *Moodle*, que combinan teoría y práctica para complementar el aprendizaje. El aspecto práctico puede abordarse de diversas formas, según el curso y las necesidades y objetivos de capacitación para obtener los resultados de aprendizaje marcados. El aspecto fundamental, especialmente para la formación de habilidades más técnicas, es cómo respaldar este aprendizaje práctico. Para que la práctica realizada sea efectiva y ayude a afianzar los conocimientos en ciberseguridad, es importante tener una base teórica sólida para comprender los conceptos y solucionar problemas. Por lo tanto, es fundamental combinar teoría y práctica para adquirir los conocimientos necesarios, independientemente de la metodología elegida.

Actualmente, existen varios métodos útiles para complementar la formación práctica en ciberseguridad, cada uno de ellos con sus ventajas y desventajas. A continuación, se destacan aquellos seleccionados como relevantes para este estudio.

- **Videotutoriales**. Permiten observar paso a paso un ataque o vulnerabilidad, pero no ofrecen interacción ni soporte. Una simple búsqueda de cualquier ataque o vulnerabilidad en Youtube genera cientos de resultados con explicaciones paso a paso detalladas.
- **Páginas web**. Ofrecen escenarios predefinidos para poner a prueba habilidades, pero pueden no estar actualizadas o carecer de una explicación teórica detallada. Algunos ejemplos: *HackThis* [6], *Google Gruyere* [7], *bWAPP* [8] *y HackThisSite* [9].
- **Plataformas o Testbeds como LOST [10]**. Proporcionan entornos seguros y controlados para aplicar conocimientos de ciberseguridad, pero no recopilan información sobre el rendimiento del estudiante. En concreto, LOST es la plataforma utilizada actualmente en La Salle Campus BCN para enseñar hacking ético.
- **Cyber Ranges (CRs).** Son entornos virtuales que simulan situaciones reales y permiten el entrenamiento y la validación de conocimientos, adaptándose a las necesidades actuales de la ciberseguridad.

En la Tabla I se presenta una comparativa de estos métodos destacando el modo de crear los escenarios para practicar (*Escenarios*), si usualmente incorpora la enseñanza de teoría y/o práctica (*Educación*) y, si dispone de capacidades de seguimiento o retroalimentación las cuales son importantes para guiar a los estudiantes en su aprendizaje (*Seguimiento*).

Los videotutoriales y las páginas web de aprendizaje no aportan las características necesarias para ser un buen complemento en la enseñanza de ciberseguridad debido a la falta de énfasis en la teoría y la ausencia de un entorno seguro. Las plataformas de aprendizaje ofrecen escenarios predefinidos, pero carecen de información para el formador sobre el progreso del alumno. En contraste, los CRs se adaptan a las últimas tecnologías, ofrecen entornos seguros y brindan datos valiosos para el seguimiento del alumno. Además, son plataformas flexibles que permiten la evolución tanto del alumno como del profesor en el campo de la ciberseguridad.

Tabla I
COMPARATIVA DE MÉTODOS DE FORMACIÓN EN CIBERSEGURIDAD

Método	Escenarios	Educación	Seguimiento
Videotu-toriales	Alumno crea su escenario. Tópicos marcados por los vídeos.	Centrado en la parte práctica y cómo realizarla.	Sin herramientas de seguimiento del progreso del alumno.
Páginas web	Escenario ya está creado. Tópicos marcados por la página web.	Centrado en la parte práctica y cómo realizarla.	Sin herramientas de seguimiento del progreso del alumno.
Plataformas (LOST)	Escenario ya está creado. Ofrece cierta libertad para elegir tópicos a enseñar.	Combina teoría y práctica.	Sin herramientas de seguimiento del progreso del alumno.
Cyber Ranges (CRs)	Escenario ya está creado. Ofrece libertad absoluta en la creación de escenarios para trabajar el tópico que se desee.	Combina teoría y práctica.	Puede mostrar resumen de ejecución y resultados a formadores y estudiantes.

III. CYBER RANGES

Se escoge el uso de CRs para complementar la formación práctica de los estudiantes. La proliferación del uso de CRs en los últimos años, ha generado cierta confusión o disparidad en su definición. En esta sección se presenta una definición que se adapta a cualquier plataforma de CR actual. También se introducen las características deseables en un CR enfocado a la formación de habilidades en ciberseguridad.

A. Definición de un Cyber Range

El concepto de un CR puede resultar difícil de comprender, ya que muchas fuentes lo definen de manera diferente, dependiendo del propósito de la organización/institución que lo usará o el ámbito de aplicación. En algunos casos, este concepto está relacionado únicamente con el desarrollo profesional, mientras que otros hacen más hincapié en su aspecto educativo. Después de evaluar varias definiciones tanto de agencias gubernamentales como de empresas del sector [11]-[19], es importante extraer las características y capacidades básicas comunes entre todas estas definiciones para formar una idea clara.

La Tabla II resume las principales características extraídas de la lectura realizada. Las conclusiones de esta tabla ayudarán a formar una definición general del concepto de Cyber Range. Aunque existen muchas otras características, la tabla incluye aquellas que son comunes en la mayoría de las fuentes y también aquellas que tienen diferentes puntos de vista dependiendo de la fuente.

Con las características principales extraídas, se puede definir un CR como: *Un Cyber Range es un entorno donde se puede llevar a cabo la capacitación en operaciones situacionales, pruebas, investigación y desarrollo educativo. Por lo tanto, el alcance de los Cyber Ranges no se limita solo a organizaciones y profesionales, sino también a estudiantes y entidades educativas. La tecnología utilizada para crear estos entornos es amplia y puede ser hardware, software o una combinación de ambos. Este entorno es cerrado y sin riesgos, lo que permite la realización de escenarios de la vida real. Adquirir habilidades prácticas, probar servicios o productos y realizar pruebas de seguridad son los principales casos de uso de los Cyber Ranges.*

Tabla II
CARACTERÍSTICAS PRINCIPALES DE UN CYBER RANGE

Característica	Descripción
Tipo de entorno	· Controlado · Sin riesgos · Legal
Objetivo / Audiencia	· Profesionales de seguridad · Empresas y personas que toman decisiones estratégicas (privadas y gubernamentales) · Agencias gubernamentales y militares · Academia (educadores, estudiantes e investigadores)
Casos de uso	· Pruebas de seguridad · Investigación en seguridad · Desarrollo de competencias · Educación en seguridad · Desarrollo de capacidades cibernéticas · Desarrollo de resiliencia cibernética · Evaluación de competencias · Contratación · Destreza digital · Competencias de ciberseguridad nacionales e internacionales
Componentes	· Hardware · Software · Virtual
Tecnologías	· Basado en hardware · Basado en virtualización · Basado en la nube · Combinación de los anteriores

B. Cyber Ranges para la formación en ciberseguridad

Como se ha mostrado en el apartado anterior, existen múltiples CRs que pueden utilizarse para diferentes propósitos. De la lectura realizada para entender los CRs, se concluye que el desarrollo de uno propio para usarlo con el propósito de capacitar a estudiantes en ciberseguridad, está fuera del alcance de la universidad (debido al alto coste, la necesidad de expertos en multitud de tecnologías y la gran cantidad de tiempo necesario para tenerlo operativo). Por consiguiente, se centran los esfuerzos en encontrar una plataforma de CR viable para la educación/formación y la evaluación de competencias.

Para cumplir este objetivo, es interesante plantear una taxonomía de CR que englobe ciertas funcionalidades y características útiles en el ámbito de la formación. Resultados como los presentados en [20] y [21] proponen taxonomías de CR basadas en el estudio de una gran cantidad de CRs presentes en el mercado. Se ha utilizado el conocimiento recopilado de las taxonomías anteriores para formar una nueva (Fig. 1), compuesta por cuatro bloques principales, que se ajuste mejor a las necesidades de formación.

El bloque de **Entorno** es el bloque principal y funcional del CR, determinando su estructura y las posibles funcionalidades y capacidades que puede tener. Se distinguen cuatro *Tipos* según si está basado en hardware, software, simulación/emulación o una combinación entre éstos. También es importante dotar de contexto a los *Usuarios* (existirán diferentes tipos según

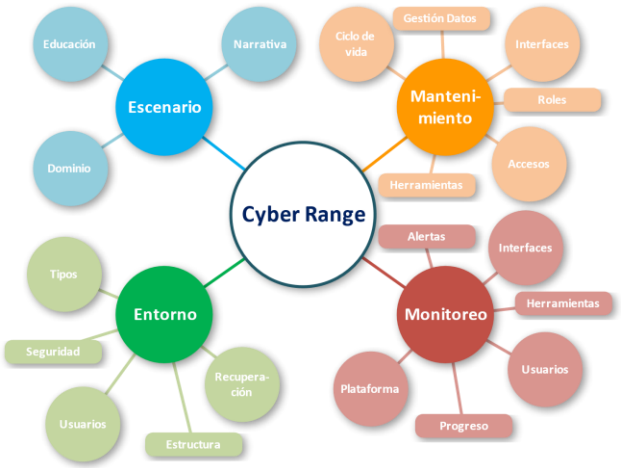


Fig. 1. Taxonomía Cyber Range Educación

los privilegios de acceso a la plataforma, las acciones que se realizarán sobre ella y la interacción con los escenarios/ejercicios ejecutados). Otros aspectos adicionales son la seguridad interna y externa para proteger la plataforma y los usuarios de posibles ataques (opción de uso de técnicas de microsegmentación y políticas *zero-trust*). Finalmente, la *Recuperación* frente a ataques y funcionamientos erróneos, para la cual se dispone de componentes y/o herramientas que permitan mejorar la disponibilidad, operabilidad y seguridad de la plataforma.

El bloque de **Escenario** permite realizar actividades y tareas eminentemente prácticas para la posterior evaluación de los alumnos en función de su rendimiento y desempeño en ellas. Para la *Educación* de los estudiantes presenta posibles tareas que pueden incluir tutoriales, demostraciones o ejercicios/retos (respaldadas mediante textos y otros elementos multimedia), contemplando una evaluación posterior del rendimiento e impacto de cada alumno. Asimismo, es importante identificar los *Dominios* en los que se distribuirán los escenarios al definir las capacidades y alcances de éstos para permitir una integración adecuada de tecnologías actuales y futuras (para que los alumnos puedan experimentar con todas ellas). Algunos ejemplos podrían ser IoT (*Internet of Things*) o *Cloud*. Por otro lado, cada escenario debe presentar una *Narrativa* clara y comprensible para transmitir el propósito de la actividad a realizar de la forma más precisa y fácil de entender posible, evitando malentendidos durante la realización de las actividades. Por último, es imprescindible contar con herramientas que faciliten la ejecución de todas las actividades y metas establecidas, las cuales son necesarias para realizar y desarrollar las funcionalidades y capacidades del bloque.

El bloque de **Monitoreo** del CR tiene como objetivo controlar y visualizar las actividades en la plataforma, así como el impacto de las acciones de los usuarios. Se monitorea la *Plataforma* para asegurar su funcionamiento y disponibilidad, detectar incidencias y notificarlas a los usuarios (estado de la plataforma, rendimiento y conexiones realizadas). También se registra la información de las acciones y datos ingresados por los *Usuarios*, así como el progreso y rendimiento de los alumnos. El monitoreo se presenta de manera gráfica y comprensible, con *dashboards* e informes automáticos (*Interfaces*) e, incluso, puede proporcionar funcionalidades adicionales para incorporar protocolos de monitoreo que permitan extraer los datos para ser evaluados (como SNMP). Las herramientas utilizadas son fundamentales para llevar a cabo estas actividades y funcionalidades.

El bloque de **Mantenimiento** permite gestionar y comunicarse con la plataforma para realizar los cambios necesarios para su correcto funcionamiento (según usuarios y privilegios, las acciones permitidas varían). Incluye la gestión de escenarios o *Ciclo de Vida* (pudiendo crear entornos, editarlos, desplegarlos, ejecutarlos y eliminarlos), la gestión de *Accesos* a la plataforma para controlarlos y detectar incidentes (p.ej. accesos remotos, locales o por VPN), la gestión de datos para el correcto funcionamiento de la plataforma y los roles de usuarios para controlar las acciones de éstos según privilegios asignados (bloque Entorno). Se busca simplificar las operaciones mediante aportes gráficos y visuales (*Interfaces*), mejorando la experiencia de usuario y pudiendo utilizar la información del Monitoreo para realizar las tareas de mantenimiento. Como en el bloque anterior, las herramientas utilizadas son fundamentales para realizar estas actividades y funcionalidades.

Una vez presentada la nueva taxonomía, el siguiente paso es llevar a cabo investigaciones y análisis exhaustivos de los diferentes CRs disponibles en el mercado. Esto implica clasificarlos con el fin de encontrar una opción similar a la definida, además de recopilar herramientas y tecnologías que no se hayan considerado hasta ahora, y que resultarán útiles durante la implementación de la plataforma elegida. Partiendo de tres estudios [20]-[22], publicados entre 2019 y 2021, basados en revisiones literarias y, que recogen un total de 169 CRs, se ha filtrado la información recogiendo aquellas plataformas centradas en el mundo académico (formación, posibilidad de realizar prácticas y ejercicios), que sean *open-source* para intentar limitar los costes de implementación y/o de dominio público para disponer de información detallada. También se ha considerado que permitan la emulación, simulación y virtualización de elementos para otorgar más posibilidades a los escenarios implementados. Y, finalmente, que dispongan de funcionalidades que den soporte a los bloques definidos en la taxonomía de la Fig.1. La búsqueda se reduce al análisis de 9 plataformas de CR y 11 artículos. Posteriormente, se han analizado más en detalle estos 9 CRs y 11 artículos [23]-[33] para conocer la arquitectura, propósito, capacidades, funcionalidades ventajas y desventajas de los CRs recogidos. Con la información extraída se han podido comparar 38 CRs y aplicativos y comprobar el grado de adecuación a la taxonomía. Finalmente, ha habido tres plataformas que cumplen con las capacidades de la taxonomía, o al menos la gran mayoría de ellas, facilitando la integración con las demás de manera sencilla: KYPO [34], i-tee [28] y CyTRONE [35]. Las tres son *open-source*, y cada una tiene características destacables, como las herramientas de ayuda en el caso de KYPO y el enfoque en el aprendizaje y seguimiento de usuarios en CyTRONE. El último paso para decantar la balanza a favor de una de las tres, ha sido la instalación y despliegue de cada una de ellas, acción que ha permitido determinar los requisitos necesarios, si la

documentación disponible es útil en todo el proceso, así como valorar la implementación final y utilización. Se descarta i-tee por ser de uso limitado a la universidad desarrolladora y tener actualizaciones muy puntuales. CyTrONE tampoco es la elegida ya que presenta dificultades en la implementación y soporte, no se actualiza desde 2017 y no cumple con todas las expectativas. KYPO es la única plataforma adecuada, ya que cumple con la taxonomía y se puede desplegar e implementar correctamente (disponiendo actualizaciones periódicas, documentación pública de consulta y un equipo de asistencia para ayudar en la resolución de problemas).

IV. KYPO Cyber Range Platform (CRP)

KYPO CRP [34], en desarrollo desde 2013 por la universidad de Masaryk en República Checa, es una plataforma flexible, escalable y sofisticada de Cyber Range, con utilidades académicas tanto para la educación de estudiantes como para profesionales de ciberseguridad. Se ha desarrollado teniendo en cuenta tecnologías disruptivas como los contenedores para la optimización de recursos en la virtualización, infraestructura como código para la automatización de la gestión de entornos y escenarios, el uso de tecnologías en la nube, así como microservicios y herramientas de código abierto para la transparencia, reducción de costes y visibilidad operativa.

KYPOCRP se compone de tres secciones principales que permiten la creación de escenarios dentro de la plataforma, así como la preparación de ejercicios y todas las actividades derivadas de éstos:

1. **Despliegue de *OpenStack* [36]**. Se utiliza *OpenStack* para lograr una implementación que proporciona una infraestructura basada en tecnologías de la nube que utiliza contenedores para la virtualización y aprovecha todos los beneficios de ésta.
2. **Creación de objetos necesarios**. Se crean los objetos necesarios dentro de la plataforma OpenStack para permitir la ejecución del CR. Esto incluye la instalación y configuración de los servidores (*KYPO Head* donde se instala la plataforma y *KYPO Proxy* para permitir el acceso a todas las virtualizaciones), y la configuración de las redes que interconectan los servidores, las virtualizaciones y la conexión a Internet.
3. **Configuración y despliegue de la infraestructura**. Se lleva a cabo la configuración y despliegue de la infraestructura, seguido de la instalación de KYPO en esta infraestructura.

La instalación de *OpenStack* se puede realizar de varias formas, ya sea manualmente o utilizando herramientas como *Kolla Ansible* [37], que busca desplegar los múltiples servicios de OpenStack y su infraestructura mediante el uso de contenedores Docker. Se recomienda seguir la guía oficial proporcionada por la organización [38] para conocer los requisitos, dependencias, pasos de instalación, configuración de *Ansible* y *OpenStack*, así como su uso.

Con el objetivo de conseguir una plataforma flexible y escalable para poder dar soporte a una clase de laboratorio de ocupación media de 20 alumnos aproximadamente, se han probado diferentes instalaciones de *Openstack*. Difieren entre ellas en el proceso de instalación y la organización de la arquitectura, proporcionando unas características y ventajas diferenciadas. En las siguientes secciones se explican los resultados obtenidos.

V. KYPO CRP sobre OpenStack All-in-one

A. Introducción y estructura

En este caso, todos los componentes de *OpenStack* se encuentran en una misma máquina física la cual será la encargada de gestionar todos los servicios de *OpenStack*, almacenar todas las imágenes e instancias y de gestionar todas las redes.

B. Instalación

La única máquina que se va a utilizar debe contener un sistema operativo Linux debido a que es el único compatible con *OpenStack*. También debe tener dos interfaces de red, una con IP y una sin IP. Esto es debido a que la interfaz con IP se va a usar para la conexión a Internet, mientras que la interfaz sin IP se va a utilizar para la conexión interna de la máquina que gestiona todo con *OpenStack* y sus instancias. Una vez preparada la máquina, se puede empezar la instalación de *OpenStack*. La instalación de *OpenStack all-in-one* no tiene ninguna peculiaridad a destacar, consiste en seguir los comandos que se pueden encontrar por Internet con una simple búsqueda. Después de instalar *OpenStack*, se instala KYPO siguiendo las instrucciones disponibles en un Gitlab oficial de la universidad de Masaryk [39].

C. Ventajas

Únicamente se usará una máquina física o virtual por lo que ésta debe disponer de los recursos necesarios para abastecer a la plataforma (la gestión de todos los procesos pasará por esta máquina).

Un beneficio que posee este tipo de arquitectura es que la instalación es más simple que otros métodos. Además, debido a que la arquitectura se centra en una única máquina, es más fácil gestionar la parte de red, así como el descubrimiento de fallos cuando surge un problema.

D. Desventajas

Los recursos disponibles son limitados debido a que solo existe una máquina en la arquitectura. No resulta una implementación viable para gestionar un entorno de producción como el necesario en una universidad.

Este tipo de instalación está pensada para destinarla a proyectos pequeños o de uso personal dado que la plataforma resultante no ofrece ni escalabilidad ni estabilidad.

En cuanto a la escalabilidad, la única opción que se posee es la de ampliar los recursos de la propia máquina, con su coste asociado y probables reinstalaciones completas de la plataforma. Por otro lado, al depender de un único elemento, no se tiene ninguna tolerancia al fallo. Es decir, que, si se produce cualquier error o problema con la única máquina de la arquitectura, se pierde el acceso a la plataforma.

E. Conclusión

OpenStack all-in-one tiene una instalación más sencilla que otras arquitecturas de *OpenStack*, pero no está pensado para que sea usado a gran escala. Es muy importante realizar previamente una buena estimación de recursos debido a que la plataforma no es escalable. Esta plataforma, está más destinada al uso particular o para la creación de un pequeño entorno de pruebas. Se ha comprobado que no es viable destinarla a un entorno de producción por su falta de estabilidad, tolerancia al fallo y escalabilidad.

VI. KYPO CRP sobre OpenStack Multi-node

A. Introducción y estructura

En el caso de trabajar sobre *OpenStack multi-node*, interviene más de una máquina física o virtual por lo que el primer paso a realizar es la planificación de la arquitectura que va a soportar la plataforma *OpenStack* sobre la que se instalará y operará KYPO.

En la arquitectura se deben de administrar tres tipos de nodos diferentes: el *Controlador*, el *Nodo de computación* y el *Nodo de red*. Se pueden agrupar varios nodos en una misma máquina física o virtual o, por el contrario, que cada máquina sea un nodo distinto. Las funciones de cada uno de los nodos de la arquitectura son [40]:

- **Controlador**. Es la base del sistema y se encarga de gestionar el funcionamiento de *OpenStack* y de almacenar todas las imágenes que hay en la plataforma.. Este nodo contiene *Nova*, que es el gestor de *OpenStack*. También suele integrar los componentes *Keystone*, *Glance* y *Horizon*.
- **Nodo de red**. Es el que se usa para gestionar las diferentes redes de *OpenStack* y puede estar instalado en una máquina física independiente o en la misma máquina que el controlador. En este nodo se instala el componente *Neutron*.
- **Nodo de computación o hipervisor**. Estos tipos de nodos dotan de recursos los distintos proyectos y también almacenan las distintas instancias creadas de *OpenStack*. En ellos se instalará el hipervisor seleccionado y una API de *Nova* para gestionarlo.

Una peculiaridad de los nodos en esta arquitectura es que pueden crecer horizontalmente, es decir, que no tiene que haber un único nodo de cada tipo en la arquitectura, sino que puede haber varios nodos que desempeñen la misma función.

Para el análisis realizado, se define una arquitectura en la que una máquina virtual es la encargada de gestionar el controlador y el nodo de red, mientras que para la parte de computación se han usado tres servidores físicos.

Cabe destacar que todos estos nodos deben estar conectados entre sí por una red independiente de administración. Por tanto, el nodo controlador que es el que gestiona *OpenStack*, debe tener dos interfaces de red. La primera sin IP para gestionar la comunicación interna con los servicios de *OpenStack* y, la segunda con IP, debe pertenecer a la red de administración para comunicarse con los otros nodos y con Internet. Por otro lado, los nodos de computación solo requieren de una interfaz con IP que pertenezca a la red de administración para poderse comunicar con los otros nodos de la arquitectura (en este caso, solamente un nodo que incluye el controlador y el nodo de red).

B. Instalación

Para poder instalar KYPO, primero se debe instalar la plataforma *OpenStack*. La correcta instalación y obtención de *OpenStack multi-node* se puede dividir en cuatro fases:

1. **Preparación de los nodos de computación.** En cada uno de los nodos se debe instalar un sistema operativo Linux debido a que es el único compatible con *OpenStack*. Una vez se ha instalado el sistema operativo, se debe asegurar que cada nodo de computación está conectado a la red de administración para que se puedan comunicar entre ellos y con el controlador.
2. **Preparación del controlador.** Se crea la máquina virtual del controlador y se realizan los pasos previos a la instalación de *OpenStack* para dejar todo listo. Para este análisis, se ha creado una máquina virtual con el sistema operativo Linux debido a que *OpenStack* solo funciona con sistemas operativos Linux. Posteriormente, se configuran las dos interfaces de red, una con IP conectada a la red de administración y otra sin IP para las comunicaciones internas de *OpenStack*.
3. **Instalación de *OpenStack* multi-node.** En esta fase se realizan todos los pasos para la instalación de *OpenStack* en el controlador y en los dos nodos de computación. Cabe destacar que toda la instalación se puede realizar a través del controlador. Es importante comprobar la correcta comunicación entre los diferentes nodos y el controlador antes de empezar a instalar *OpenStack*.
4. **Ajuste de configuraciones *OpenStack*.** Después de instalar *OpenStack*, se deben realizar unos ajustes para poder obtener el máximo potencial de la plataforma. Cuando se accede a la plataforma de *OpenStack* instalada, primero se ve el apartado de visión general y dentro de esta vista, se observa un apartado que recoge los valores de recursos como instancias, VCPU y RAM. Es necesario entender que los valores mostrados son los que limitarán los recursos que va a tener *OpenStack* al trabajar con KYPO. Por ejemplo, si se dispone de 60 VCPU dentro de los nodos de computación, pero dentro de la visión general solo se muestran 20VCPU, ese va a ser el límite y se van a desperdiciar 40VCPU. No obstante, estos valores se pueden modificar y adaptar a las necesidades del usuario. Para modificar estos valores, se realiza desde el controlador mediante las siguientes instrucciones [41]:
    - **Instancias:** `openstack quota set --instances <número de instancias que se desean> <nombre del proyecto>`
    - **VCPU:** `openstack quota set --cores <número de VCPU que se desean> <nombre del proyecto>`

- **RAM:** `openstack quota set –ram <tamaño de la memoria RAM en MB> <nombre del proyecto>`

Con estas instrucciones, se pueden modificar los tres valores para usar el máximo de recursos disponibles para KYPO. Cabe destacar que el nombre del proyecto se encuentra en la ventana de visión general, arriba a la izquierda, al lado del logo de *OpenStack*. Otro aspecto a tener en cuenta es que la RAM se muestra en la visión general en GB, pero al realizar la instrucción para modificar el límite, se debe introducir en MB. Una vez configurados correctamente los valores según los recursos disponibles que ofrecen los nodos de computación, se puede dar por finalizada la instalación de *OpenStack* y ya se dispone de una arquitectura *OpenStack multi-node* completamente operativa.

Al terminar las cuatro fases de instalación de *OpenStack*, se puede instalar KYPO siguiendo las instrucciones disponibles en un Gitlab de la propia universidad de Masaryk [39].

C. Ventajas

Después de haber visto cómo se estructura el *OpenStack multi-node* y su instalación, se presentan las ventajas de éste.

- **Escalabilidad infinita.** Los nodos de computación son los que ofrecen los recursos disponibles en el proyecto de *OpenStack*. Estos nodos pueden crecer horizontalmente, aunque ya haya sido instalado *OpenStack*, de modo que si se necesitan más recursos basta con añadir nuevos nodos de computación.
- **Estabilidad.** Al poder tener diferentes nodos de computación, aunque falle uno, se van a seguir teniendo recursos para poder crear instancias y trabajar con la plataforma. El nodo controlador que es el más importante, también puede crecer de forma horizontal evitando que deje de funcionar la plataforma si uno de los nodos controlador falla.
- **Adaptación según necesidades.** Esta arquitectura no tiene una forma definida de trabajar y se pueden organizar los nodos en función de las necesidades del proyecto.
- **Entornos de producción.** Debido a todas las ventajas mencionadas anteriormente, esta forma de instalación y gestión de *OpenStack* es de las más idóneas para crear un entorno de producción sobre una plataforma estable, escalable y que se adapta a las necesidades del usuario.

D. Desventajas

Las desventajas de *OpenStack multi-node* son las siguientes:

- **Instalación y gestión.** La instalación es compleja al tener que trabajar con diferentes máquinas físicas y tener que realizar más pasos que en otros métodos de instalación de *OpenStack.*
- **Diseño de la arquitectura.** Es muy importante diseñar correctamente la arquitectura y los nodos que se necesitan para el proyecto. *OpenStack multi-node* es capaz de adaptase según necesidades mientras se haya realizado un buen diseño previo y una buena instalación.

Cabe destacar que con los conocimientos necesarios se pueden solventar estas desventajas y se puede obtener una plataforma que ofrece mucho potencial al usuario.

E. Pruebas realizadas

Una vez demostrado que KYPO sobre *Openstack multi-node* puede ser una solución viable, se han realizado diferentes pruebas para comprobar el funcionamiento de la plataforma creando muchos laboratorios a la vez, simulando lo que pasaría en una formación en ciberseguridad. Estas pruebas han permitido detectar problemas en la virtualización de recursos y el balanceo de carga entre nodos de computación (no se implementa con la configuración de *Openstack* por defecto).

En las primeras pruebas solo se ha conseguido subir a la plataforma 3 escenarios/laboratorios con 3 máquinas virtuales. Después de solucionar los problemas mencionados y con la última instalación y modificaciones de la configuración de *Openstack multi-node*, actualmente se pueden subir hasta 8 escenarios/laboratorios con 9 máquinas virtuales cada uno. Hay que tener en cuenta que un laboratorio con 3 máquinas virtuales consume menos recursos por lo que se pueden subir más de 8 laboratorios con las características de las primeras pruebas.

Finalmente, se han ejecutado todos los laboratorios a la vez simulando una clase y la plataforma ha respondido correctamente. Se puede afirmar que la plataforma actual es estable.

F. Conclusión

KYPO sobre *OpenStack multi-node* ofrece muchas ventajas respecto a la instalación de *OpenStack all-in-one*. *OpenStack multi-node* está pensada para trabajar en un entorno de producción debido a su escalabilidad, estabilidad y tolerancia a los fallos. Es una plataforma muy potente donde no hay límites y la plataforma puede seguir creciendo mientras se disponga de los recursos económicos suficientes. Se puede trabajar perfectamente con KYPO en esta plataforma y es la opción indicada para crear un entorno de CR usable y efectivo para la formación en ciberseguridad.

VII. CONCLUSIONES

En los últimos años, se han multiplicado exponencialmente las ofertas de empleo en el sector de la ciberseguridad. Existe una gran necesidad de expertos en este campo y, consecuentemente, es necesaria formación técnica de ciertas habilidades demandadas por el mercado actual. La Salle Campus BCN lleva muchos años formando a los futuros expertos en ciberseguridad, sin embargo, se ha determinado que es necesario complementar la formación práctica de los estudiantes para crear perfiles aún más competentes a nivel europeo y mundial.

El objetivo principal enmarcado en este artículo es conseguir una plataforma escalable para mejorar la adquisición de competencias y el proceso de evaluación en el ámbito de la ciberseguridad, complementando y

mejorando las metodologías actualmente implementadas para formar a los alumnos.

Existen múltiples metodologías para la formación en ciberseguridad como el *Entrenamiento práctico*, el *Aprendizaje basado en escenarios*, la *Gamificación*, el *Aprendizaje mixto* o los *Programas de certificación*. La que mejor se adapta a las necesidades de la universidad es el *Aprendizaje mixto* pudiéndose implementar combinando clases en formato *face-to-face*, ejercicios prácticos presenciales en modo guiado y tipo *challenge* y recursos en línea accesibles a través de VLEs.

Uno de los componentes cruciales para adquirir las competencias técnicas en ciberseguridad necesarias es la parte práctica. Ésta se puede realizar mediante *Videotutoriales*, *páginas web* de esta temática, *plataformas o testbeds* específicos diseñados para este propósito y/o *Cyber Ranges*. Después de comparar los diferentes métodos se concluye que los *Cyber Ranges* se adaptan perfectamente al objetivo buscado dada su flexibilidad y adaptación a nuevas tecnologías, así como las funcionalidades clave que otorgan para un correcto seguimiento y evaluación de las competencias de los alumnos.

Se ha estudiado el concepto de *Cyber Range*, sus características y funcionalidades para entender exactamente qué son este tipo de plataformas y para qué se utilizan. Se ha podido observar que existen múltiples propósitos, casos de uso, tecnologías involucradas, etc. De modo que se ha optado por recoger toda la información de múltiples estudios previos de CRs para formar una *taxonomía* que contemple todas las características y funcionalidades para la formación en ciberseguridad. Después de valorar lo que supone el diseño y desarrollo de una plataforma como esta, se ha determinado que no se dispone de recursos suficientes para crear un CR desde cero. De modo que se decide realizar un estudio de los CRs disponibles en el mercado y compararlos con la taxonomía definida, para encontrar aquél que mejor se adapta a la misma. Las conclusiones del estudio apuntan a que KYPO CRP es la plataforma de CR que mejor se adapta a las características y funcionalidades buscadas.

KYPO es una plataforma *open-source* flexible, escalable y sofisticada, con utilidades académicas tanto para la educación de estudiantes como para profesionales de ciberseguridad. Se apuesta por la utilización de KYPO, pero antes es necesario instalarla y probar su funcionamiento. KYPO se apoya en tecnología *cloud* como *Openstack* y puede instalarse en una arquitectura *Openstack all-in-one* o *multi-node*. Probadas ambas instalaciones se han valorado las ventajas y desventajas de cada una.

KYPO sobre *Openstack all-in-one* ofrece una instalación simple de la plataforma, pero concentra los recursos en un único nodo. Esto limita la escalabilidad y además supone un único punto de fallo (si falla el nodo en el que está todo instalado y funcionando, la solución deja de estar disponible). Esta arquitectura es adecuada para pequeños proyectos, uso particular y entonos de pruebas. Sin embargo, no es lo suficientemente flexible, escalable y estable para entornos de producción como la universidad (donde el objetivo es tener una plataforma que se pueda utilizar en prácticas de laboratorio de inicialmente unas 20 personas).

Con el funcionamiento de KYPO sobre *OpenStack multi-node* se ha conseguido obtener una plataforma escalable, estable y con soporte de diferentes sistemas operativos. La plataforma es escalable debido a que los nodos pueden crecer de forma horizontal. Por tanto, si se necesitan más recursos, se pueden añadir nuevos nodos de computación a la arquitectura ya montada sin necesidad de volver a reinstalar toda la plataforma. La plataforma es estable debido a que, si se definen correctamente los recursos que va a usar *OpenStack*, nunca serán insuficientes y también porque, si un nodo falla, dado que se tendrán varios nodos dedicados a la misma función, el sistema sigue estando disponible. Por otro lado, la plataforma permite el soporte de nuevos sistemas operativos para trabajar en los escenarios creados en KYPO. Se debe tener en cuenta que en *OpenStack*, todas las imágenes de los sistemas operativos se almacenan en el controlador por lo que éste debe tener los recursos suficientes para almacenarlos y, a su vez, instalar y gestionar *OpenStack*. Por tanto, a diferencia de *OpenStack all-in-one* que el controlador debía también almacenar las instancias de *OpenStack*, se gastan menos recursos y se pueden destinar a almacenar más imágenes de sistemas operativos. Finalmente, destacar que el límite de crecimiento de la plataforma está ligado a la capacidad económica para adquirir nuevas máquinas o nodos de computación y no a la arquitectura o tecnología utilizada.

Disponiendo de la plataforma KYPO CRP funcionando correctamente, se planifican futuras formaciones con un conjunto medio de usuarios (entre 20 y 30 alumnos) que pondrán realmente a prueba la solución planteada para corroborar su flexibilidad, estabilidad y escalabilidad y, comprobar la viabilidad económica.

AGRADECIMIENTOS

Este trabajo se ha llevado a cabo gracias a la financiación recibida por el proyecto Cybersecurity Skills Alliance – A New Vision for Europe (REWIRE) - Grant Agreement 621701-EPP-1-2020-1-LT-EPPKA2-SSA-B, el cual tiene como objetivo principal construir un Blueprint para la industria de la Ciberseguridad y una estrategia europea concreta para la adquisición de habilidades en Ciberseguridad.

REFERENCIAS

[1] REWIRE Project, Cybersecurity Skills Alliance – A New Vision for Europe. https://rewireproject.eu/

[2] Sánchez, J., Mallorquí, A., Briones, A., Zaballos, A., & Corral, G. (2020). An integral pedagogical strategy for teaching and learning IoT cybersecurity. Sensors, 20(14), 3970.

[3] CONE, Benjamin D., et al. A video game for cyber security training and awareness. computers & security, 2007, vol. 26, no 1, p. 63-72.

[4] RAMAN, Raghu; LAL, Athira; ACHUTHAN, Krishnashree. Serious games based approach to cyber security concept learning: Indian context. En 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE). IEEE, 2014. p. 1-5.

[5] CyberSec4Europe Project. "D3.19 Guidelines for Enhancement of Societal Security Awareness". https://cybersec4europe.eu/wp-content/uploads/2022/04/D3.19-Guidelines-for-Enhancement-of-Societal-Security-Awareness_v1.0_submitted.pdf

[6] HackThis. https://hackthis.hackillinois.org/

[7] Google Gruyere - App Engine, *Web Application Exploits and Defenses.* https://google-gruyere.appspot.com/

[8] bWAPP, *A buggy web application.* http://www.itsecgames.com/

[9] Hack This Site. https://www.hackthissite.org/

[10] Abella, J., Corral, G., & Zaballos, A. (2012, October). LOST project, a learning platform for security training. In 2012 International Symposium on Computers in Education (SIIE) (pp. 1-6). IEEE.

[11] Nist.gov. 2021. The Cyber Range: A Guide. [online] Available at: https://www.nist.gov/system/files/documents/2020/06/25/The%20Cyber%20Range%20-%20A%20Guide%20%28NIST-NICE%29%20%28Draft%29%20-%20062420_1315.pdf

[12] Ariessecurity.com. 2020. What is a Cyber Range? A Definitive Guide and Definition | Aries Security. [online] Available at: https://www.ariessecurity.com/what-is-a-cyber-range-a-definitive-guide-and-definition/

[13] Cyberwiser.eu. 2021. What is a Cyber Range? | CYBERWISER.eu. [online] Available at: https://www.cyberwiser.eu/content/what-cyber-range

[14] GovTech. 2018. Cyber Range: Who, What, When, Where, How and Why?. [online] Available at: https://www.govtech.com/blogs/lohrmann-on-cybersecurity/cyber-range-who-what-when-where-how-and-why.html

[15] CloudShare. 2021. What Is a Cyber Range? | CloudShare. [online] Available at: https://www.cloudshare.com/virtual-it-labs-glossary/what-is-a-cyber-range/

[16] Urias, V. E., Stout, W. M. S., Van Leeuwen, B., & Lin, H. (2018). Cyber Range Infrastructure Limitations and Needs of Tomorrow: A Position Paper. Proceedings - International Carnahan Conference on Security Technology, 2018-Octob, 1–5. https://doi.org/10.1109/CCST.2018.8585460

[17] Karjalainen, M., & Kokkonen, T. (2020). Comprehensive Cyber Arena; the Next Generation Cyber Range. Proceedings - 5th IEEE European Symposium on Security and Privacy Workshops, Euro S and PW 2020, 11–16. https://doi.org/10.1109/EuroSPW51379.2020.00011

[18] Deckard, G. M. (2018). Cybertropolis: Breaking the paradigm of cyber-ranges and testbeds. 2018 IEEE International Symposium on Technologies for Homeland Security, HST 2018, 1–4. https://doi.org/10.1109/THS.2018.8574134

[19] ECSO. European Cyber Security Organization. Understanding cyber ranges: From hype to reality. WG5, March 2020. [online] Available at: https://www.ecs-org.eu/documents/uploads/understanding-cyber-ranges-from-hype-to-reality.pdf

[20] Muhammad Mudassar Yamir. Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. Computers & Security, pages 1-60, October 2019. http://dx.doi.org/10.1016/j.cose.2019.101636

[21] Mohamed A.B. Farah Ukwandu, Elochukwu, Hanan Hindy, David Brosset, Dimitris Kavallieros, Robert Atkinson, Christos Tachtatzis, Miroslav Bures, Ivan Andonovic, and Xavier Bellekens. A review of cyber-ranges and test-beds: Current and future trends. Sensors, 20(7148):1_35, December 2020. https://doi.org/10.3390/s20247148

[22] Nestoras Chouliaras. Cyber ranges and testbeds for education, training and research. Applied Sciences, 11(4):1-23, February 2021. https: //doi.org/10.3390/app11041809

[23] Jan Vykopal Pavel Celeda, Jakub Cega and Daniel Tovarnák. Kypo - a platform for cyber defence exercises. Munich (Germany), STO-MP-MSG-133: M&S Support to Operational Tasks Including War Gaming, Logistics, Cyber Defence, p. nestránkováno, pages 1_12, 2015. http://dx.doi.org/10.14339/STO-MP-MSG-133-08-doc

[24] Lorenzo Cavallaro Nicholas CHilders, Bryce Boe, Ludovico Cavedon, Marco Cova, Manuel Egele, and Giovanni Vigna. Organizing large scale hacking competitions. Detection of Intrusions and Malware, and Vulnerability Assessment, 7th International Conference, DIMVA 2010, Bonn, Germany, July 8-9, 2010. Proceedings, pages 132_152, July 2010. http://dx.doi.org/10.1007/978-3-642-14215-4_8

[25] Jon Davis and Shane Magrath. A survey of cyber ranges and testbeds. DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION EDINBURGH (AUSTRALIA) CYBER AND ELECTRONIC WARFARE DIV, page 38, October 2013. https://apps.dtic.mil/sti/citations/ADA594524

[26] Johannes Tammekand Margus Ernits and Olaf Maennel. i-tee: A fully automated cyber defense competition for students. SIGCOMM '15: Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, pages 113_114, August 2015. https://doi.org/10.1145/2785956.2790033

[27] Ian Chapman Sylvain P. Leblanc, Andrew Partington and Mélanie Bernier. An overview of cyber attack and computer network operations simulation. Proceedings of the 2011 Military Modeling & Simulation Symposium, pages 92_100, April 2011. https://dl.acm.org/doi/10.5555/2048558.2048572

[28] Rain Ottis Kaie Maennel and Olaf Maennel. Improving and measuring learning e_ectiveness at cyber defense exercises. Springer International Publishing AG 2017, pages 123_138, 2017. https://doi.org/10.1007/978-3-319-70290-2_8

[29] Zdenek Eichler Radek Oslejsek, Dalibor Toth and Karolína Burská. Towards a uni_ed data storage and generic visualizations in cyber ranges. Proceedings of the 16th European Conference on Cyber Warfare and Security ECCWS 2017, pages 298_306, 2017. https://www.muni.cz/vyzkum/publikace/1385031.

[30] Christos Siaterlis and Marcelo Masera. A survey of software tools for the creation of networked testbeds. International Journal On Advances in Security, 3(1-2):1_12, 2010. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.686.974&rep=rep1&type=pdf.

[31] Ambareen Siraj Joseph Stites and Eric L. Brown. Smart grid security educational training with thundercloud: A virtual security test bed. Information Security Curriculum Development Conference 2013, pages 105_1

[32] Pavel Celeda Jan Vykopal, Radek Oslejsek, Martin Vizvary, and Daniel Tovarnak. Kypo cyber range: Design and use cases. Proceedings of the 12th International Conference on Software Technologies (ICSOFT 2017), pages 310-321, 2017. http://dx.doi.org/10.5220/0006428203100321.

[33] Martin Vizvary Jan Vykopal and Radek Oslejsek. Lessons learned from complex hands-on defence exercises in a cyber range. 2017 IEEE Frontiers in Education Conference, pages 1_8, December 2017. https://doi.org/10.1109/FIE.2017.8190713

[34] Masaryk University. Kypo cyber range platform. https://crp.kypo.muni.cz/

[35] Dat Tang Razvan Beuran, Cuong Pham, Ken ichi Chinen, Yasuo Tan, and Yoichi Shinoda. Cytrone: An integrated cybersecurity training framework. 3rd International Conference on Information Systems Security and Privacy, pages 1-15, June 2018. http://dx.doi.org/10.5220/0006206401570166

[36] Docs OpenStack. https://docs.openstack.org/2023.1/?_ga=2.90516698.1522290031.1687337080-966599635.1687337080

[37] OpenStack. Kolla ansible - quick start. https://docs.openstack.org/kolla-ansible/latest/user/quickstart.html

[38] OpenStack. Software - openstack deployment tools. https://www.openstack.org/software/project-navigator/deployment-tools

[39] Gitlab KYPO Muni. https://gitlab.ics.muni.cz/muni-kypo-crp

[40] Serrano, M. (2017). Cómo funciona el sistema multi-nodo en Openstack. Virtualiza desde Zero. https://virtualizadesdezero.com/multi-nodo-en-openstack/

[41] OpenStack Docs: quota. (s. f.). https://docs.openstack.org/python-openstackclient/pike/cli/command-objects/quota.html

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Caracterización y cuantificación automática del nivel de implicación (*engagement*) de los estudiantes en entornos virtuales de aprendizaje síncronos. Resultados preliminares

Xavier Solé-Beteta, Joan Navarro
Departamento de Ingeniería,
La Salle Campus Barcelona - Universitat Ramon Llull
Quatre Camins, 30. 08022, Barcelona.
xavier.sole@salle.url.edu, jnavarro@salleurl.edu

Resumen

Los entornos virtuales de aprendizaje síncronos, limitan en gran medida la percepción del nivel de implicación (*engagement*) de los estudiantes por parte del docente. Este fenómeno obstaculiza el despliegue de prácticas docentes adaptativas al estado de los alumnos. Aunque las particularidades intrínsecas de dichos entornos virtuales puedan imponer limitaciones, su naturaleza digital también ofrece fuentes únicas de datos, que tratadas expresamente para el propósito pueden devenir una ventaja y consecuente contribución al éxito del proceso enseñanza-aprendizaje. Este trabajo describe una herramienta software que utiliza los datos procedentes de los dispositivos inherentes a los entornos virtuales de aprendizaje síncronos para caracterizar y cuantificar el nivel de implicación de los estudiantes en este tipo de espacios educativos. Los experimentos preliminares conducidos hasta la fecha dejan entrever el potencial de la herramienta para el docente y su grado de aceptación en cuanto a intrusividad se refiere para los alumnos.

***Palabras Clave*—Formación en línea, Atención, Implicación, *Engagement*, Herramienta de soporte a la enseñanza.**

I. INTRODUCCIÓN

El nivel de implicación (*engagement*) de los estudiantes tiene repercusión directa en la adquisición de conocimientos y consecuente impacto en el rendimiento académico [1], [2]. Su observación es de una importancia vital para los docentes con el objetivo de contribuir al éxito del proceso enseñanza-aprendizaje. El docente decide cómo tratar los contenidos de acuerdo con los objetivos académicos, seleccionando actividades formativas acordes con la esperanza de mantener o (incluso) aumentar la implicación de los estudiantes en el proceso de aprendizaje [3]. Esto da lugar al despliegue de prácticas docentes adaptativas al nivel de implicación de los estudiantes. Este despliegue debe adaptarse no sólo a los alumnos y a las capacidades del docente sino que también a las características del espacio de aprendizaje, lo que pone en relieve las diferencias, limitaciones y restricciones que presentan los entornos de aprendizaje presenciales respecto a los entornos virtuales de aprendizaje síncronos.

En un entorno presencial, dada la innata coexistencia física entre ambos actores partícipes (docentes y estudiantes), los docentes perciben el nivel de implicación mediante la observación natural (e inconsciente) de expresiones faciales, gestos, posiciones corporales, tonos de voz, así como otros indicadores que permiten identificar comportamientos y emociones específicos que inducen niveles de implicación. Dicha percepción está supeditada a las capacidades del observador, es subjetiva y también dependiente de su experiencia en el campo. Contrariamente, en los entornos virtuales de aprendizaje síncronos esta práctica no es fácilmente reproducible [4]. Las limitaciones inherentes a la no convivencia física en el espacio con los alumnos y la propia naturaleza de las herramientas digitales dificultan en gran medida la cuantificación del nivel de implicación de los estudiantes. Este aspecto compromete las capacidades docentes del profesor, limitando el potencial de su experiencia, así como originando un posible impacto negativo en el nivel de adquisición de conocimientos [5]. Un ejemplo reciente de este aspecto, fueron los trances experimentados por la transición forzada de la docencia presencial a virtual durante la pandemia de la COVID-19 [6].

No obstante, los entornos virtuales de aprendizaje síncronos presentan unas características constitutivas que ofrecen oportunidades únicas para poder capturar, medir y monitorizar el nivel de implicación de forma automática

y objetiva. Con este objetivo, en [7] se propuso el desarrollo conceptual de una herramienta software capaz de, utilizando datos capturados por los dispositivos inherentes a los entornos virtuales de aprendizaje síncronos (pantalla, cámara y micrófono), caracterizar y cuantificar el nivel de implicación de los estudiantes en este tipo de espacios educativos. El propósito de este trabajo es compartir la evolución en el desarrollo de la herramienta y evaluar los resultados obtenidos tras ser utilizado en 26 sesiones con un total de 291 alumnos. Estos resultados preliminares están focalizados en evaluar el nivel de aceptación e intrusividad (entendida como la percepción de tener un sistema software monitorizando continuamente los indicadores asociados al nivel de implicación) percibidos por parte de los alumnos al saber que hay una herramienta automática capturando constantemente los datos que se desprenden de la sesión virtual síncrona para mejorar su experiencia de aprendizaje. Tanto el desarrollo como las pruebas se han llevado a cabo en el contexto del proyecto europeo HOTSUP (Erasmus+).

El resto del trabajo está organizado como sigue. La Sección 2 describe los elementos digitales que la herramienta software [7] tiene en cuenta para caracterizar el nivel de *engagement* de los alumnos en entornos virtuales de aprendizaje síncronos. A continuación, la Sección 3 muestra los resultados preliminares obtenidos en cuanto a intrusividad de la herramiente se refiere. Finalmente, la Sección 4 resume las principales conclusiones obtenidas y discute las posibles líneas futuras de investigación.

II. CARACTERIZACIÓN DE LA IMPLICACIÓN (*ENGAGEMENT*)

Para cuantificar el nivel de implicación de los estudiantes, y tomando como restricciones la no adición de dispositivos artificiales ni la integración con otros sistemas más allá del entorno virtual de aprendizaje síncrono, se han llevado a cabo dos tareas principales: (1) la identificación de aquellos parámetros digitales sobre los que se apoya el *engagement* de los estudiantes y (2) la ideación de un modelo analítico para su cuantificación. El detalle de cada una de estas fases, así como los resultados obtenidos, son explicados a continuación.

A partir de las señales de audio, texto y vídeo, el propósito de esta fase ha sido definir un conjunto de características digitales que permitan modelar las dimensiones de la implicación [8], [9] que son obtenibles durante el transcurso de una sesión virtual síncrona. Cabe notar que estas son las referentes a la implicación conductual (*behavioral engagement*) y emocional (*emotional engagement*), puesto que la cognitiva (*cognitive engagement*) precisaría integración con otros sistemas más allá de un entorno de videoconferencia.

El proceso para la obtención del conjunto de características digitales ha comprendido, además de la propia identificación, su categorización. Concretamente, se han identificado un total de 46 características digitales, las cuales han sido categorizadas en 10 categorías. Esto ha permitido mejorar tanto su comprensión como manejabilidad.

Cuadro I
CATEGORIZACIÓN DE LAS CARACTERÍSTICAS DIGITALES IDENTIFICADAS.

Categoría digital	Tipo de implicación (*engagement*)
Asistencia	Conductual / Emocional
Uso de cámara	Conductual / Emocional
Interacciones de voz	Conductual
Mano levantada	Conductual
Compartición de pantalla	Conductual
Interacciones de chat	Conductual / Emocional
Análisis de sonido	Conductual
Emoción facial	Emocional
Movimiento de los labios	Conductual / Emocional
Seguimiento ocular	Emocional

Las categorías definidas y correspondiente descripción de las características digitales son las siguientes:

- **Asistencia.** Agrupa todos aquellos aspectos vinculados con el número de participantes en la sesión y sus variaciones.
- **Uso de cámara**. Comprende eventos referentes al estado y cambio de estado de la cámara.
- **Interacciones de voz.** Aglutina características de las intervenciones de voz, como lo son la duración, el número y silencio.
- **Manos Levantadas y Compartición de pantalla.** Agrupan las detecciones de mano levantada y compartición de pantalla respectivamente.
- **Interacciones de chat.** Unifica características propias del uso de chat, como el número de mensajes, el tipo de *stickers* y el número de preguntas.
- **Análisis de sonido.** Reúne los aspectos referentes a la calidad del sonido como ruido de fondo, volumen y discontinuidad.
- **Emoción facial.** Incluye características respectivas a la detección de las emociones de los asistentes, como la emoción principal (contento, neutral…) y el número de estudiantes para cada una de las emociones.
- **Movimiento de los labios.** Incluye detecciones como el bostezo y la interacción con otras personas ajenas a la videoconferencia.
- **Seguimiento ocular.** Relaciona el número de estudiantes que abandona la zona de captura de la cámara y el hecho de no estar mirando la pantalla.

La relación entre las características y el tipo de implicación puede verse en la Tabla I.

Tomando las características digitales presentadas anteriormente como punto de partida, se ha procedido con la definición del modelo analítico capaz de cuantificar el nivel de implicación de los estudiantes.

Para asegurar la adecuación del modelo al propósito, y tomando en consideración las propiedades intrínsecas de la implicación [10] así como el hecho de que la herramienta debe ser un soporte a la práctica docente, se han establecido un conjunto de premisas que han guiado el diseño. En concreto:

1. El nivel de implicación detectado debe ser de fácil y rápida interpretación.

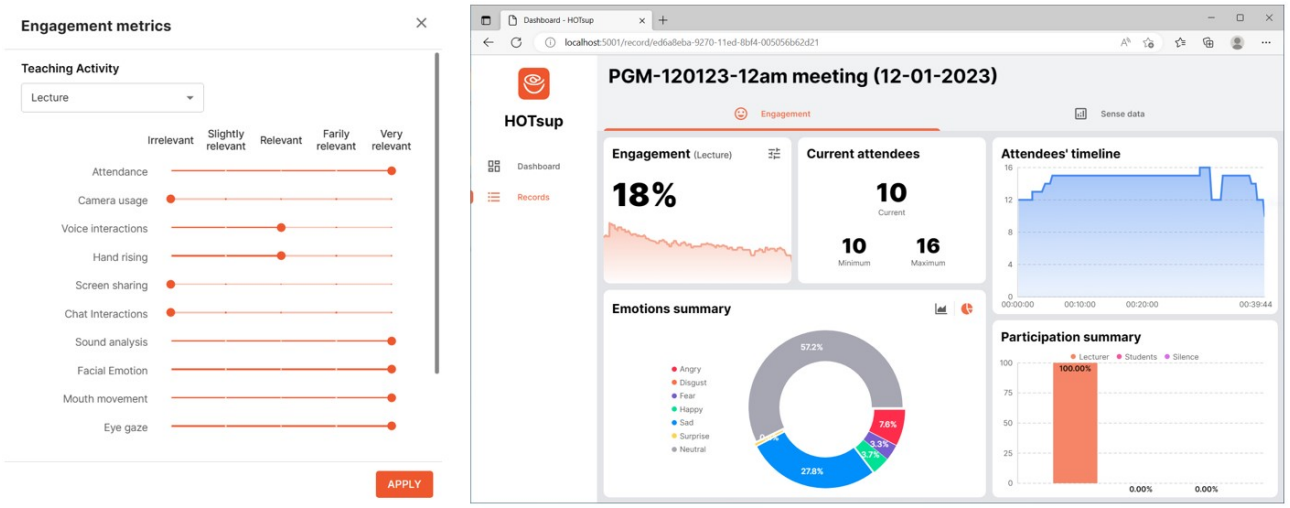


Figura 1. Captura de pantalla de la herramienta docente para la monitorización del nivel de implicación.

2. Los comportamientos y las emociones individuales de los estudiantes, así como sus cambios, deben implicar variaciones en el nivel de implicación detectado.
3. El cálculo del nivel de implicación debe ser configurable según las preferencias del docente y el tipo de sesión (p.ej., clase magistral, clase de dudas, trabajo rn grupo, seminario, ...).
4. El nivel de implicación de los estudiantes debe tener una tendencia a decrecer a lo largo del tiempo, y esta disminución debe ser configurable por el docente.

Para dar respuesta al primer argumento, se ha establecido que el nivel de implicación sea representado mediante un número entero llamado ENQUA (*ENgagement QUAntification*). Este puede tomar valores porcentuales de 0 a 100. El valor 0 indica que no se ha detectado ninguna implicación, mientras que 100 revela un estado en el que todos los comportamientos y las emociones detectados indican implicación total.

En cuanto a la segunda restricción, referente a reflejar tanto los comportamientos como emociones detectadas, así como sus variaciones, la ecuación toma características digitales identificadas en la fase anterior agrupadas según las (10) categorías definidas. De esta forma el modelo es fiel a la realidad detectada y al mismo tiempo susceptible a sus variaciones. Esto puede observarse en el segundo término de la Ecuación 1.

$$ENQUA(t) = \alpha * (e^{-t/k}) + \beta * (\sum_{i=1}^{10} W_i * CD_i) \quad (1)$$

Respecto a la tercera característica de diseño, su alto grado de configurabilidad, se han realizado las siguientes acciones: (1) la creación de una escala de valores para ponderar (W_i) la influencia de las categorías digitales en el cálculo de la implicación y (2) la definición de un conjunto de valores de influencia por defecto (y manualmente ajustable por el docente) para un conjunto de actividades académicas contrastadas (como lo son clase magistral, tutorial o laboratorio).

En último lugar, con relación a la inclusión de la disminución natural de la implicación con el transcurso del tiempo [10] y su adaptabilidad según preferencias del profesor, se ha introducido el termino exponencial de la Ecuación 1. Tal y como puede observarse, t modela el tiempo y K denota el grado de tendencia a la pérdida de implicación del alumnado. Asimismo, en la Ecuación 1, y en línea con la premisa 3, cabe destacar que los parámetros α y β permiten configurar el efecto de esta pérdida natural de la implicación en el cálculo global del nivel de implicación detectado objetivamente.

III. RESULTADOS PRELIMINARES

El modelo analítico ha sido implementado mediante una herramienta software (véase Fig. 1). Esta herramienta software se conecta automáticamente a la sesión virtual síncrona, como si fuera un alumno más de la sesión, y captura todos los datos disponibles detallados en la Sección 2 para presentarlos de una forma inteligible al docente a través de una interfaz web. Así, este software permite observar: la evolución del nivel de implicación global de los alumnos (Engagement), el número de asistentes, así como su evolución (Current attendees y Attendees' timeline), el resumen de las emociones principales detectadas (Emotions summary) y las intervenciones de voz (Participation summary). El software ha sido probado en 26 sesiones, permitiendo obtener unos resultados preliminares tanto de docentes (14) como de alumnos (291) obtenidos a partir de encuestas en línea.

Las Fig. 2 y Fig. 3 muestran los resultados obtenidos a 2 preguntas. La primera, referente al nivel de intrusividad y la segunda, sobre el condicionamiento de la implicación mostrada sabiendo sobre la existencia del software de medición. Es de merecida mención que solamente el 6,3 % de los encuestados cree que el nivel de intrusividad de la herramienta es elevado, y también el hecho que un 32 % afirma no haber mostrado un comportamiento condicionado por ser conscientes de la medición de la implicación. A pesar de ello, se detecta que hace falta un

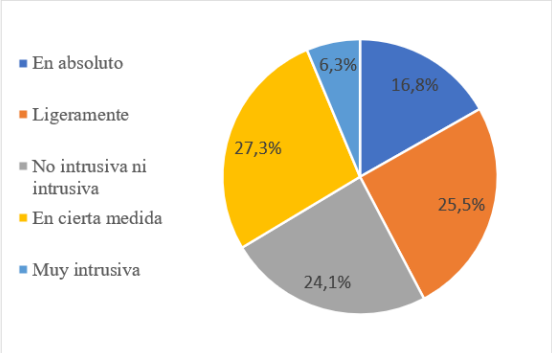


Figura 2. Resultados obtenidos a la pregunta "¿Considera el software intrusivo?".

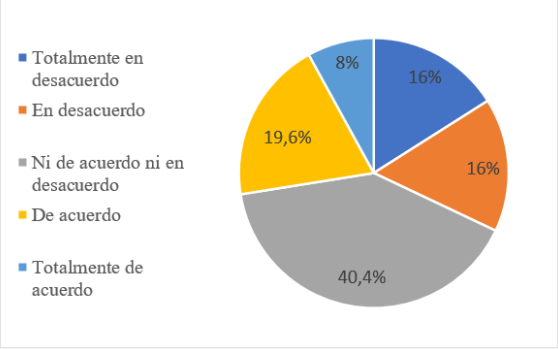


Figura 3. Respuestas obtenidas, según escala Likert, a la pregunta ¿Hasta qué punto está de acuerdo en que ha estado más involucrado en las clases sabiendo que su implicación fue monitoreada?.

esfuerzo didáctico más efectivo para transmitir qué hace y cómo funciona la herramienta. Además, en el apartado de respuestas abiertas de las encuestas los docentes que utilizaron la herramienta coincidieron en que la información que este software les muestra no la tienen actualmente disponible por otros medios.

IV. CONCLUSIONES Y LÍNEAS DE FUTURO

El recurso docente presentado pretende ser un soporte al docente en respuesta a las limitaciones inherentes a los entornos virtuales de aprendizaje síncronos para la cuantificación del nivel de implicación de los estudiantes. Minimizar este tipo de diferencias entre los entornos presenciales y virtuales síncronos contribuye a que el docente pueda llevar a cabo una docencia adaptativa a la implicación del alumnado, así favoreciendo la adquisición de conocimientos. Un aspecto crítico de este tipo de sistemas es la intrusividad percibida. Los resultados obtenidos dejan entrever la aceptación de la herramienta en gran medida por parte de los alumnos. También, de acuerdo con los resultados obtenidos, el alumnado parece no mostrar un comportamiento condicionado por su existencia.

En cuanto a líneas de futuro, en este momento se abren varios posibles itinerarios. Por un lado, se debe contrastar que los resultados mostrados por la herramienta software son útiles y están correlacionados con la realidad percibida por el docente. Para ello sería de gran ayuda llevar a cabo entrevistas (por ejemplo de tipo *Bipolar Laddering* [11]) con los docentes, los alumnos e incluso algún observador externo de la sesión. Por otro lado, sería muy interesante explorar la posibilidad de que la herramienta ofreciera una recomendación fiable para los valores de los pesos de cada categoría digital (véase parte izquierda de la Fig. 1). Actualmente, el profesor/a de la sesión ajusta estos pesos manualmente en función de su conocimiento o expectativas que pueda tener antes de la sesión. Con un volumen de datos suficientemente alto recogido de varias sesiones, se estima que sería posible obtener valores coherentes para estos pesos mediante un sistema de inteligencia artificial. Por último, se debería contrastar mediante pruebas empíricas hasta qué punto el uso de esta herramienta por parte del profesor/a mejora la experiencia de aprendizaje de los alumnos. Para ello, se propone llevar a cabo una batería de experimentos de tipo A/B y analizar las respuestas tanto de docentes como de alumnos.

AGRADECIMIENTOS

Los autores quieren agradecer a los programas Erasmus+ (Referencia: 020-1-PL01-KA226-HE-096456 - HOTSUP: HOlistic online Teaching SUPport) y Aristos Campus Mundus (Referencia: ACM2023_15) por financiar parte de esta investigación.

REFERENCIAS

[1] J.-S. Lee, "The relationship between student engagement and academic performance: Is it a myth or reality?" *The Journal of Educational Research*, vol. 107, no. 3, pp. 177–185, 2014.

[2] A. P. Delfino, "Student engagement and academic performance of students of partido state university." *Asian Journal of University Education*, vol. 15, no. 1, p. n1, 2019.

[3] H. A. El-Sabagh, "Adaptive e-learning environment based on learning styles and its impact on development students' engagement," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, pp. 1–24, 2021.

[4] L. Durán López, D. Gutiérrez Galán, E. Cerezuela Escudero, J. A. Ríos Navarro, and J. P. Domínguez Morales, "Semipresencialidad en tiempos de covid-19: adaptación de la docencia en el ámbito de fundamentos de informática," *JENUI 2021: XXVII Jornadas sobre la Enseñanza Universitaria de la Informática (2021), pp. 315-318.*, 2021.

[5] S. Kurbakova, Z. Volkova, and A. Kurbakov, "Virtual learning and educational environment: New opportunities and challenges under the covid-19 pandemic," in *2020 The 4th International Conference on Education and Multimedia Technology*, 2020, pp. 167–171.

[6] D. M. Cretu and Y.-S. Ho, "The impact of covid-19 on educational research: A bibliometric analysis," *Sustainability*, vol. 15, no. 6, p. 5219, 2023.

[7] X. Solé-Beteta, J. Navarro, B. Gajšek, A. Guadagni, and A. Zaballos, "A data-driven approach to quantify and measure students' engagement in synchronous virtual learning environments," *Sensors*, vol. 22, no. 9, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/9/3294

[8] J. A. Fredricks, M. Filsecker, and M. A. Lawson, "Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues," pp. 1–4, 2016.

[9] J. A. Fredricks, *Eight myths of student disengagement: Creating classrooms of deep learning*. Corwin Press, 2014.

[10] C. A. Boulton, E. Hughes, C. Kent, J. R. Smith, and H. T. Williams, "Student engagement and wellbeing over time at a higher education institution," *PloS one*, vol. 14, no. 11, p. e0225770, 2019.

[11] M. Pifarré and O. Tomico, "Bipolar laddering (bla) a participatory subjective exploration method on user experience," in *Proceedings of the 2007 Conference on Designing for User eXperiences*, 2007, pp. 2–13.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Detección de valores atípicos en el uso de las redes móviles a través de espacios de baja dimensión

David Cortés-Polo[(1)], Jesús Calle-Cancho[(1)], Luis Ignacio Jiménez[(2)], Francisco Javier Rodríguez-Pérez[(1)], Mihaela I. Chidean [(3)].

dcorpol@unex.es, jesuscalle@unex.es, nacho.jimenez@uva.es,
fjrodri@unex.es, mihaela.chidean@urjc.es

[(1)]Dpto. de Ingeniería de Sistemas Informáticos y Telemáticos. Universidad de Extremadura. Cáceres, España.
[(2)]Grupo de Sistemas Inteligentes, Dpto. de Informática. Universidad de Valladolid, Valladolid, España.
[(3)]Dpto. de Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación, Universidad Rey Juan Carlos, Fuenlabrada, España.

Las redes de datos son uno de los pilares fundamentales en la cuarta revolución industrial. Las redes 5G y *Beyond 5G* (B5G) conectarán personas, equipos y objetos, por lo que entender su rendimiento es crucial. Los operadores pueden planificar y tomar decisiones basadas en la información extraída de la red. Debido a la cuantiosa cantidad de información que se captura de la red, encontrar mecanismos que permitan analizar el comportamiento estándar de los usuarios y encontrar eventos que sean interesantes de analizar se vuelve un gran desafío para los operadores. En este artículo, se propone un mecanismo cuya ventaja es analizar y extraer el comportamiento básico de los usuarios de la red en una zona, así como detectar valores atípicos o anómalos de uso de la red. Gracias a esta técnica no sólo se resaltan los valores anómalos, sino que también el comportamiento temporal de cada celda.

***Palabras Clave*—clustering, 5G, análisis de la red, espacios de baja dimensionalidad, análisis de eventos**

I. INTRODUCCIÓN

El número de dispositivos conectados a la red ha experimentado un crecimiento exponencial cada año, llegando a aproximadamente 4.9 mil millones de usuarios activos en 2021. Para este año 2023 se esperan aproximadamente 29.3 mil millones de dispositivos conectados [1], entre los que se incluyen dispositivos de bajo consumo, teléfonos inteligentes, tabletas, sensores, identificación por radiofrecuencia, etc. Estos nuevos dispositivos también aprovechan las capacidades y servicios novedosos que ofrecen las redes 5G y B5G, como un mayor ancho de banda, una baja latencia y la inclusión de nuevas entidades en la red, como la computación en el borde de la red. Como resultado de estas innovaciones, se espera que se mejore e incremente el desarrollo de servicios y aplicaciones industriales de próxima generación, como vehículos autónomos, robótica e Internet de las Cosas (IoT). Además, estas redes de próxima generación mejoran la recopilación de información en múltiples capas de la red, permitiendo a los proveedores de servicios mejorar las tareas de gestión mediante técnicas de aprendizaje automático para el análisis de datos [2].

Una de las bases de datos más importantes donde los operadores almacenan información de la red es la conocida como *Call Detail Record* (CDR). Esta herramienta almacena información geolocalizada y es utilizada para la obtención de información sobre el uso de la red y su posterior análisis. Tradicionalmente, el análisis de los CDR se enfocaba en el estudio de la actividad humana y su movilidad, pero se han obtenido resultados prometedores en diversas áreas como la planificación urbana, medicina o en la detección de anomalías y la clasificación de patrones de tráfico [3]. Por tanto, se puede observar que los CDR contienen información muy valiosa que permite a los investigadores y a los operadores analizar el uso de la red desde múltiples puntos de vista.

La información almacenada en cada CDR generalmente se puede estructurar como un cubo de datos tridimensional: dos dimensiones para la ubicación espacial y una dimensión para las características de la red. Además, los CDR también incluyen información temporal que establece un orden de ocurrencia. Gracias a esta organización de la información se pueden usar nuevas técnicas de análisis de información, agregando diferentes

perspectivas a las tareas de gestión al incluir el patrón de comportamiento del usuario. Ejemplos de este enfoque se incluyen en trabajos previos como [3] y [4], donde se analiza un CDR real de una red celular que incluye datos de uso recopilados durante dos meses completos con diferentes metodologías, basadas en modelos de mezcla lineal. Los resultados de estos trabajos incluyen vectores descriptores de diferentes niveles de uso de la red, llamados "comportamientos" o "*comportments*". Además, los "comportamientos" ordenados (en función de su módulo) reflejan la escala relativa entre los valores mínimos y máximos de intensidad de uso de la red y facilitan la detección de eventos atípicos a nivel de celda. La principal limitación de estos y otros trabajos anteriores es que el análisis se realiza únicamente sobre las anomalías, pero no se obtienen los patrones del uso de cada zona.

Es por esto que en este trabajo se propone un método que permita extraer el comportamiento base de una zona, de forma que todos aquellos eventos que queden fuera de ese comportamiento puedan ser clasificados como posibles eventos con valores atípicos o anomalías de la red. Para ello, la principal aportación de este trabajo se basa en métodos de análisis de datos dispersos como son las técnicas de descomposición de matrices de bajo rango. Esta aproximación permite descomponer el comportamiento de la zona en valores de baja dimensionalidad para describir su patrón general de uso, y unos valores "resto" que representan los posibles eventos o anomalías de la red. Esto introduce una ventaja a la hora de analizar los datos, ya que el disponer de un patrón base de cada zona analizada, permite clasificar nuevos eventos de forma más rápida así como, incorporando nuevas fuentes de datos como información poblacional o de tráfico humano, pueda extrapolar comportamientos de uso de los servicios ofrecidos por la red.

El resto del artículo está organizado de la siguiente manera. La sección II presenta las particularidades de los CDR y la metodología Análisis de la Red usando Proyección Ortogonal (OPNA), describiendo y analizando sus fases. La sección III detalla el método propuesto para la descomposición de baja dimensionalidad de los datos. Los resultados de los experimentos propuestos sobre un conjunto de datos del área metropolitana de Milán se detallan en la Sección IV. Finalmente, en la sección V se presentan las conclusiones de este trabajo.

II. ANÁLISIS DEL USO DE LA RED A TRAVÉS DE LA METODOLOGÍA OPNA

Como se ha explicado anteriormente, los conjuntos de datos CDR son grandes bases de datos que almacenan información georeferenciada del uso de la red con información subyacente muy importante para los operadores e investigadores. Entre esa información subyacente se encuentran los patrones de "comportamiento" de una ubicación que define si se han usado los recursos de la red de forma intensiva, o en qué periodos se ha hecho un mayor uso de los mismos.

Para encontrar esos "comportamientos" la información se suele organizar como un cubo donde cada celda se define por diferentes valores que caracterizan su comportamiento. El análisis de datos con múltiples características presenta problemas relacionados con la información espacial y las características del área analizada. Además, a estos problemas hay que añadirle la complejidad del manejo de las series temporales. Esto hace que los "comportamientos" detectados en las zonas principalmente agrícolas sean completamente diferentes a las zonas muy pobladas del centro de la ciudad. Para solventar estos problemas, en trabajos previos, se desarrolló la metodología OPNA [4] que representa cada celda del CDR como una combinación de "comportamientos" extraídos del análisis del conjunto de los datos. Gracias a la ponderación de éstos basándose en su módulo, se consigue ordenarlos describiendo la intensidad del uso de la red.

El cálculo de la intensidad de la red a través del análisis de "comportamientos" es un proceso complejo que comienza con la extracción de puntos clave del CDR en cada intervalo de tiempo. Para cada uno de esos intervalos se busca los "comportamientos" más extremos que describen al conjunto de datos y que permiten describir cada punto del conjunto de datos en ese intervalo como una combinación lineal de los "comportamientos" extraídos por el algoritmo. De forma que esos vectores describen cada "comportamiento" de la red para un determinado intervalo de tiempo.

Una vez extraídos los puntos clave, se asigna el "comportamiento" más cercano a cada subárea a través de un proceso de minimización de la distancia euclídea. Esta distancia se calcula como una norma-2 en un espacio l-dimensional. Por último, para establecer el orden entre los diferentes "comportamientos", se utiliza el valor de su norma, estableciendo la relación de los niveles de uso de la red para cada celda. La Figura 1 muestra de forma visual el proceso para clasificar el "comportamiento" de cada celda.

En la primera fase, la metodología carga los datos almacenados en el CDR ya tratados y normalizados. A través de la técnica Proyección Ortogonal de los vectores en el Subespacio (OSP) se consigue extraer los vectores que representan al conjunto de los datos analizados, siendo los datos más extremos del simplex que resuelve el sistema de ecuaciones propuesto [4]. Estos vectores serán los "comportamientos" que describen el uso de la red, mientras que el resto de puntos se aproximarán al que tenga menor distancia euclídea, como se ha explicado anteriormente. Finalmente, la sucesión de los diferentes "comportamientos" que describen el área a lo largo del tiempo, permite describir el uso de la red a lo largo del periodo analizado. De esta forma, estos valores serán los usados en este trabajo para la extracción del patrón base y la detección de comportamientos atípicos en la red.

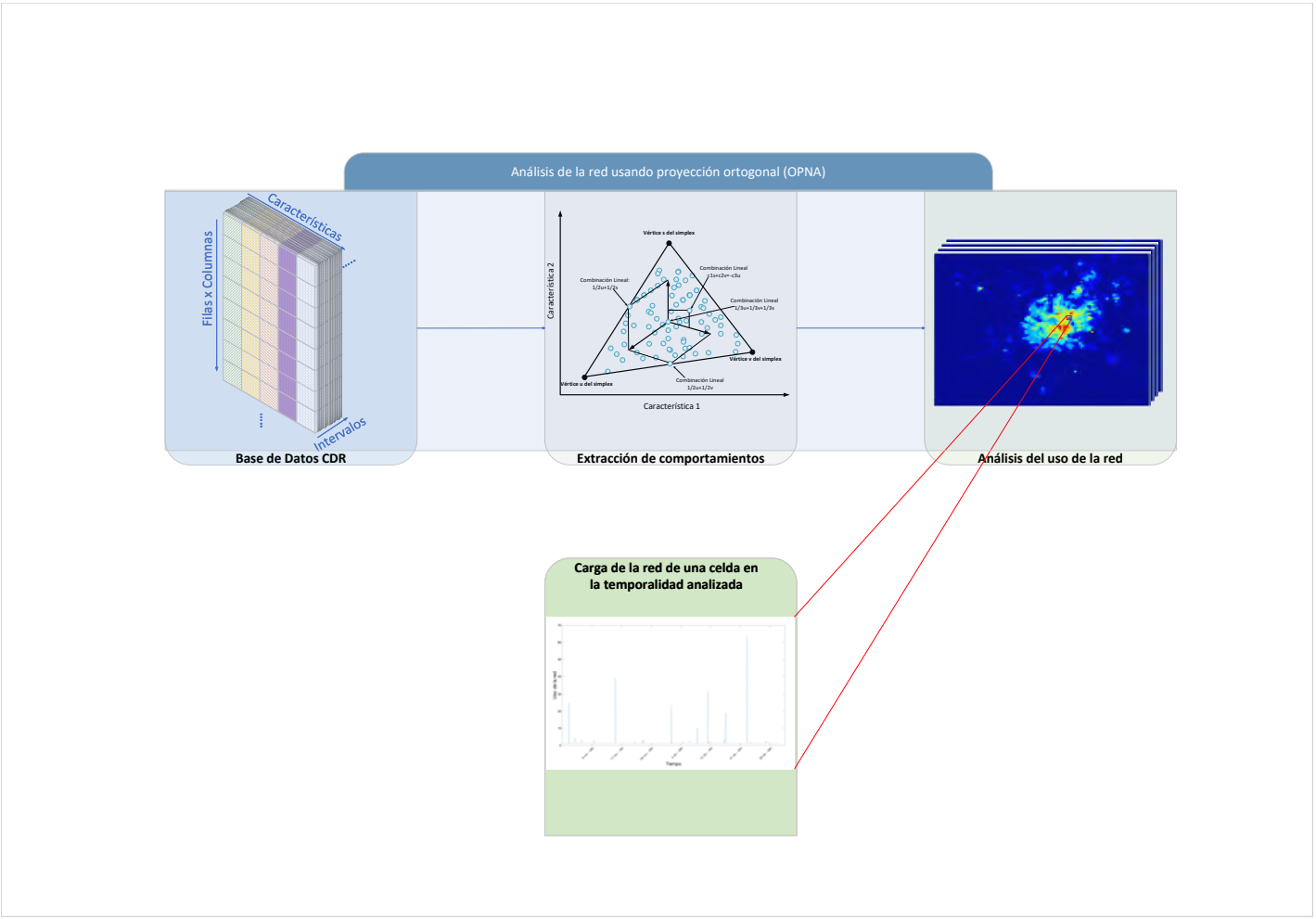


Fig. 1. Representación esquemática del análisis realizado usando la metodología OPNA.

III. MÉTODO PARA EXTRACCIÓN DE PATRÓN BASE Y ANÁLISIS DE EVENTOS EN PUNTOS CLAVE DEL CONJUNTO DE DATOS

En el análisis de datos y aprendizaje automático, uno de los enfoques clásicos para reducir una gran colección de datos es usar el Análisis de Componentes Principales (PCA). De forma que dada una colección de puntos, PCA calcula una proyección lineal de los puntos en un subespacio de baja dimensión que minimiza el error usando la norma-2 entre los puntos originales y los puntos proyectados. Sin embargo, debido a este cálculo, el subespacio de baja dimensión proporcionado por PCA es sensible a valores atípicos. En particular, los valores atípicos tienden a desviar el subespacio hacia ellos debido a la norma-2, lo que hace que la distancia entre el subespacio calculado y los valores atípicos que se desean detectar sea indeseablemente pequeña.

A. Obtención del espacio de baja dimensionalidad para extracción del patrón base

En ciertos problemas como el analizado en este artículo, donde en cada celda existen una mezcla muy alta de valores normales y anómalos (o con eventos que incrementan el uso de la red en ciertos momentos) debido, principalmente, a la gran variedad de usuarios que hacen uso de los recursos de la red. Dado que el objetivo de partida de este trabajo es encontrar los comportamientos básicos de cada celda, así como aislar los valores anómalos, se requiere de un análisis más profundo y no tan susceptible a los valores anómalos.

Para evitar las desventajas comentadas anteriormente, se han desarrollado otros métodos como el Análisis de Componentes Principales Robusto (RPCA) que se enfoca en proyectar los puntos originales sobre un espacio de baja dimensión y es especialmente indicado para datos que contienen valores atípicos [5]. Esta descomposición tiene profundas implicaciones para muchos problemas modernos como son la vigilancia de vídeo, el reconocimiento facial, el procesamiento del lenguaje natural, o los problemas de clasificación.

Este algoritmo presupone que existe un subespacio de baja dimensión $\mathbf{L}$ y un conjunto de valores atípicos $\mathbf{S}$ en esos datos, de forma que la matriz de datos original $\mathbf{Y}$, y su descomposición en un subespacio de baja dimensión es descrita por la Ecuación 1:

$$\mathbf{Y} = \mathbf{L} + \mathbf{S} \tag{1}$$

Matemáticamente, el fin último del algoritmo es

encontrar $\mathbf{L}$ y $\mathbf{S}$ que satisfagan lo siguiente:

$$\min_{L,S} rank(\mathbf{L}) + \|S\|_0 \quad sujeto\ a \quad \mathbf{L} + \mathbf{S} = \mathbf{Y} \qquad (2)$$

Dado que la resolución de esta optimización es convexa, se puede resolver el sistema con una relajación convexa de la optimización:

$$\min_{L,S} \|L\|_* + \lambda\|S\|_1 \quad sujeto\ a \quad \mathbf{L} + \mathbf{S} = \mathbf{Y} \qquad (3)$$

Donde $\|L\|_*$ es la norma nuclear de $\mathbf{L}$, definida como la suma de los valores singulares de $\mathbf{L}$ y $\lambda = \sqrt{max(n, m)}$, siendo n y m las dimensiones de la matriz $\mathbf{Y}$.

B. Clasificación de los comportamientos base del conjunto de datos

Una vez resuelto la optimización a través del algoritmo RPCA, la matriz $\mathbf{L}$ contendrá la información mínima que define el comportamiento básico de cada una de las subáreas analizadas. En este trabajo, se empleará el algoritmo k-means [6] para realizar la clasificación de los comportamientos base del conjunto de datos. K-means es un algoritmo no supervisado que asigna cada vector de comportamientos base a los k grupos predefinidos. El objetivo es reducir la varianza dentro de cada grupo al asignar cada muestra al grupo más cercano en función de su norma-2.

El único valor configurable del algoritmo es la elección del parámetro k que será el número de clústeres que intentará construir el algoritmo. La selección del valor de k es crucial en k-means y suele requerir métodos adicionales para determinar el óptimo. En este estudio, se utiliza el criterio de evaluación de clústeres de Calinski-Harabasz (CHI) [7] para esta tarea. El CHI evalúa la cohesión del clúster calculado, midiendo las distancia entre los diferentes puntos del clúster con el centroide calculado. Además se mide la separación con otros clústeres calculados.

Sin embargo, el método del CHI presenta limitaciones, ya que evalúa el valor óptimo de k después de realizar el agrupamiento. Además, debido a los mecanismos inicialización de los centroides en k-means, existe la posibilidad de obtener un valor de k subóptimo. Para superar esta limitación, en este trabajo se ha adoptado un enfoque empírico. Se han realizado múltiples ejecuciones de k-means y cálculos del CHI para diferentes valores de k analizando de forma estadística cuál es el valor de k más adecuado.

C. Análisis de los puntos clave y detección de valores atípicos en el conjunto de datos

Tras el análisis de la matriz $\mathbf{L}$ con la información mínima del uso de la red, se procederá al análisis de la matriz $\mathbf{S}$, la cual contendrá información muy distribuida que atiende a los usos esporádicos de la red. Bajo esta premisa, esta matriz contendrá, por un lado, información de las pequeñas diferencias con respecto al comportamiento base, así como los eventos que sean atípicos.

Para discernir entre las pequeñas diferencias y los eventos atípicos, este trabajo va a utilizar un método que permita detectar aquellos eventos fuera de lo normal. Para este análisis, el primer paso que se realizará será aproximar el conjunto de datos analizados a una distribución exponencial que describa el comportamiento de esa celda. Una vez ajustados los datos, donde la gran mayoría serán pequeñas variaciones del comportamiento base, se deberá encontrar el límite que permita clasificar un evento como atípico o no. Para ello se utilizará el cálculo del rango intercuartílico (IQR) [8], el cual va a permitir analizar la información de la matriz $\mathbf{S}$. Los valores que se encuentran fuera del intervalo definido por (percentil 25 - 1.5 veces el rango intercuartílico) y (percentil 75 + 1.5 veces el rango intercuartílico) serán considerados como valores atípicos.

IV. EXPERIMENTACIÓN Y RESULTADOS

A. Conjunto de datos

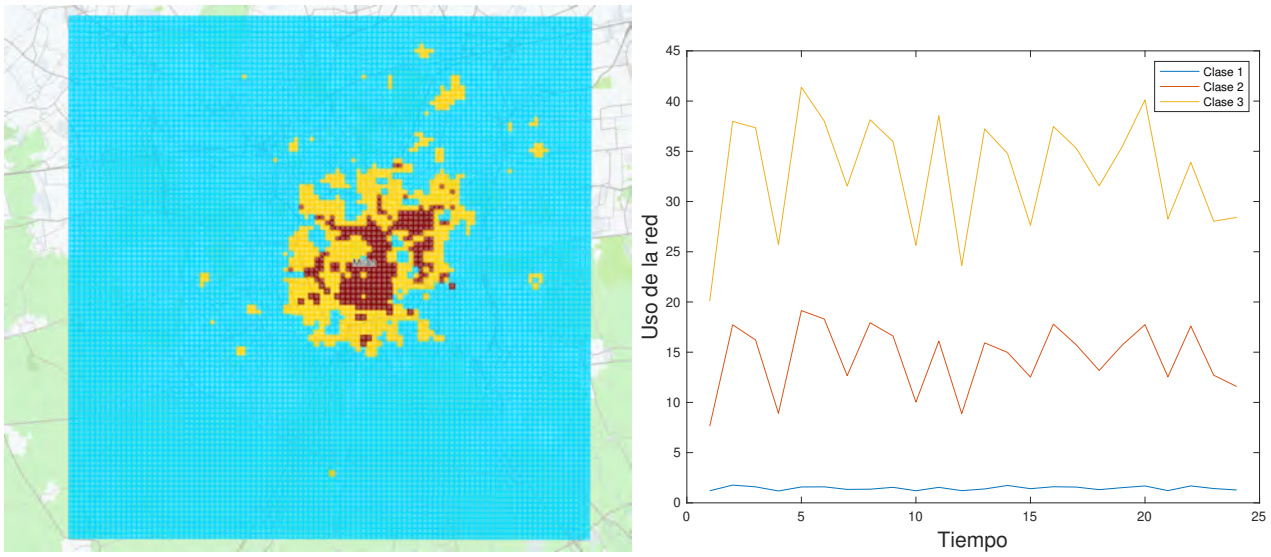
El trabajo se ha realizado sobre un conjunto de datos CDR que fue desarrollado por Telecom Italia en colaboración con varias empresas y publicado a través del primer desafío de Big Data [9]. El conjunto de datos liberado para este desafío es el primer conjunto abierto para el uso en investigación por parte de un operador y únicamente incluye información de telecomunicaciones y datos sociales de la ciudad de Milán y la provincia de Trento, capturando la información durante 2 meses y conteniendo los siguientes atributos:

- SMS recibidos
- SMS enviados
- Llamadas entrantes
- Llamadas salientes
- Conexión a internet

El conjunto de datos utilizados es el más completo liberado por un operador de telecomunicaciones para su uso público. Su importancia radica en el hecho de que contiene información real sobre el uso de la red móvil que se puede utilizar para desarrollar nuevas metodologías. El análisis desarrollado en este trabajo se va a centrar en las áreas metropolitanas y suburbanas de Milán. En el conjunto de datos liberado, el área geográfica se divide en una cuadrícula regular de 100×100 celdas cuadradas, con un tamaño de 235 metros por lado, que cubre un total de 552 km^2. El período de muestreo es de 10 minutos, lo que da lugar a un total de 8784 muestras para cada celda. Para reducir el número de intervalos, en este trabajo se hará un análisis horario de los datos, de forma que se ha calculado la media del uso de la red en cada una de las horas analizadas del conjunto de datos.

B. Clasificación de los comportamientos base

Partiendo del conocimiento adquirido al aplicar OPNA sobre el conjunto de datos, y buscar los comportamientos base a través de RPCA separando en dos diferentes conjuntos de datos los patrones de uso y los posibles



(a) (b)

Fig. 2. Análisis de las clases obtenidas a través del método k-means. La subfigura 2(a) muestra la clasificación sobre el mapa de Milán usando k=3 proyectado sobre un mapa real de Milán. Se pueden observar las tres clases detectadas (zona centro ciudad - roja, zona residencial del centro de la ciudad y partes centrales de las ciudades dormitorio - Amarillo, resto de zonas - azul. La subfigura 2(b) muestra los centroides encontrados para cada una de las clases analizadas.

valores atípicos, se hace necesario clasificar todos los "comportamientos" extraídos de los patrones del conjunto de datos para descubrir si existen relaciones entre las diferentes subáreas.

Como se ha descrito en la sección anterior, para realizar la clasificación se utilizará el algoritmo k-means, asignando al parámetro k el mejor valor que se ajuste al número de clústeres que conformen el conjunto de datos analizados (en este caso son 10000 celdas que conforman el área metropolitana de Milán y sus alrededores). Se han realizado 100 repeticiones de la selección del parámetro k para evitar los efectos de la semilla aleatoria usada por k-means a la hora de inicializar los centroides. En todas las repeticiones, usando el criterio de evaluación de clústeres de Calinski-Harabasz el número de clústeres óptimo es 3.

De esta forma, una vez clasificados todos los puntos del conjunto de datos entre las tres clases, se puede observar en la Figura 2 la asignación de cada una de las subáreas a una clase.

En la clasificación existen tres clases bien diferenciadas que están muy ligadas con el tipo de subárea que describen. Gracias a la información que proporcionan los centroides, en la Figura 2(b), se puede observar la tendencia del uso de la red, ya que estos vectores no representan un comportamiento real, sino un valor que ha calculado el algoritmo y que tiene la menor distancia con el resto de puntos clasificados en el mismo clúster, pero aporta información con respecto al uso promedio de la red.

De forma general, la clase 1 está principalmente localizada en las subáreas de las afueras de Milán y es la zona con menor carga de la red debido, principalmente, a que estas áreas son zonas agrícolas o poblaciones pequeñas. La clase 2 está relacionada con las zonas de viviendas de Milán y las ciudades dormitorio que se encuentran a las afueras. Estas zonas tienen una carga de la red más alta que las zonas clasificadas como clase 1, localizadas en las afueras.

Por último, la clase 3 está muy relacionada con la zona centro de Milán y la zona residencial más poblada de la ciudad. Estas subáreas son las que más carga de red tienen debido, principalmente, a la densidad de usuarios que viven o hacen uso de los servicios de red, por ejemplo, turistas en las zonas más características de la ciudad.

Como conclusión de este análisis se puede desprender que, separando las pequeñas diferencias que separan las diferentes subáreas, el comportamiento base es muy similar dependiendo de la cantidad de usuarios potenciales que se tengan en un momento dado del día. Con esta información es muy sencillo implementar patrones en la red que puedan reservar recursos para satisfacer las necesidades de los usuarios de forma automática.

C. Análisis de puntos clave y detección de valores atípicos

Una vez que se ha analizado, de forma general, la información contenida en el conjunto de datos, se va a analizar un punto clave bien conocido para detectar valores atípicos, partiendo de la búsqueda de los datos de baja dimensionalidad y los datos clasificados por el algoritmo RPCA como datos distribuidos de forma "aleatoria".

Concretamente, nos vamos a centrar en el estadio de futbol Giuseppe Meazza (también conocido como San Siro) para analizar los resultados obtenidos con la metodología presentada. El estadio se encuentra en una zona a las afueras de Milán, rodeado por otros recintos deportivos como un hipódromo y zonas verdes, tal y como se puede observar en la Figura 3. La subárea analizada es la celda que ocupa la mayor parte del estadio y que está remarcada de color azul.

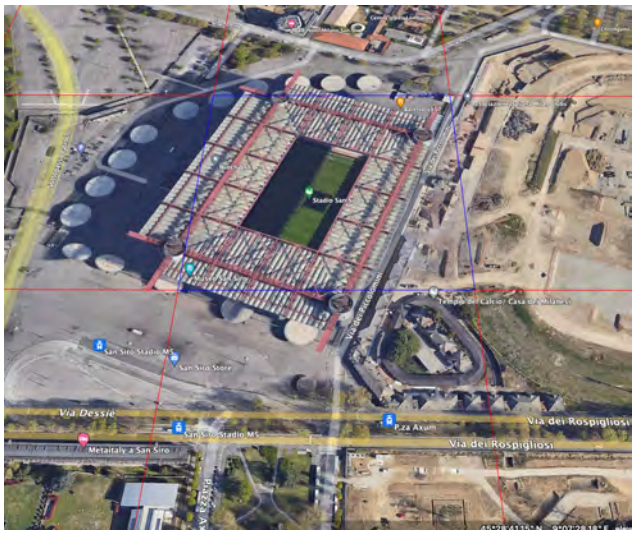


Fig. 3. Localización del Estadio Guiseppe Meazza

Este punto clave ha sido bien analizado en los trabajos anteriores [3] y [4], donde se describe los diferentes eventos que generan valores atípicos del uso de la red.

La Tabla I muestra los eventos de los que se tienen constancia y han sido validados por la información pública dispuesta en Internet.

La celda analizada está clasificada como clúster 1, es decir, tiene una carga de red muy baja durante todo el día, con lo que es fácil relacionar los valores atípicos con los eventos sucedidos en esa subárea.

La Figura 4 muestra el análisis de la información de la subárea ocupada por el estadio en la matriz $\mathbf{S}$ a lo largo del tiempo. Como se puede observar, hay algunos valores aleatorios con pequeña carga de uso de la red que pueden ser debidos a la conexión esporádica de cualquier dispositivo móvil a la red, pero existe un conjunto de valores atípicos que claramente incrementan la carga de la red. A través del análisis usando el método IQR, se obtiene que el mecanismo de detección de valores atípicos tiene una precisión del 81% al detectar 9 de los 10 eventos recogidos en la Tabla I, mientras que también detecta 2 eventos no recogidos en esa tabla. Estos resultados obtienen la misma precisión que los presentados en el trabajo [3] usando diferentes métodos para detectar los valores atípicos de la red.

Tabla I
LISTA DE EVENTOS CONOCIDOS DURANTE EL PERIODO ANALIZADO EN EL ESTADIO GUISEPPE MEAZZA.

Fecha	Evento
02/11/13	AC Milan - AC Fiorentina
09/11/13	FC Inter. Milano - Livorno
15/11/13	Italy - Germany
23/11/13	AC Milan - Genoa
01/12/13	FC Inter. Milano - UC Sampdoria
04/12/13	FC Inter. Milano - Trapani
08/12/13	FC Inter. Milano - Parma
11/12/13	AC MIlan - Ajax
16/12/13	AC Milan - AS Roma
22/12/13	FC Inter. Milano - AC Milan

V. CONCLUSIONES

En este trabajo se ha analizado la aplicación de una herramienta matemática bien conocida en otros campos para separar la información básica de cada subárea de los datos atípicos en el conjunto de datos, con el objetivo de detectar eventos en la red. Para ello se ha utilizado un conjunto de datos de telecomunicaciones almacenados en una base de datos CDR proporcionada por Telecom Italia. Gracias al conocimiento adquirido en análisis previos, mediante la metodología OPNA para la extracción de los "comportamientos" presentes en cada celda del conjunto de datos analizado, se puede buscar el patrón base de esa subárea que permita definir cómo se va a utilizar de forma general la red en esa zona. Así mismo, esa separación permitirá encontrar eventos de interés para que los gestores y administradores de la red estudien la motivación que ha producido el aumento puntual en el uso de los recursos.

Gracias a este trabajo y, como línea futura, se puede usar este conocimiento adquirido para entrenar sistemas inteligentes en la red que, una vez comiencen a detectar eventos que requieran de una mayor cantidad de recursos en la red, permitan reconfigurarla para satisfacer esos recursos sin necesidad de sobre aprovisionar la red, llegando a una configuración de red basada en *zero-touch*, o sin intervención humana.

En este trabajo se ha llevado un análisis novedoso sobre unos datos bien estudiados, de forma que ha permitido validar la capacidad del método propuesto a la hora de encontrar un "patrón base" de la red. Además, a través de un análisis en un punto de interés bien conocido, se ha realizado un estudio que demuestra el potencial de esta metodología a la hora de encontrar eventos de interés en las redes de próxima generación.

AGRADECIMIENTOS

Este trabajo ha sido financiado, en parte, por la Unión Europea NextGenerationEU/PRTR, con el proyecto TED2021-131699B-I00 (AEI/FEDER,UE), por el Ministerio de Ciencia e Innovación, con el proyecto PID2020-112545RB-C54 y la Univ. Rey Juan Carlos (Ref. F920).

REFERENCIAS

[1] "Cisco Systems, Inc., Cisco Annual Internet Report (2018–2023)," March 2020. [Online]. Available: http://www.cisco.com/

[2] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Network*, vol. 34, pp. 134–142, 5 2020.

[3] D. Cortes-Polo, L. I. Jimenez, M. E. Paoletti, J. Calle-Cancho, and J. A. Rico-Gallego, "Orthogonal projection for anomaly detection in networking datasets," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2023.

[4] D. Cortés-Polo, L. I. J. Gil, J. Calle-Cancho, and J.-L. González-Sánchez, "A novel methodology based on orthogonal projections for a mobile network data set analysis," *IEEE Access*, vol. 7, pp. 158 007–158 015, 2019.

[5] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, jun 2011. [Online]. Available: https://doi.org/10.1145/1970392.1970395

[6] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

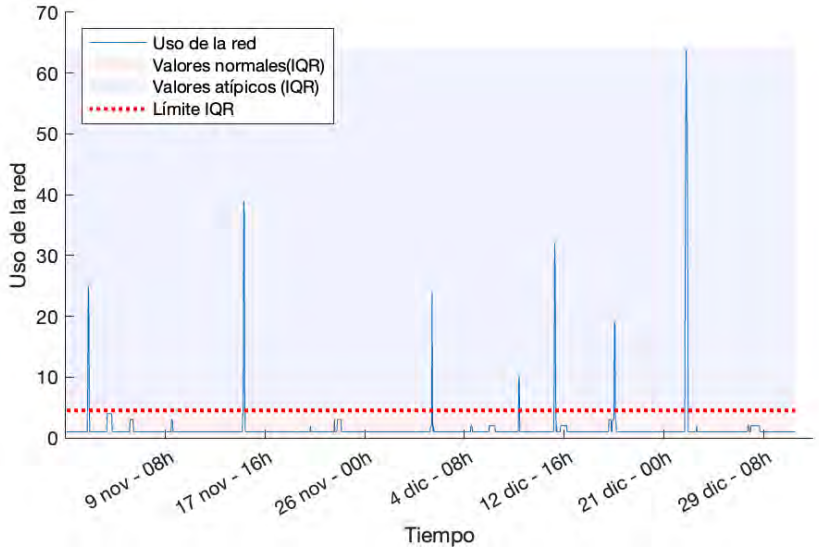


Fig. 4. Análisis de la celda 5738, referente al Estadio Guiseppe Meazza usando el cálculo del rango intercuartílico (IQR)

[7] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/03610927408827101

[8] M. Seyedan, F. Mafakheri, and C. Wang, "Cluster-based demand forecasting using bayesian model averaging: An ensemble learning approach," *Decision Analytics Journal*, vol. 3, p. 100033, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772662222000066

[9] G. Barlacchi, M. De Nadai, R. Larcher, A. Casella, C. Chitic, G. Torrisi, F. Antonelli, A. Vespignani, A. Pentland, and B. Lepri, "A multi-source dataset of urban life in the city of Milan and the Province of Trentino," *Scientific data, Nature*, vol. 2, no. 1, pp. 1–15, 2015.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Estrategias de offloading en arquitecturas Fog-Cloud: Un esquema basado en Lyapunov

Neco Villegas*, Luis Diez*, Idoia de la Iglesia†, Marco González-Hierro†, Ramón Agüero*
*Departamento Ingeniería de Comunicaciones, Universidad de Cantabria
e-mail: villegasn@unican.es, {ldiez, ramon}@tlmat.unican.es
†IoT and Digital Platforms Department, Ikerlan Technology Research Centre
e-mail: {idelaiglesia, marco.gonzalez}@ikerlan.es

En este trabajo se introducen políticas de offloading para arquitecturas Fog-Cloud que tienen en cuenta diferentes parámetros de rendimiento. Se afronta el diseño y desarrollo de una plataforma de tres niveles, utilizando técnicas de virtualización, que se puede utilizar para desplegar escenarios con nodos que tienen diferentes características, imitando aquellas de elementos Fog y Cloud. A continuación, se emplea la Teoría de control de Lyapunov para introducir políticas de offloading que equilibren el consumo de energía en los nodos de Fog y el coste monetario de utilizar el Cloud. El esquema propuesto es capaz de encontrar un equilibrio entre estos dos parámetros, garantizando al mismo tiempo la estabilidad del sistema y los requisitos de retardo. A continuación, se compara el algoritmo propuesto con soluciones de referencia (round-robin) y los resultados demuestran que es capaz de ofrecer un mejor rendimiento, incluso en situaciones de elevada demanda y requisitos de energía estrictos. Ajustando los parámetros operativos del algoritmo, los resultados obtenidos demuestran que es capaz de adaptar su comportamiento a diferentes objetivos, evaluando su rendimiento en configuraciones realistas.

***Palabras Clave*—fog, cloud, offloading, Lyapunov, energía, modelado**

I. INTRODUCCIÓN

Hoy en día, el número de servicios en la nube está aumentando continuamente, especialmente aquellos ofrecidos por los principales proveedores, como Microsoft Azure, Amazon Web Services (AWS) y Google Cloud. Este incremento en la demanda viene motivado, entre otras razones, por el aumento de servicios de Internet of Things (IoT) e Industrial Internet of Things (IIoT). Al mismo tiempo, existe una creciente implantación de redes 5G, cuyas tecnologías subyacentes ofrecen varias ventajas, por ejemplo, en términos de latencia, disponibilidad o confiabilidad. Por todo ello, son muchos los sectores que ven ahora una oportunidad para implementar diferentes servicios IoT e IIoT. De hecho, junto con la masiva implementación de redes celulares 5G, el número de conexiones IoT ya ha alcanzado los $14.6 \cdot 10^3$ millones y se espera que supere los $30 \cdot 10^3$ millones en 2027[1]. La gran cantidad de datos generados por estos dispositivos requiere encontrar una arquitectura de sistema adecuada, que sea capaz de acometer su procesado y almacenamiento.

Como resultado, existe un interés creciente en una integración de servicios IoT con Cloud Computing, puesto que la combinación de estas dos tecnologías presenta un gran potencial. Sin embargo, el fuerte incremento de servicios IoT e IIoT y su implementación en nuevos sectores con requisitos más estrictos, pueden llevar a situaciones en las que el Cloud Computing no sea suficiente. En este contexto el Fog Computing ha surgido como su extensión natural que acerca los recursos computacionales a los dispositivos.

El Cloud Computing proporciona alta disponibilidad de recursos computacionales con un consumo de energía relativamente alto (centro de datos), mientras que el Fog Computing proporciona disponibilidad moderada de recursos con un consumo de energía más bajo (servidores pequeños, routers, switches, gateways, etc.). Los entornos Cloud y Fog se pueden usar de manera independiente, pero ambas soluciones se complementan entre sí y, por tanto, la cooperación entre ellas conduce a un uso óptimo de los recursos. Este enfoque conduce a arquitecturas de tres niveles (IoT-Fog-Cloud).

Las principales contribuciones de este trabajo se resumen brevemente a continuación:

1) Utilizando una arquitectura IoT-Fog-Cloud de tres niveles, se propone un esquema de offloading, asignación dinámica de cargas de trabajo, que considera el consumo de energía y el coste monetario en un entorno aleatorio y no controlado.

[1]https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/iot-connections-outlook

2) Se aborda el problema de optimización estocástica resultante mediante la aplicación de la Teoría de Lyapunov, de modo que se reduce a un problema de estabilización del sistema de colas, que puede resolverse con el algoritmo "drift-plus-penalty", esto es, una secuencia de problemas de programación lineal entera (ILP, por sus siglas en inglés).
3) Se realiza un análisis exhaustivo del esquema propuesto en diferentes escenarios y bajo condiciones heterogéneas.

El resto del documento se estructura de la siguiente manera. En la Sección II, se discuten los trabajos existentes relacionados con la combinación de IoT, Fog y Cloud, y otros algoritmos de offloading, señalando en qué se diferencia la propuesta que aquí se presenta. En la Sección III, se describe el modelo del sistema y la solución propuesta para el algoritmo de offloading. A continuación, en la Sección IV se describe la plataforma desplegada para llevar a cabo la evaluación, mientras que en la Sección V se analiza el rendimiento de la solución propuesta. Por último, la Sección VI concluye el documento, resumiendo los aspectos fundamentales, y proporciona una perspectiva del trabajo futuro que surge a partir de la metodología diseñada.

II. TRABAJOS RELACIONADOS

La combinación de IoT e IIoT con Fog y Cloud Computing ha atraído recientemente la atención de la comunidad científica desde diferentes ángulos. Con una perspectiva global, algunos trabajos han propuesto arquitecturas adaptadas a estos entornos [1]–[3]. Aunque estos trabajos comparten el mismo escenario de aplicación que el presentado en este documento, su ámbito de aplicación se sitúa a nivel de arquitectura, mientras que el principal interés de la propuesta descrita se sitúa en la lógica de offloading del procesado de tareas de cómputo y soluciones algorítmicas para obtener comportamientos óptimos. Otros trabajos se han centrado más específicamente en enfoques de offloading [4]–[8].

A diferencia de estos trabajos, la propuesta aquí descrita se centra no sólo en el consumo energético, sino también en el coste monetario, manteniendo la estabilidad de las colas y reduciendo así el retardo. Además, al igual que otras soluciones que utilizan la Teoría de Lyapunov, también tiene en cuenta la evolución temporal del escenario, pudiendo producirse eventos aleatorios. En la Tabla I se compara la propuesta de este trabajo con enfoques similares de la literatura, al menos en sus objetivos. Se seleccionan aquellas soluciones que asumen entornos aleatorios (no controlados) y proponen técnicas para proporcionar una adaptación instantánea. La comparación se realiza en términos del algoritmo de decisión y de los parámetros y métricas de rendimiento que considera. Se incluyen los que se enumeran a continuación:

- **Retardo.** Se refiere al retardo que sufren las tareas o servicios de computación, desde que se generan hasta que se procesan completamente.

Tabla I: Parámetros y métricas de rendimiento considerados en algoritmos de la literatura reciente con escenarios similares al algoritmo propuesto.

	Algoritmo	Retardo	Energía	Coste	Estabilidad
[4]	Optimización distribuida basada en ADMM.	✓	✓	✗	✗
[5]	Programación estocástica basada en Lyapunov.	✓	✗	✗	✓
[6]	Programación estocástica basada en Lyapunov.	✓	✓	✗	✓
[8]	Offloading predictivo.	✓	✓	✗	✓
Prop.	Programación estocástica basada en Lyapunov.	✓	✓	✓	✓

- **Consumo de energía.** Es consecuencia, principalmente, del procesado. Se suele tener en cuenta para los nodos Fog, que tienen capacidades más limitadas.
- **Coste monetario.** Corresponde al coste de utilizar la capacidad de procesamiento del Cloud. Se considera un modelo de "pago por uso", puesto que es lo que ofrecen la mayoría de los proveedores (Amazon, Microsoft, IBM, Google, etc.).
- **Estabilidad.** Hace referencia a la estabilidad de las colas de memoria en el sistema global.

Como puede observarse, la solución propuesta es la única que considera conjuntamente la energía (en los nodos Fog) y el coste monetario (en los nodos Cloud), manteniendo la estabilidad del sistema. Por ello, es posible concluir que complementa y amplía el estado del arte relacionado con la distribución de tareas de computación en despliegues Fog-Cloud.

III. MODELO DE SISTEMA

Esta sección describe el modelo del sistema, así como el diseño del algoritmo propuesto basado en la Teoría de Lyapunov. En la Tabla II se resumen los símbolos utilizados en el modelo propuesto y su significado. Se utiliza el término "servicio" para referirse a conjuntos de paquetes sobre los que se necesita aplicar cierta computación.

Se considera un sistema compuesto por nodos Fog y Cloud con diferentes capacidades de procesamiento. En este escenario, múltiples aplicaciones de usuario generan servicios, compuestos por paquetes, que se envían a los nodos Fog. A continuación, los servicios pueden procesarse localmente (en los mismos nodos Fog) o ser enviados al Cloud. El sistema incluye también un orquestador, o nodo Master, que toma las decisiones de offloading, dependiendo de la política particular implementada.

Sea M el número de aplicaciones que generan servicios. Se supone que el tiempo transcurre en slots y que cada aplicación genera un servicio en cada slot. A su vez, el número de paquetes por servicio sigue una

Tabla II: Símbolos y variables del modelo del sistema.

Notación	Descripción
N	número de puntos de procesamiento, tales como CPUs en el Fog e instancias Cloud.
M	número de aplicaciones independientes generando servicios para ser procesados.
$a_m(t)$	llegadas a la cola de la aplicación m en el slot t, medido en paquetes.
$b_m(t)$	salidas desde la cola de la aplicación m en el slot t, medido en paquetes.
$Q_m(t)$	tamaño de la cola de la aplicación m en el slot t, medido en paquetes.
$\alpha_{m,n}(t)$	decisión para la aplicación m y la CPU n en el slot t, medido en número de paquetes.
$\alpha(\mathbf{t})$	$M \times N$ matriz de las variables de decisión.
$\mathcal{A}(t)$	conjunto de decisiones admisibles en el slot t.
$\omega_n(t)$	tasa de transferencia de la opción de procesamiento n en el slot t. Refleja la variación de la capacidad de procesamiento disponible.
$g_m(t)$	complejidad de procesado de los servicios de la aplicación n en el slot t.
$C(t)$	coste monetario por el uso del Cloud en el slot t.
$k_n(t)$	coste genérico de usar el punto de procesamiento en el slot t.
$E_n(t)$	coste de energía del punto de procesamiento n en el slot t.
E_nth	umbral de energía del punto de procesamiento n.
$G_n(t)$	cola virtual relativa al consumo de energía en el punto de procesamiento n en el slot t.
$\hat{\bullet}$	se utiliza para indicar una función arbitraria que produce una variable $\bullet$.

cierta distribución aleatoria. Los paquetes de los servicios generados se almacenan localmente en las colas de las aplicaciones. Se supone que el escenario tiene N alternativas de procesamiento, incluyendo procesadores locales (alternativas $1, \ldots, N-1$) y un Cloud (alternativa N). En cada slot el nodo Master establece la cantidad de datos de cada aplicación a procesar en cada alternativa, satisfaciendo algunas restricciones. La política de offloading debe garantizar que las colas de aplicaciones permanezcan estables, con el fin de evitar incrementar el retardo. En este escenario, se propone usar la Teoría de Lyapunov, ampliamente utilizada en optimización estocástica para garantizar la estabilidad del sistema.

Sea $a_m(t)$ la cantidad de paquetes que llegan a la cola de la aplicación m, $m \in \{1, \ldots, M\}$, en el slot t y $b_m(t)$ el número de paquetes salientes como consecuencia de la política que se aplica. La dinámica de colas viene dada por la Ec. (1), donde $Q_m(t+1)$ es la cola de espera de la aplicación m en el slot t.

$$Q_m(t+1) = \max[Q_m(t) - b_m(t), 0] + a_m(t) \quad (1)$$

El objetivo es garantizar la estabilidad (promedio) de las colas de aplicaciones, según se describe en la Definición 1.

Definición 1 (Estabilidad de la tasa media): Una cola es estable en tasa media si:

$$\lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\{Q(t)\} = 0 \quad (2)$$

donde $Q(t)$ es el tamaño de la cola en el slot t y $\mathbb{E}$ es la esperanza matemática.

Sea $\alpha(t)$ una matriz MxN, tal que el elemento $\alpha_{m,n}(t)$ se corresponde con la cantidad de datos de la aplicación m que se asigna al procesador n en el slot t. En cada slot se toma una decisión $\alpha(t)$, dentro de un conjunto $\mathcal{A}(t)$ de posibles elecciones. Además, se supone que existe una tasa de servicio (velocidad de procesado de datos) para cada opción, que dicta cuántos bytes se pueden aceptar en un slot determinado. Es posible definir $b_m(t)$ según la Ec. (3), donde $\omega(t)$ es la tasa de servicio de cada procesador en el slot t, en bytes por slot. En general, se supone que la tasa varía con el tiempo, siguiendo una distribución arbitraria. Como puede observarse, la cantidad de datos drenados por cada aplicación es función de la decisión y de la tasa de servicio de los procesadores. Cabe señalar que esta última puede verse influida tanto por la capacidad de cálculo de la CPU como por la capacidad de comunicación entre la cola de aplicación y el procesador. Por ejemplo, en el procesamiento local la tasa de servicio estaría dominada por la capacidad de cálculo, mientras que en el caso de un procesado remoto (datos que deben enviarse al Cloud) estaría limitada por la capacidad de comunicación.

$$b_m(t) = \hat{b}(\alpha(t), w_1(t), w_2(t), \ldots) = \hat{b}(\alpha(t), \omega(t)) \quad (3)$$

Para evitar la asignación de paquetes inexistentes, en cada slot se impone que la cantidad total de bytes asignados de una cola de aplicación i en la slot t no supere los bytes que hay en dicho instante, como se indica en la Ec. (4).

Además, se asegura que la asignación no supere la tasa de servicio, con la restricción definida en la Ec. (5), donde $g_i(t)$ denota un factor de escala genérico. Por ejemplo, en el caso de una tasa de servicio limitada por la capacidad de cálculo, este parámetro estaría relacionado con la complejidad de dicho cálculo. De este modo, los servicios más complejos producirían tiempos de cálculo más lentos para el mismo número de bytes, lo que se representa escalando el número de bytes, mientras se mantiene constante la capacidad de cálculo.

$$\sum_{j=1}^{N} \alpha_{ij}(t) \leq Q_i(t) \quad \forall i \in \{1, \ldots, M\}, \forall t \quad (4)$$

$$\sum_{i=1}^{M} g_i(t) \cdot \alpha_{ij}(t) \leq w_j(t) \quad \forall t, \forall j \quad (5)$$

Ahora se puede reescribir la salida $b_m(t)$, como se muestra en la Ec. (6).

$$b_m(t) = \hat{b}(\alpha(t), \omega(t)) = \sum_{j=1}^{N} \alpha_{m,j}(t) \quad (6)$$

También se consideran las penalizaciones por utilizar las distintas alternativas de procesamiento. Se utiliza el símbolo $k_i(t)$ para denotar el coste de utilizar la alternativa de procesamiento i en el slot t. Las penalizaciones se definen para las CPU del Cloud y del Fog de formas distintas. Para el procesamiento en el Cloud el objetivo

es minimizar el coste monetario, que está relacionado con la potencia de cálculo necesaria, tal y como se define en la Ec. (7).

$$C(t) = \hat{C}\left(\sum_{i=1}^{M} g_i(t) \cdot k_N(t) \cdot \alpha_{i,N}(t)\right) \tag{7}$$

donde $g_i(t)$ se corresponde, como ya se ha mencionado, con la complejidad de cálculo de los servicios generados por la aplicación i, mientras que $k_N(t)$ es el coste monetario de utilizar el Cloud en el slot t. Como puede verse, el coste es proporcional a la cantidad de tráfico enviado a la instancia del Cloud, escalado por la complejidad computacional. En general, se supone que tanto la complejidad computacional de los servicios de cada aplicación como la tarifa del Cloud pueden variar con el tiempo, siguiendo distribuciones aleatorias arbitrarias.

En cambio, el procesamiento local en el Fog se ve penalizado por el consumo de energía. En este caso, no se busca minimizar este parámetro, sino garantizar que, en promedio, se mantenga por debajo de un determinado valor. Esto sería necesario, por ejemplo, para los dispositivos alimentados por baterías que pueden recargarse periódicamente. Se define la restricción energética de la Ec. (8),

$$E_j(t) = \sum_{i=1}^{M} g_i(t) \cdot k_j \cdot \alpha_{i,j}(t) \quad \forall j \neq N \tag{8}$$

donde k_j representa un mapeo general entre el número de bytes que hay que procesar, escalado por la complejidad del procesamiento, y la energía necesaria para dicho procesamiento. A diferencia del coste del Cloud, el coste asociado a la energía (k_j) dependería principalmente de las características del hardware del procesador, por lo que se supone que no varía con el tiempo: $k_j(t) = k_j \ \forall t \, \forall j \neq N$. En ambos casos, se tienen en cuenta las penalizaciones a lo largo del tiempo, por lo que se utilizan sus expectativas temporales promedio, $\overline{C}$ y $\overline{E}$, que se definen en las Ec. (9) y (10), respectivamente.

$$\overline{C} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \mathbb{E}\{C(t)\} \tag{9}$$

$$\overline{E} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \mathbb{E}\{E(t)\} \tag{10}$$

En conjunto, se busca una política de control que minimice el problema de optimización enunciado en el Problema 1.

Problema 1:

$$\min_{\alpha(t)} \quad \overline{C} \tag{11}$$

$$\textbf{s.t.} \quad \overline{E}_j \leq E_j^{Th} \quad \forall j \in \{1, \ldots, N-1\} \tag{12}$$

$$\alpha(t) \in \mathcal{A}(t) \tag{13}$$

donde E_j^{Th} es el umbral de energía definido para cada procesador de un nodo Fog y $\mathcal{A}(t)$ se cumple para el conjunto de restricciones definidas en las Ec. (4) y (5) en cada slot. Utilizando el entorno de optimización estocástica desarrollado en [9], las desigualdades relacionadas con la limitación del consumo de energía pueden convertirse en colas virtuales, al igual que las colas de aplicaciones definidas anteriormente. Con ello, la actualización de la cola virtual asociada a la energía del procesador de Fog j, G_j, se define en la Ec. (14).

$$G_j(t+1) = \max\{G_j(t) + (E_j(t) - E_j^{Th}), 0\} \tag{14}$$

La cola virtual se introduce como un método para garantizar que se satisface la restricción de consumo promedio de energía. Así, se puede definir el conjunto de colas (de aplicaciones y la cola virtual) como $\Theta(t)$. La función de Lyapunov $L(\Theta(t))$ y su deriva (drift) $\Delta(\Theta(t))$ se definen como se muestra en (15) y (16).

$$L(\Theta(t)) = \frac{1}{2}\left(\sum_{j=1}^{N} G_j(t) + \sum_{i=1}^{M} Q_i(t)\right) \tag{15}$$

$$\Delta(\Theta(t)) = \mathbb{E}\{L(\Theta(t+1)) - L(\Theta(t)) | \Theta(t)\} \tag{16}$$

La solución al Problema 1 es el algoritmo drift-plus-penalty. En cada slot t, se observa el estado de las colas, y se toma una decisión que resuelve el Problema 2, donde V es un factor de ponderación positivo que establece el compromiso entre la deriva y la penalización. Se trata de un problema de programación lineal entera (ILP), que puede resolverse con herramientas existentes.

Problema 2:

$$\min_{\alpha(t)} \quad V \cdot C(t) + \sum_{i=1}^{M} Q_i(t)[a_i(t) - b_i(t)] + \tag{17}$$

$$\sum_{j=1}^{N} G_j(t)(E_j(t) - E_j^{Th}) \tag{18}$$

$$\textbf{s.t.} \quad \sum_{j=1}^{N} \alpha_{ij}(t) \leq Q_i(t) \quad \forall i \in \{1, \ldots, M\}, \forall t \tag{19}$$

$$\sum_{i=1}^{M} g_i(t) \cdot \alpha_{ij}(t) \leq w_j(t) \quad \forall t, \forall j \tag{20}$$

IV. PLATAFORMA DE EVALUACIÓN

Existen varias alternativas para desplegar y gestionar instancias de Fog y Cloud. La mayoría de las grandes empresas tecnológicas proporcionan servicios en la nube, como AWS, Azure, Linode, etc. Por otro lado, existen alternativas que permiten el despliegue de instancias Fog-Cloud propietarias y autogestionables, tanto comerciales (ej. VmWare) como open source (ej. OpenStack, Apache CloudStack, Proxmox). Sin embargo, estas tecnologías no están diseñadas para probar o evaluar el rendimiento de diferentes soluciones de orquestación, sino para gestionar servicios en ejecución. En este sentido, es necesario desarrollar frameworks que llenen el vacío existente entre la evaluación analítica y la planificación, y la evaluación del

rendimiento esperado en entornos controlados. Existen algunos trabajos relacionados en los que se han desarrollado plataformas de arquitectura de tres niveles como [10]–[12], pero tenían algunas limitaciones para los objetivos anteriormente descritos.

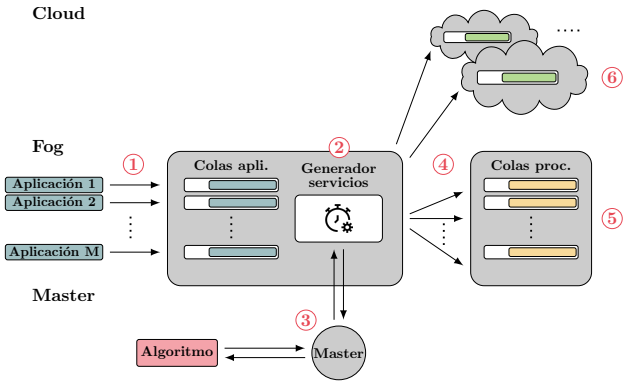


Fig. 1: Vista general de la plataforma Fog-Cloud.

Por ello, se ha desarrollado una plataforma, cuyo diseño se muestra Fig. 1. Abarca tres tipos de elementos que imitan Fog, Cloud y un nodo Master. Los nodos Fog generan flujos de tráfico sintéticos independientes, pertenecientes a diferentes aplicaciones, para lo que se utiliza distribuciones aleatorias configurables (Poisson, uniforme, Lognormal, etc.). El tráfico generado de cada aplicación se almacena en un búfer de entrada, que se representa en la Fig. 1 como el paso 1. A partir del tráfico generado, los nodos Fog definen servicios (tareas de procesamiento) que engloban una serie de paquetes. La generación de servicios, representada en la Fig. 1 con el paso 2, también es configurable, utilizando distribuciones aleatorias. Cuando se definen los servicios, el nodo Fog consulta al Master y este indica el punto de procesado según el algoritmo implementado (paso 3). A continuación, el nodo Fog envía los datos del servicio a procesadores locales o a instancias remotas (paso 4). Por último, se generan registros del rendimiento del sistema (pasos 5 y 6) para cada servicio y la evolución temporal de los estados de todos los dispositivos. Cabe destacar que la plataforma implementa una interfaz genérica para comunicarse con el nodo Master, que es independiente del algoritmo de decisión adoptado. Además, el algoritmo del nodo Master se implementa como un plugin, fácilmente modificable, permitiendo así la comparación de diferentes esquemas de decisión bajo las mismas circunstancias.

Para garantizar una plataforma escalable y ligera todos los nodos se han desplegado en contenedores utilizando Docker. Esto permite a los usuarios desplegar rápidamente múltiples contenedores personalizados en el mismo host. En esta plataforma, cada uno de los contenedores funciona como un nodo Fog, Cloud o Master.

V. RESULTADOS

En esta sección se lleva a cabo una campaña de experimentos utilizando la plataforma desarrollada para analizar el rendimiento de la propuesta de algoritmo de offloading descrito previamente. En primer lugar, se estudian las propiedades del algoritmo en dos escenarios sintéticos, con y sin opción de procesamiento en el Cloud. De esta forma, el primer escenario se centra en el equilibrio entre las limitaciones relacionadas con el consumo de energía y el coste monetario, mientras que la última configuración presta especial atención al impacto que la limitación de energía puede tener sobre el tráfico entrante de la aplicación. A continuación, se evalúa el esquema propuesto en un tercer escenario más realista, en términos de capacidad de procesamiento y generación de tráfico.

Tabla III: Configuración común de la plataforma para la campaña de simulaciones.

Parámetro	Valor
Slots simulados	1000
Duración del slot	1 s
Capacidad de procesamiento de una CPU en el Fog	1 kB/s
Capacidad de procesamiento en un Cloud	100 kB/s
Longitud de un paquete	$200 + 12$ bytes
Tasa de tráfico media agregada	Poisson $[6, \ldots, 30]$ pkt/s
Tasa de generación de servicios	1 serv/s
Factor de energía del Fog (k_j)	1
Coste del Cloud (k_N)	1
Complejidad de la aplicación (g_m)	1
Umbral de energía	$[1, \ldots, 5]$
Número de aplicaciones	3
Número de CPUs en un nodo Fog	2

En estas tres configuraciones, el rendimiento de la solución propuesta se compara con el observado con un algoritmo simple basado en round-robin. Para todas las pruebas se opta por configurar un único nodo Fog con tres aplicaciones y dos procesadores, un nodo Cloud, y un nodo Master que ejecuta los algoritmos. Los detalles de la configuración base que se ha utilizado para las dos primeras pruebas se muestran en la Tabla III. Como puede observarse, en todos los casos se realizan ejecuciones que abarcan 1000 slots de 1 segundo cada uno. En cada slot, las aplicaciones generan un servicio consistente en un número aleatorio de paquetes. En la Tabla III se muestra la tasa media agregada de las aplicaciones y, en cada escenario, se especificará la tasa particular de cada aplicación. Como se puede ver, algunas variables que en el modelo pueden ser aleatorias se definen como constantes, para simplificar la interpretación de los resultados, aunque la utilización de otras opciones no supondría ninguna modificación de la solución propuesta.

A. *Colaboración Fog y Cloud*

Con la primera configuración se pretende analizar el equilibrio de procesamiento entre los nodos Fog y Cloud en diferentes configuraciones. Cabe destacar que la capacidad de procesamiento del nodo Cloud, teniendo en cuenta el tráfico generado por las aplicaciones y la alta capacidad que típicamente tienen los centros de datos, se considera infinita. Se refleja así que, en cualquier caso, será siempre significativamente mayor que la de los nodos Fog.

En primer lugar, se procede a evaluar el impacto del parámetro V, el cual ajusta el compromiso entre con-

sumo de energía y coste monetario. La Fig. 2 muestra la proporción de tráfico enviado al nodo Cloud en función de distintos valores de la tasa de tráfico media agregada. En esta configuración, se fija el valor umbral de energía, E_{th}, en 2. Además, se realizan simulaciones para distintos valores de V. La figura también muestra, con líneas discontinuas, los resultados obtenidos al utilizar el algoritmo round-robin (RR) y una variante de round-robin que tiene en cuenta el umbral de energía prefijado (RRe). Así, la primera opción de round-robin consume toda la capacidad de procesamiento del Fog, enviando el excedente al Cloud. La segunda hace uso de los procesadores del nodo Fog sin superar el umbral de energía, enviando el resto de paquetes al Cloud.

Como era de esperar, se puede observar que una mayor tasa de tráfico conlleva un mayor uso del Cloud. Además, a medida que aumenta el valor de V, vemos una tendencia decreciente en el uso del Cloud. Se identifican 3 regiones de operación en función de dicho parámetro, delimitadas por los algoritmos round-robin. En la primera región, se puede ver que se envía más tráfico al Cloud que con el algoritmo RRe. Esto ocurre con valores de V inferiores a 1, donde se da muy poco peso al coste monetario. El resultado es que no se utiliza la máxima capacidad de procesamiento disponible en el Fog, aunque no se alcance el umbral de energía. La segunda región corresponde a valores de V comprendidos entre 2 y 1000, y el rendimiento observado se sitúa entre las dos versiones round-robin. En esta región, los valores más altos de V reducen significativamente el uso del Cloud y, a su vez, conducen a una eventual saturación de los procesadores del Fog. Con el objetivo de no rebasar el umbral de energía, las soluciones propuestas mantienen el tráfico en las colas de aplicaciones y equilibran las decisiones para garantizar la estabilidad del sistema, teniendo en cuenta las colas de aplicaciones, la energía y el coste.

Este efecto se analiza además estudiando cómo afectan las distintas configuraciones al indicador de rendimiento energético. En la Fig. 3 se representa, con un diagrama de barras, el consumo energético promedio producido por las soluciones propuestas para distintas configuraciones de la tasa de tráfico agregada y para diversos valores de V. Como puede observarse, cuando V se encuentra dentro de la primera región observada en la Fig. 2 ($V = 1$), el consumo está muy por debajo del umbral, independientemente de la tasa de tráfico. A medida que se aumenta el valor de V se observa que el esquema propuesto no es capaz de mantener la energía por debajo del umbral, debido al alto coste de uso de la instancia Cloud. Además, los resultados muestran que el impacto de V también varía con la tasa de tráfico. En este sentido, el consumo de energía se satura con $V = 1e3$ para la tasa de tráfico más baja (6 $\texttt{pkt/s}$), mientras que este valor de saturación aumenta para tasas más altas.

Este primer conjunto de resultados valida el adecuado comportamiento del esquema propuesto, mostrando que es capaz de equilibrar la carga computacional, considerando diferentes parámetros (energía, coste y colas de aplicaciones). Además, se puede configurar para fomentar diferentes comportamientos, gracias al parámetro de operación V.

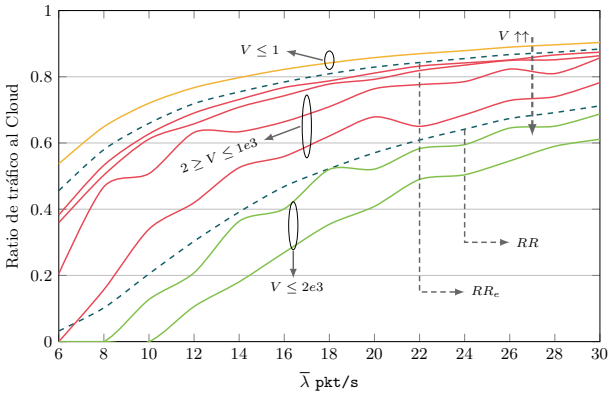


Fig. 2: Uso del Cloud frente a la variación de la tasa de tráfico agregada.

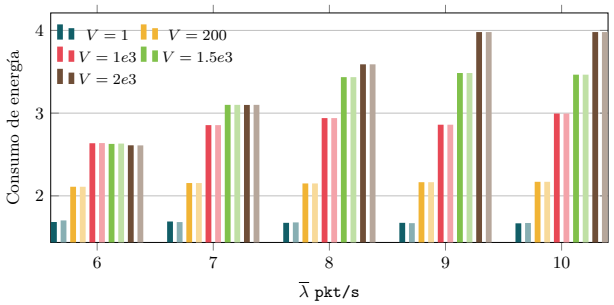
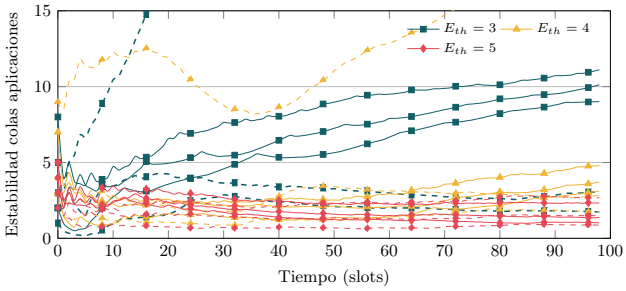


Fig. 3: Coste de energía promedio frente a la variación de la tasa de tráfico agregada.

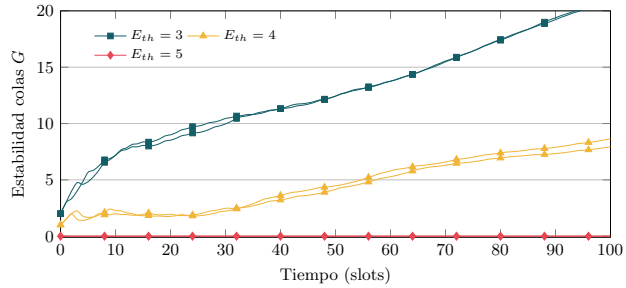
B. Análisis de rendimiento en Fog

Los resultados que se presentan a continuación se centran, con mayor detalle, en el impacto que tienen las distintas configuraciones sobre la energía y las colas de aplicaciones en el Fog. Como se ha visto en la sección anterior, valores bajos de V evitarían la sobrecarga de los nodos Fog, puesto que obligan a enviar muchos servicios al Cloud. Teniendo esto en cuenta, se considera una situación en la que el Cloud y su alta capacidad de procesamiento no estén disponibles, lo que se correspondería con un valor de V alto. Sin el Cloud, es posible evaluar configuraciones más críticas en términos de capacidad de procesamiento, lo que permite observar más de cerca la estabilidad de las colas de aplicaciones.

En este caso, la tasa media agregada de tráfico se fija en 7 paquetes por slot. En concreto, la primera, segunda y tercera aplicación generan 1, 2 y 4 paquetes por slot, respectivamente. La Fig. 4 muestra la estabilidad temporal de las colas de aplicaciones y energía, utilizando la Ec. (2), para distintos valores del umbral de energía. Cabe señalar que umbrales altos son equivalentes a no imponer ningún límite al consumo de energía. Para una mejor representación de los resultados, se muestran los valores observados en los 100 primeros slots. La figura ilustra la estabilidad de la cola de las aplicaciones obtenida



(a) Colas de las aplicaciones. round-robin está representado con líneas discontinuas.



(b) Colas virtuales de energía.

Fig. 4: Evolución de las colas debido a la variación del umbral de energía.

con el algoritmo propuesto y la obtenida con RRe (líneas discontinuas). Los distintos colores corresponden a las estabilidades de las distintas colas de las aplicaciones.

En general, se puede observar que el algoritmo propuesto consigue mantener la estabilidad de todas las colas de las aplicaciones, independientemente de los límites establecidos en el consumo de energía. Como se pone de manifiesto, cuando se utiliza round-robin con el umbral de energía fijado en 3 y 4, hay una cola muy inestable, que corresponde a la aplicación con mayor tasa, mientras que las demás muestran valores bastante bajos. Por otra parte, la solución propuesta es capaz de adaptarse a las tasas de tráfico, como demuestra el hecho de que todas las colas presenten valores similares. Cuando se utiliza un umbral de 3, no se puede garantizar la estabilidad, pero, con restricciones energéticas más suaves, la estabilidad se alcanza rápidamente. La Fig. 4b muestra la estabilidad de la cola virtual de energía (G_j) de cada CPU del nodo Fog. Los resultados muestran una tendencia similar a la observada para las colas de aplicaciones. A medida que se relaja la restricción energética, el esquema propuesto es capaz de estabilizar ambos tipos de colas, mientras que penaliza aquellas con requisitos más estrictos.

C. Entorno realista de generación de tráfico

A continuación se amplía la evaluación utilizando un entorno más realista. En concreto, se ajustan las distribuciones de tráfico de las aplicaciones y, por tanto, las capacidades de los dispositivos. La configuración concreta se muestra en la Tabla IV. Los valores se han obtenido de [13], donde los autores caracterizaron la carga de trabajo de entrada/salida y la distribución de datos de AliCloud, uno de los mayores proveedores de Asia. Como se muestra en la Tabla IV, el tráfico sigue una distribución lognormal. En concreto, el valor esperado y la varianza de la distribución normal subyacente se fijan en 4.5 y 0.8, respectivamente. A su vez, se obtiene una tasa media de tráfico de 125 `pkt/s`[2] Además, se escala la capacidad del nodo Fog en función de las nuevas tasas de tráfico, de forma que pueda hacer frente cómodamente, en media, al tráfico generado por las tres aplicaciones. Así, el escenario configurado permite que el umbral de energía desempeñe un papel importante en los resultados de la simulación.

Tabla IV: Configuración de la simulación del entorno realista.

Parámetro	Valor
Slots simulados	1000
Duración del slot	1 `s`
Capacidad de procesamiento de una CPU en el Fog	100 `kB/s`
Capacidad de procesamiento en un Cloud	10^6 `kB/s`
Umbral de energía	$[75, \ldots, 120]$
Distribución de tráfico para cada aplicación	Lognormal $(4.5, 0.8)$ $(\log pkt/s)$

Bajo esta configuración se pretende identificar configuraciones óptimas del algoritmo según las limitaciones de coste monetario y consumo de energía asequibles. En este sentido, la Fig. 5 muestra una representación en dos ejes del coste monetario (eje izquierdo) y el consumo de energía (eje derecho) con líneas sólidas y discontinuas, respectivamente. Las líneas representan el valor medio obtenido a partir de 30 experimentos independientes, cada uno de los cuales dura 1000 slots. Junto con los valores medios, también se representan el máximo y el mínimo obtenidos durante las simulaciones (fondo sombreado en cada una de las líneas). Se muestran los resultados a medida que se aumenta el valor del umbral de energía E_{th} y para distintos valores del parámetro V.

Como era de esperar, al relajar el umbral de energía el coste disminuye, ya que se procesan más datos en el Fog, mientras que el consumo de energía crece. Las intersecciones corresponden a los puntos (configuraciones) en los que ambos costes son iguales. En este sentido, el esquema propuesto permite establecer configuraciones (V y E_{th}) que igualan los costes de energía y monetarios. Como puede observarse, para el escenario considerado, E_{th} debe fijarse entre 90 y 115 para todo el rango de valores de V. En la Tabla V se indican los puntos de intersección de más valores de V que no se incluyeron en la Fig. 5, para simplificar su representación gráfica. Se puede realizar un análisis similar para diferentes relaciones entre el consumo de energía y el coste del Cloud, o fijando el umbral de energía en lugar del coste del Cloud.

[2]El valor esperado de la distribución lognormal $\mathcal{X}$ viene dado por $\mathbb{E}\{\mathcal{X}\} = \exp(\mu + \frac{\sigma^2}{2})$, donde μ y σ^2 son el valor esperado y la varianza de la distribución normal $\mathcal{N}$, de modo que $log(\mathcal{X}) \sim \mathcal{N}(\mu, \sigma^2)$.

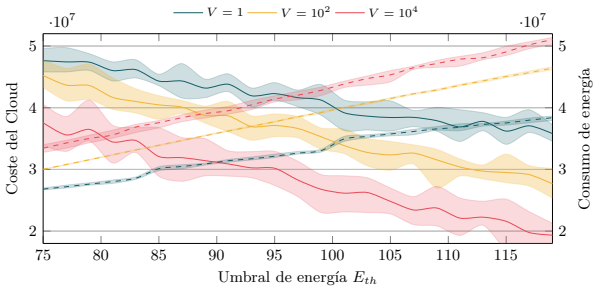


Fig. 5: Coste del Cloud y consumo de energía en función de V y E_{th}.

Tabla V: Puntos de intersección en función del parámetro V.

V	1	10	10^2	10^3	10^4
E_{th}	80.13	91.61	92.96	108.61	113.54
Coste ($\cdot 10^6$)	35.33	37.05	36.96	37.07	37.36

VI. CONCLUSIONES

El enorme volumen de datos generado por los dispositivos IoT precisa encontrar una arquitectura de sistema adecuada capaz de procesar y almacenar la gran cantidad de servicios desplegados. Si bien las arquitecturas basadas en el Cloud se utilizan actualmente con ese propósito, se vislumbra que el nuevo paradigma de Fog Computing permitirá escalar y optimizar las infraestructuras de IoT.

En este contexto se propone, en primer lugar, un modelo de sistema genérico que asume patrones de tráfico variables arbitrarios, capacidad de cálculo disponible y coste en el Cloud. A continuación, se formula un problema de optimización estocástica y se emplea la Teoría de Lyapunov para convertirlo en una secuencia temporal de problemas ILP, que pueden resolverse fácilmente con herramientas existentes. El esquema propuesto se aplica posteriormente a una variedad de escenarios Fog-Cloud, para validar su comportamiento bajo diferentes configuraciones del sistema. Los resultados demuestran que el esquema propuesto es capaz de equilibrar el uso de instancias de Fog y Cloud, consiguiendo regular el consumo de energía y el coste monetario debido al uso del Cloud. Además, se observa que la solución propuesta es capaz de adaptarse a cargas de tráfico desequilibradas, garantizando la estabilidad del sistema incluso bajo situaciones de elevada demanda o para restricciones de energía más estrictas.

Como líneas futuras se plantea ampliar el modelo de diferentes maneras. En primer lugar, se analizará el rendimiento al agregar funciones más complejas (ej. logarítmicas) al problema de optimización, para fomentar diferentes compromisos entre las limitaciones de costo y energía. Además, se analizará la posibilidad de tener en cuenta la ocupación de las colas de los procesadores dentro del modelo del sistema, haciendo que evolucione hacia una red de colas interconectadas, donde se puedan aplicar algoritmos de tipo back-pressure. En este sentido, otra línea futura que subyace es la aplicabilidad de esta estrategia, modificándola para que se base en el retardo de los servicios de computación.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, MINECO-FEDER, a través del proyecto SITED: *Semantically-enabled Interoperable Trustworthy Enriched Data-spaces (PID2021-125725OB-I00).*

REFERENCIAS

[1] M. Aazam, S. Zeadally, and K. A. Harras, "Deploying fog computing in industrial internet of things and industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4674–4682, 2018.

[2] C. Mouradian, D. Naboulsi, S. Yangui, R. H. Glitho, M. J. Morrow, and P. A. Polakos, "A comprehensive survey on fog computing: State-of-the-art and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 416–464, 2018.

[3] P. Bellavista, L. Foschini, and D. Scotece, "Converging mobile edge computing, fog computing, and iot quality requirements," in *2017 IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)*, 2017, pp. 313–320.

[4] Y. Xiao and M. Krunz, "Qoe and power efficiency tradeoff for fog computing networks with fog node cooperation," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, pp. 1–9.

[5] X. Duan, F. Xu, and Y. Sun, "Research on offloading strategy in edge computing of internet of things," in *2020 International Conference on Computer Network, Electronic and Automation (ICCNEA)*, 2020, pp. 206–210.

[6] J. Xu, L. Chen, and P. Zhou, "Joint service caching and task offloading for mobile edge computing in dense networks," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 207–215.

[7] Y. Qi, L. Tian, Y. Zhou, and J. Yuan, "Mobile edge computing-assisted admission control in vehicular networks: The convergence of communication and computation," *IEEE Vehicular Technology Magazine*, vol. 14, no. 1, pp. 37–44, 2019.

[8] X. Gao, X. Huang, S. Bian, Z. Shao, and Y. Yang, "Pora: Predictive offloading and resource allocation in dynamic fog computing systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 72–87, 2020.

[9] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, ser. Synthesis Lectures on Communication Networks. Morgan & Claypool Publishers, 2010. [Online]. Available: http://dx.doi.org/10.2200/S00271ED1V01Y201006CNT007

[10] H. Gupta, A. Vahid Dastjerdi, S. K. Ghosh, and R. Buyya, "ifogsim: A toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments," *Software: Practice and Experience*, vol. 47, no. 9, pp. 1275–1296, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2509

[11] A. Kertész, T. Pflanzner, and T. Gyimothy, "A mobile iot device simulator for iot-fog-cloud systems," *Journal of Grid Computing*, vol. 17, 09 2019.

[12] I. Lera, C. Guerrero, and C. Juiz, "Yafs: A simulator for iot scenarios in fog computing," *IEEE Access*, vol. 7, pp. 91 745–91 758, 2019.

[13] Z. Ren, W. Shi, J. Wan, F. Cao, and J. Lin, "Realistic and scalable benchmarking cloud file systems: Practices and lessons from alicloud," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 11, pp. 3272–3285, 2017.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

El problema de la escalabilidad en el posicionamiento con Wi-Fi RTT

Nestor Gonzalez Diaz, Enrica Zola, Israel Martin-Escalona, Francisco Barcelo-Arroyo
Departamento de Ingeniería Telemática,
Universitat Politècnica de Catalunya (UPC)
C. Jordi Girona 1-3, 08034 Barcelona.
(nestor.gonzalez.diaz, enrica.zola, israel.martin, francisco.barcelo)@upc.edu

Resumen

Este artículo explora la localización en interiores mediante tecnologías Wi-Fi, con un enfoque en el estándar IEEE 802.11mc y las técnicas de *fingerprinting*. A pesar de la probada precisión en las estimaciones de Round Trip Time (RTT), la necesidad de inyectar tramas FTM para su obtención limita notablemente la escalabilidad del sistema, abriendo múltiples líneas en las que investigar. Una de ellas propone la hibridación con señales pasiva (cuya obtención no requiera de tráfico de localización). El mayor exponente en este sentido es el nivel de señal recibido (RSSI), presente en todos los dispositivos Wi-Fi. De esta forma, se pretende generar un sistema *fingerprinting* heterogéneo RTT/RSS que permita graduar el tráfico de localización a la vez que se maximiza la precisión. Los resultados preliminares demuestran un mayor impacto en la precisión al añadir muestras RTT frente a RSSI, si bien esta última fuente de información se revela como un complemento destacable en situaciones de oscuridad parcial en términos de RTT.

***Palabras Clave*—Localización en interiores, fingerprint, RTT, RSSI, machine learning, KNN, escalabilidad.**

I. INTRODUCCIÓN

En los últimos años, los sistemas de posicionamiento en interiores (IPS) han despertado el interés de la comunidad científica debido a, por un lado, la creciente demanda de localización precisa en entornos interiores, y por otro lado, la incapacidad del sistema de posicionamiento global (GPS) para cubrir esos espacios. La importancia de investigar en ese ámbito yace en que las personas pasan la mayor parte de su tiempo en interiores; por lo tanto, mejorar la precisión en IPS puede resultar en tiempos de respuesta reducidos en caso de emergencia, mayor seguridad y un uso más eficiente de los recursos. Sin embargo, lograr un posicionamiento preciso y de bajo coste en estos entornos es un reto complejo y abierto.

Hoy en día, existen diversas tecnologías empleadas en el posicionamiento en interiores, como Assisted-GPS, Radio-frequency identification (RFID), Ultra-Wide Band (UWB), Bluetooth y Wireless Fidelity (Wi-Fi) [1]. Entre estas tecnologías, el Wi-Fi destaca por disponer de una amplia cobertura y penetración (p.ej., la mayoría de los dispositivos móviles soportan esta tecnología). Históricamente, los sistemas de localización se han decantado por el uso del nivel de señal Wi-Fi (RSSI) por ser ésta una información presente en la mayoría de dispositivos [2].

En 2016, el IEEE presentó el estándar 802.11mc [3], que introduce la capacidad de realizar medidas temporales precisas (Fine Time Measurement (FTM)), y por ende de obtener estimaciones de distancias a partir de medidas de ida y vuelta de una señal (Round-Trip Time (RTT)) entre dos dispositivos Wi-Fi (p.ej., entre un usuario y el Punto de Acceso (AP)). Sin embargo, la estimación del RTT acarrea varios errores debido a múltiples factores: ancho de banda limitado, propagación multicamino, *drift* y el *jitter* del oscilador, etc [2]. El estándar permite promediar múltiples muestras de RTT para mitigar estos errores, aunque esto conlleva un aumento en el tiempo de respuesta y el tráfico de localización en la red y por consiguiente, limita tanto la cantidad de dispositivos que se pueden localizar concurrentemente (p.ej. 34 si la frecuencia de localización es de 1 Hz [4]), como el ancho de banda disponible para los servicios de datos regulares. Además, la mayoría de los trabajos en la literatura proponen usar la multilateración [5] como método para hallar la posición a partir de las estimaciones de Wi-Fi RTT; de esa forma, el usuario necesita empezar un proceso de medidas RTT con al menos tres APs para hallar una posición en 2D. Todo eso conlleva un evidente problema de escalabilidad intrínseco de la tecnología Wi-Fi RTT.

Otro método ampliamente estudiado en la literatura es el Fingerprinting (FP) RSSI, el cual se basa en construir una base de datos a partir de las medidas de la señal Wi-Fi (RSSI) [6] obtenidas en puntos concretos dentro de un mapa, para posteriormente cotejar las muestras obtenidas con las previamente almacenadas en la base de datos. Aunque la técnica FP ofrece un posicionamiento

potencialmente preciso, el RSSI es un dato muy volátil, lo que limita la capacidad de los algoritmos de Machine Learning (ML) a la hora de proporcionar una correcta estimación. Recientemente se ha demostrado que es posible emplear medidas RTT en sistemas FP, consiguiendo alcanzar mejores resultados comparado con FP RSSI, especialmente en escenarios poco densos [7].

El objetivo de este trabajo es el de profundizar en el uso del FP RTT e investigar diferentes soluciones para mejorar la escalabilidad del sistema. En la sección II se presentan las líneas de investigación que consideramos más relevantes y que se deberían abordar en un futuro para abordar el problema de la escalabilidad en los sistemas de posicionamiento Wi-Fi que emplean IEEE 802.11mc. En la sección III se abordan los resultados preliminares obtenidos tras aplicar una de las posibles soluciones a este problema. Finalmente, la sección IV presenta las principales líneas de investigación que se seguirá en un futuro próximo a tenor de los resultados alcanzados.

II. LÍNEAS DE INVESTIGACIÓN

Tal como se ha comentado en la introducción, Wi-Fi RTT presenta un problema intrínseco de escalabilidad puesto que cualquier usuario que quiera localizarse necesita forzar un intercambio de tramas de localización (i.e. FTM) con los puntos de acceso Wi-Fi cercanos. Cuantos más usuarios necesiten localizarse, mayor es el número de tramas FTM que deberán enviarse, en detrimento del tráfico asociado a los servicios de datos para los que la red de comunicaciones Wi-Fi fue concebida. Reducir el número de RTTs necesarios para alcanzar una estimación de la posición suficientemente precisa es crucial para poder aprovechar de forma eficiente los recursos disponibles. Este trabajo pretende, en primer lugar, estudiar las soluciones propuestas hasta ahora en la literatura a tal efecto y evaluar la viabilidad de las mismas para, posteriormente, proponer nuevas estrategias con las que favorecer la escalabilidad de los sistemas de posicionamiento FP Wi-Fi.

Recientes trabajos han abordado la fusión de métricas Wi-Fi RTT con otras de naturaleza pasiva, como pueden ser RSSI, channel state information (CSI), información del campo magnético, información del entorno a través de imágenes, sensores inerciales, etc. Por ejemplo En [8] se propone un sistema de localización FP híbrido que fusiona RSSI y RTT, demostrando mejorar los resultados en comparación con los modelos que utilizan solo RTT o solo RSSI. Otros estudios como el presentado en [9] muestran que la fusión de las predicciones obtenidas utilizando diferentes técnicas de localización (p.ej., multilateración y FP a partir de medidas de RSSI), es capaz de mejorar la precisión de las estimaciones.

Por otra parte, en [10] se proponen la fusión de Wi-Fi RTT con medidas RSSI y de los sensores inerciales. Sin embargo, el uso de sensores inerciales conlleva un aumento en la complejidad de los algoritmos de estimación que debe tenerse en presente y que hacen languidecer este tipo de soluciones frente a otras más simples.

Siguiendo la línea propuesta en [8], se plantea el uso de un sistema Wi-Fi FP que conjugue medidas RTT y RSSI. Pese a la alta disponibilidad de las medidas RSSI, su alta volatilidad suscita la duda de en qué proporción deberían fusionarse estas medidas con medidas RTT que han demostrado ser mucho más estables [7], para proporcionar un adecuado compromiso entre escalabilidad y precisión. Se plantea analizar el impacto de la cantidad de medidas RTT y RSSI en modelos de ML destinados al posicionamiento, estudiando su rendimiento en función del número de estimaciones RTT y RSSI empleados en el cálculo de la estimación de la posición. Los resultados preliminares de este estudio se presentan en la sección III, para lo que ha desarrollado una aplicación que integra un amplio número de tecnologías y permite la obtención de forma conjunta tanto de RTT como RSSI [11].

Además de esta línea de investigación, se prevé explorar el desarrollo de filtros que ayuden a reducir el ruido de las muestras de RTT empleadas en el posicionamiento [12][13]. De igual forma, el uso de sistemas de posicionamiento jerárquico se considera otra vía a la hora de acometer una fusión de métricas [14].

La técnica de FP requiere construir una base de datos previa al despliegue, lo que supone un hándicap destacable. Además, los cambios en el entorno de posicionamiento conllevan necesariamente una actualización de dicha base de datos. El coste temporal y de recursos que esto supone limita en gran medida despliegues a gran escala de esta técnica. Por ello, los autores en [15] proponen un modelo de predicción Gaussiano con el que estimar la base de datos FP. Un sistema Wi-Fi RTT (puro o híbrido) podría adoptar también esta aproximación, más aún teniendo en cuenta la estabilidad del RTT, para así reducir en gran medida el coste de la construcción de la base de datos FP.

III. RESULTADOS PRELIMINARES

En esta sección se muestran los primeros resultados obtenidos al aplicar una solución FP que fusiona observables RTT y RSSI. Para ello se propone el uso de un algoritmo supervisado de ML para clasificar las medidas sobre el conjunto de puntos de referencia de que consta la base de datos FP. El objetivo es entender qué mejoras puede aportar el uso de más características (i.e. medidas RSSI o RTT), especialmente en términos de error de posicionamiento.

Para ello, se utilizan los datos de la campaña de medidas descrita en [7], realizada en el auditorio de la Escuela de Ingeniería de Telecomunicación y Aeroespacial (EETAC), un espacio rectangular (19.2 x 8 m^2), con sillas, mesas y marcos metálicos, aunque se garantiza visión directa entre los distintos Puntos de referencia (RPs) y APs. Sobre él se ha establecido una cuadrícula de 5x5 RPs (i.e. 25 en total) a distancia de 4.8 m en el lado largo, y 2 m en el corto, en los que se toman las medidas. Para ello, se desplegaron 7 APs compatibles con Wi-Fi RTT: 4 en las esquinas, 2 en la mitad de las paredes largas, y 1 en la mitad de una de las paredes cortas. En cada uno de los 25 RPs, se tomaron 100 muestras con cada AP disponible. Se observaron un total de 216 direcciones MAC, de las

cuales 16 corresponden a los siete APs desplegados[1]. Se han obtenido así 2500 muestras por característica, es decir, para cada una de las 216 direcciones MAC de diferentes APs de las que obtenemos un RSSI y/o un un RTT.

Teniendo en cuenta la cantidad de características y para evitar el sobreajuste (*overfitting*), como primer paso, se han seleccionado aquellas características consideradas más relevantes, empleando para ello el algoritmo *Tree-based* [16]. Con ello se obtienen dos listas, con las 9 características más relevantes en términos de RTT y RSSI, respectivamente. De esta manera, se reduce el conjunto de datos inicial, a un máximo de 18 características.

Como algoritmo de predicción, se ha escogido el clasificador K-Nearest Neighbors (KNN) por ser uno de los mejores exponentes en cuanto a sencillez y rendimiento en el ámbito de la localización [17]. Para obtener un mejor rendimiento del KNN, se deja al algoritmo optimizar los hiperparámetros con los que realizar posteriormente las predicciones. El análisis del comportamiento de estos hiperparámetros y su influencia en el rendimiento, así como su generalización, son temas abiertos pospuestos a futuras investigaciones.

La Fig. 1 muestra la exactitud alcanzada por el algoritmo KNN al usar un determinado número de características RTT (con colores diferentes), en función del número de características RSSI. En dicha figura, la exactitud representa la proporción de predicciones correctas (i.e. RPs acertados) entre el número total de predicciones; si el algoritmo es capaz de acertar todas las predicciones, la exactitud del modelo es 1.

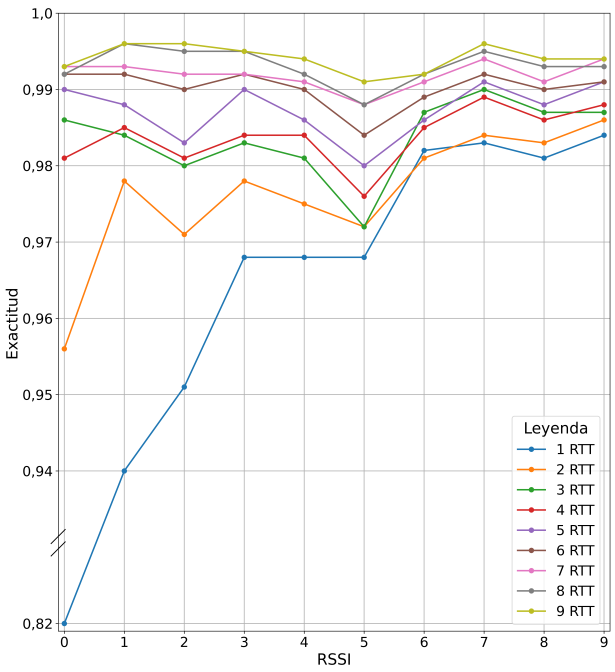


Figura 1. Exactitud del algoritmo KNN en función del número de características utilizadas en el modelo de clasificación.

Los resultados del estudio muestran que, en el modelo que emplea únicamente RTTs (i.e. RSSI igual a 0), 2 o más RTTs son suficientes para obtener buenos resultados, con una exactitud mayor a 0,956. A partir de 3 RTTs el beneficio obtenido al aumentar en el número de RTTs es muy paulatino y puede no llegar a justificar el incremento del *overhead* de localización (i.e. tráfico de localización en la red).

En cuanto a los modelos combinados con RSSI, el mayor incremento en exactitud se observa al emplear un único RSSI, tal y como se refleja en la Fig. 1. Debe notarse que, para mejorar la visualización de los resultados en dicha figura, no se ha usado el mismo escalado a lo largo de todo el eje y. Exceptuando el caso de un RTT (en azul), el añadir más de un RSSI no comporta una mejora significativa de la predicción y, además, tiende a generar fluctuaciones en la exactitud, lo que sugiere la presencia de un sobreajuste en los modelos empleados. Este sobreajuste se comienza a observar, de forma general, en modelos que emplean más de 4 características. Por ejemplo, 2 RTTs (en naranja) y 2 RSSIs, o 3 RTTs (en verde) y 1 RSSI, o 4 RTTs (en rojo) y 2 RSSIs, etc. A pesar de que la exactitud se mantiene por encima de 0,97, añadir sistemáticamente un mayor número de RSSIs no parece traducirse siempre una mejora significativa de la exactitud. Por ejemplo, para el caso de 1 RTT (en azul) se observa que la exactitud aumenta a medida que aumenta el número de RSSI considerados pero, entre 3 y 5 RSSIs, la exactitud es constante, para luego aumentar por encima del 0,98 y empezar a oscilar (sobreajuste).

Para entender mejor el rendimiento del algoritmo de localización propuesto, es importante analizar el error de posicionamiento cometido en las estimaciones. Para ello se ha calculado la raíz del error cuadrático medio (Root Mean Squared Error (RMSE)) de las posiciones estimadas, para así poder reflejar el comportamiento medio del modelo y resaltar aquellos en los que se han producido errores de mayor magnitud. La Fig. 2 muestra el comportamiento del RMSE en función del número de características RSSI (eje horizontal) y RTT (representado con colores diferentes). Nótese que también en este caso, en el eje vertical se ha cambiado el escalado para mostrar el valor más alto observado (2,48 m), sin perjudicar la visualización del resto de curvas. Se puede observar que, con solo añadir un RSSI, el RMSE baja en todos los modelos RTT, exceptuando el caso con 3 RTTs (en verde) donde añadir RSSIs no parece aportar un beneficio apreciable en términos de error de posicionamiento . Además, mientras mayor sea el número de RTTs empleado, menor el aporte por parte de los RSSIs. Exceptuando el caso de un RTT (en azul), si se añade un **único** RSSI a las características RTT el RMSE es siempre inferior a los 55 cm, lo cual es un resultado muy satisfactorio para la gran mayoría de las aplicaciones de localización.

De forma análoga a la exactitud, los valores de RMSE para cada curva empiezan a oscilar a medida que se añaden más características RSSI. Concretamente, se observa que a partir de 6 RTTs (en marrón), añadir un RSSI permitiría reducir la RMSE más que añadiendo un RTT adicional. En

[1]Los AP desplegados transmiten en múltiples bandas frecuenciales, empleando en cada una de ellas una dirección MAC diferente.

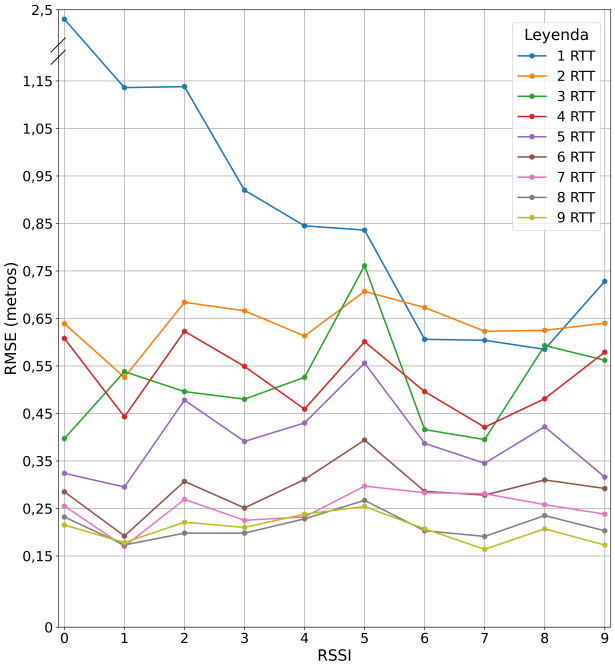


Figura 2. Evolución del RMSE en las predicciones obtenidas con KNN y en función del número de características.

el caso de 1 RTT (en azul), pese a que exactitud en el caso de añadir 2 RSSIs se ve incrementada si se compara con el uso de un único RSSI, los valores RMSE permanecen invariables. Se requiere elevar el número de RSSIs hasta 6 para converger a valores mínimos (alrededor de 55cm). Sin embargo, en escenarios poco densos, ese valor se puede alcanzar con 2 RTTs y 1 RSSI, a costa eso sí de un mayor overhead.

IV. CONCLUSIÓN Y TRABAJO FUTURO

Los hallazgos actuales sugieren que la incorporación de métricas RSSI a los modelos basados en RTT permiten mejorar su rendimiento y escalabilidad, lo que abre nuevas vías de exploración en la correlación de las métricas de RTT y RSSI. En los resultados presentados, se ha mostrado que un modelo que integra un RTT y un RSSI (no necesariamente tomados a partir del mismo AP) obtiene resultados prometedores, alcanzando una exactitud de 0,94 y una RMSE ligeramente por encima de un metro. Otra línea de investigación igualmente interesante es la explorar la viabilidad de utilizar las dos métricas, RTT y RSSI, tomadas del mismo AP. Con ello se podría aprovechar el hecho de que habitualmente la estimación del RTT viene acompañado por una medida de RSSI, simplificando el modelo y optimizando los recursos de red.

Otra línea de investigación futura consiste en analizar otros algoritmos de ML para comparar el beneficio que podría suponer emplear algoritmos quizás más complejos y cuantificar como esa mayor complejidad podría impactar en la capacidad del sistema para escalar.

Este estudio abre además la posibilidad de comparar qué combinaciones de métricas con RTT ofrecen un mejor rendimiento y eficiencia (Bluetooth, la red celular, mediciones de campo magnético, etc), y ver si estos valores son trasladables a otras tecnologías. Así mismo, también se plantea estudiar el impacto de medidas RSSI de esas otras tecnologías sobre modelos basados en RTT Wi-Fi.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Generalitat de Catalunya con la subvención 2021-SGR-00594.

REFERENCIAS

[1] H. Obeidat, W. Shuaieb, O. Obeidat, and R. Abd-Alhameed, "A review of indoor localization techniques and wireless technologies," Wireless Personal Communications, vol. 119, pp. 289–327, 2021.

[2] B. K. Horn, "Doubling the Accuracy of Indoor Positioning: Frequency Diversity," Sensors, vol. 20, no. 5, 2020.

[3] "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 802.11-2016 (Revision of IEEE Std 802.11-2012), pp. 1–3534, 2016.

[4] I. Martin-Escalona and E. Zola, "Passive Round-Trip-Time Positioning in Dense IEEE 802.11 Networks," Electronics, vol. 9, no. 8, 2020.

[5] C. Ma, B. Wu, S. Poslad, and D. R. Selviah, "Wi-Fi RTT ranging performance characterization and positioning system design," IEEE Transactions on Mobile Computing, 2020.

[6] X. Zheng, R. Cheng, and Y. Wang, "RSSI-KNN: A RSSI Indoor Localization Approach with KNN," in 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), 2023, pp. 600–604.

[7] I. Martin-Escalona and E. Zola, "Improving Fingerprint-Based Positioning by Using IEEE 802.11 mc FTM/RTT Observables," Sensors, vol. 23, no. 1, p. 267, 2023.

[8] H. Rizk, A. Elmogy, and H. Yamaguchi, "A Robust and Accurate Indoor Localization Using Learning-Based Fusion of Wi-Fi RTT and RSSI," Sensors, vol. 22, no. 7, p. 2700, 2022.

[9] G. Guo, R. Chen, F. Ye, X. Peng, Z. Liu, and Y. Pan, "Indoor Smartphone Localization: A Hybrid WiFi RTT-RSS Ranging Approach," IEEE Access, vol. 7, pp. 176 767–176 781, 2019.

[10] G. Guo, R. Chen, X. Niu, K. Yan, S. Xu, and L. Chen, "Factor Graph Framework for Smartphone Indoor Localization: Integrating Data-Driven PDR and Wi-Fi RTT/RSS Ranging," IEEE Sensors Journal, vol. 23, no. 11, pp. 12 346–12 354, 2023.

[11] I. Martin-Escalona, "Fingerprinting Map Builder," https://play.google.com/store/apps/details?id=edu.upc.grxca.wifimapbuilder&hl=en_ZA&gl=US, accessed: 2023-06-06.

[12] J. Fu, Y. Fu, and D. Xu, "Application of an Adaptive UKF in UWB Indoor Positioning," in 2019 Chinese Automation Congress (CAC), 2019, pp. 544–549.

[13] X. Liu, B. Zhou, P. Huang, W. Xue, Q. Li, J. Zhu, and L. Qiu, "Kalman Filter-Based Data Fusion of Wi-Fi RTT and PDR for Indoor Localization," IEEE Sensors Journal, vol. 21, no. 6, pp. 8479–8490, 2021.

[14] A. Alitaleshi, H. Jazayeriy, and J. Kazemitabar, "Affinity propagation clustering-aided two-label hierarchical extreme learning machine for wi-fi fingerprinting-based indoor positioning," Journal of Ambient Intelligence and Humanized Computing, vol. 13, no. 6, pp. 3303–3317, Jun 2022. [Online]. Available: https://doi.org/10.1007/s12652-022-03777-1

[15] H. Zou, M. Jin, H. Jiang, L. Xie, and C. J. Spanos, "WinIPS: WiFi-Based Non-Intrusive Indoor Positioning System With Online Radio Map Construction and Adaptation," IEEE Transactions on Wireless Communications, vol. 16, no. 12, pp. 8118–8130, 2017.

[16] A. K. Panja, S. F. Karim, S. Neogy, and C. Chowdhury, "A novel feature based ensemble learning model for indoor localization of smartphone users," Engineering Applications of Artificial Intelligence, vol. 107, p. 104538, 2022.

[17] A. Nessa, B. Adhikari, F. Hussain, and X. N. Fernando, "A Survey of Machine Learning for Indoor Positioning," IEEE Access, vol. 8, pp. 214 945–214 965, 2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Contención del movimiento lateral por análisis epidemiológico de grafos de Directorio Activo

David Herranz Oliveros, Iván Marsá Maestre, Enrique de la Hoz y Marino Tejedor Romero
Universidad de Alcalá, Dpto. de Automática
28805 Alcalá de Henares, España,
{david.herranz,ivan.marsa,enrique.delahoz,marino.tejedor}@uah.es

José Manuel Giménez Guzmán
Dpto. de Comunicaciones
Universitat Politècnica de València
46022 Valencia,España,
jmgimenez@upv.es

El ciberataque típico en una red corporativa, conocido como cadena de ataque o 'kill-chain', incluye varios pasos de movimiento lateral mediante los cuales los atacantes se desplazan desde su punto de entrada hacia un activo de alto valor en la red (generalmente, privilegios de administración de dominio). La capacidad de una red para ralentizar el movimiento lateral de los atacantes es crucial para su resiliencia, ya que proporciona tiempo para implementar contramedidas reactivas o rastrear el origen del ataque. En este trabajo, utilizamos un modelo de contagio epidemiológico para demostrar cómo el modelado de grafos y el análisis de las relaciones presentes en una red de Directorio Activo nos permite identificar pequeños subconjuntos de nodos en la red, cuyo fortalecimiento puede aumentar significativamente el tiempo que un atacante tarda en comprometer una fracción determinada de la misma.

***Palabras Clave*—movimiento lateral, grafos, Directorio Activo, resiliencia, epidemiología para ciberseguridad**

I. INTRODUCCIÓN

El movimiento lateral es el proceso y las técnicas utilizadas por un ciberatacante para avanzar a través de una red en busca de objetivos de alto valor. Es fundamental ralentizar este proceso para garantizar la resiliencia de la red, ya que brinda la oportunidad de detectar el ataque, mitigar su impacto y trazar su origen [1].

En este estudio, nos centramos en aprovechar la estructura de grafo que existe en una red corporativa para identificar puntos estratégicos donde implementar contramedidas preventivas. Nuestra hipótesis es que al actuar sobre un pequeño subconjunto de nodos o aristas de la red, lograremos mejorar significativamente su resiliencia.

Para ello, utilizamos métricas de epidemiología aplicadas a un modelo de grafo que representa una infraestructura de Directorio Activo (AD). Con este enfoque, identificamos nodos candidatos para recibir protección específica y evaluamos el impacto de proteger estos nodos basándonos en el tiempo que un atacante tarda en comprometer una fracción determinada de la red. Además, exploramos diferentes perfiles de capacidades del atacante.

Las contribuciones de este artículo son las siguientes:

- Modelamos el movimiento lateral utilizando un enfoque epidemiológico, donde los objetos y las relaciones de AD son como nodos y aristas en el modelo.
- Combinamos técnicas de epidemiología como el agrupamiento basado en densidad y algoritmos de estratificación del grafo para identificar nodos potenciales que deben ser protegidos.
- Modelamos las capacidades del atacante como subconjuntos de las aristas del grafo de AD que el atacante puede explotar, y comparamos el desempeño de nuestras técnicas propuestas con varios perfiles de capacidades del atacante utilizados como referencia.

El resto del artículo se estructura de la siguiente manera: en la sección II se revisa la literatura relevante relacionada con este tema. En la sección III, describimos nuestro modelo y enfoque propuesto, el cual evaluamos en la sección IV. Finalmente, en la sección V concluimos resumiendo nuestras contribuciones.

II. TRABAJO RELACIONADO

El artículo seminal que inspira nuestro trabajo sobre el estudio del movimiento lateral en redes modeladas como grafos es [2]. En este artículo, se propone modelar la red como un grafo dirigido y analizar cómo los

Este trabajo ha sido parcialmente financiado por el proyecto PID2019-104855RB-I00/AEI/10.13039/501100011033 del Ministerio de Ciencia e Innovación de España, por el proyecto SBPLY/19/180501/000171, de la Junta de Comunidades de Castilla-La Mancha (España), y por los proyectos UCeNet (CM/ JIN/2019-031) y WiDAI (CM/JIN/2021-004) de la Comunidad de Madrid y la Universidad de Alcalá. David Herranz también está financiado por una beca FPU de la Universidad de Alcalá. La publicación también forma parte del proyecto TED2021-131387B-I00 financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea "NextGenerationEU"/PRTR y del proyecto PID2021-123168NB-I00 financiado por MCIN/AEI/10.13039/501100011033/FEDER, UE.

atacantes se desplazan lateralmente a gran escala en una red que ha sido comprometida, considerando el papel de cada nodo en la conectividad del grafo. Además de este artículo, los modelos de grafo se han utilizado en ciberseguridad en otros trabajos. Por ejemplo, en [3], los autores utilizan esta metodología para construir un modelo de amenazas que, basado en grafos multicapa, permite reducir el riesgo al que están expuestos los activos de una red. Asimismo, en [4], se propone un análisis conductual basado en grafos del tráfico web que tiene lugar entre las entidades de una red para detectar anomalías que impliquen un comportamiento malicioso. Asimismo, existen trabajos recientes que utilizan grafos para detectar ataques de movimiento lateral [5], [6], [7]. En [1], se propone una técnica basada en comportamiento y no supervisada para detectar ataques de movimiento lateral mediante la identificación de conexiones anómalas entre sistemas. De manera similar, [8] propone una metodología para detectar amenazas de movimiento lateral en redes informáticas empresariales mediante el aprendizaje de grafos no supervisado y utilizando información de prácticas de registro o *logging* estándar de la industria.

Finalmente, el uso de modelos de contagio epidemiológico en ciberseguridad tiene cierta historia [9]. Seguimos un enfoque similar a [10], modelando el ataque como un proceso de infección probabilístico. Un trabajo interesante que utiliza la métrica de K-shell en epidemiología es [11], donde los autores señalan que, en el caso de un proceso de infección o transmisión en cualquier red, la propagación tiende a ser similar, siempre que comience o tenga lugar a través de nodos pertenecientes al mismo K-shell.

III. EPIDEMIOLOGÍA Y MITIGACIÓN DEL MOVIMIENTO LATERAL

La hipótesis inicial de este trabajo es que el enfoque de [2], basado en la identificación de vértices con centralidad atípica, puede ser ampliado aprovechando otras técnicas utilizadas en epidemiología, como K-shell. Posteriormente, analizamos los beneficios de aplicar salvaguardas de seguridad sobre los activos de red representados por los nodos identificados mediante ambas técnicas, utilizando modelos de contagio epidemiológico [2]. Además, tenemos la intención de ampliar este enfoque diferenciando el impacto de estas medidas sobre el movimiento lateral según las capacidades de explotación del adversario.

A. Identificación de nodos relevantes

Identificar nodos con conectividades atípicas en es muy valioso para determinar las áreas de la red en las que enfocarse para una mitigación eficiente del riesgo. Por lo tanto, basándonos en [2], proponemos identificar aquellos nodos con valores de centralidad atípicos. De esta manera, se puede lograr una reducción significativa del área de la red que debemos proteger.

Obtenemos los grafos para nuestros modelos epidemiológicos de la salida de la herramienta Bloodhound [12], que extrae la información de los objetos de AD (por ejemplo, equipos, usuarios, dominios...) y las relaciones (por ejemplo, membresía, privilegios de administración...) de una red dada. Dado que los grafos que hemos utilizado representan redes reales de AD, no los incluimos aquí.

La primera técnica que utilizamos para la identificación de nodos es el algoritmo DBSCAN [13] para identificar los vértices con valores de centralidad atípicos. Para cada vértice del grafo de autenticación, se calculan las siguientes métricas: número de descendientes, excentricidad, centralidad de intermediación, centralidad de vector propio, centralidad de grado y cercanía [14]. Luego aplicamos DBSCAN sobre estas métricas para encontrar los nodos que proporcionan una mayor conectividad a la red.

En paralelo al uso de DBSCAN, descomponemos la red utilizando K-shell [11]. Definimos una K-shell como el subgrafo más grande que puede conformarse a partir del grafo original, con vértices de al menos grado K. Esto nos permite estratificar el grafo según lo nuclear que es cada vértice con respecto al conjunto, o en otras palabras, lo lejos que está de la periferia [15]. De esta manera, los nodos con un valor más alto de K-shell serán aquellos ubicados en las capas más centrales y, por lo tanto, tendrán un mayor potencial como grandes propagadores durante un proceso de infección o movimiento lateral. Teniendo esto en cuenta, utilizamos K-shell para identificar los nodos que pertenecen a la capa más profunda de la red, observando los nodos que conforman el subgrafo de mayor grado que se puede conformar a partir del grafo original de red.

B. Grafos de autenticación

Como se menciona en la sección II, obtenemos información del entorno de AD de las infraestructuras TI y lo modelamos construyendo un grafo dirigido. Para esta investigación, tenemos la información completa de la red de una infraestructura real y anónima. Esta información se recopiló utilizando la herramienta BloodHound. El grafo original generado por BloodHound se compacta para facilitar el análisis. Utilizamos dos mecanismos de compactación diferentes. En la compactación basada en equipos (CBC), los nodos del grafo resultante representan equipos, y las aristas unen aquellos que tienen un camino directo en el grafo de autenticación original sin ningún equipo intermedio. En la compactación basada en usuarios (UBC), realizamos un proceso análogo, pero obtenemos un grafo donde los nodos son usuarios del sistema. El grafo UBC es demasiado grande para que DBSCAN se aplique con éxito, pero se puede aplicar la descomposición por K-shell a ambos grafos. La tabla I muestra el recuento de vértices para el grafo original y los dos grafos compactados, junto con el número de nodos candidatos identificados por DBSCAN (cuando corresponda) y K-shell. Podemos observar que el conjunto de nodos etiquetados como candidatos para la inmunización es muy selectivo, representando menos del 0.1% del recuento de vértices original.

IV. EVALUACIÓN DE RENDIMIENTO

A. Configuraciones experimentales

Para la evaluación experimental, aplicamos un modelo *Susceptible-Infected* (SI) [16], donde las simulaciones

Tabla I
ESTADÍSTICAS SOBRE LA IDENTIFICACIÓN DE VÉRTICES

Fuente	Recuento de vértices	DBSCAN	K-SHELL	Ratio
Grafo CBC	3236	44	16	1.36%
Grafo UBC	102589	–	32	0.03%
Grafo original	107368	44	48	0.07%

comienzan con uno o más nodos inicialmente comprometidos y continúan hasta que la infección alcanza un objetivo predefinido. Sin pérdida de generalidad, establecemos el objetivo en el compromiso del 50% de la red, y definimos el *Median Time to Half Compromise* (MTTHC) como nuestra principal métrica de rendimiento.

Como se mencionó anteriormente, queremos validar el efecto de nuestra estrategia de identificación de nodos para frenar el movimiento lateral en redes AD. Para hacerlo, *inmunizamos* los vértices identificados, aplicando un factor de inmunización al peso de sus aristas circundantes en el modelo de contagio SI. Este factor de inmunización se establece en el 80% del peso original de esas aristas. Vale la pena señalar que este es un caso muy conservador, ya que la inmunización solo reduce en un 20% la probabilidad de éxito dado un intento de explotación. Esto representa la aplicación de contramedidas preventivas sobre los usuarios o equipos representados por estos nodos.

Ejecutamos el modelo SI sobre el grafo original y registramos el MTTHC para las siguientes estrategias de inmunización (dos *baselines* y nuestra propuesta):

- *Sin inmunización:* Ejecutamos el modelo SI sobre el grafo AD original, con los pesos originales.
- *Estrategia de inmunización aleatoria (RImS):* Inmunizamos un subconjunto aleatorio de nodos de igual cardinalidad que nuestro subconjunto seleccionado (0,07% del grafo original). En lugar de simplemente elegir nodos al azar, inmunizamos un vecino de cada nodo seleccionado al azar. Esta es una estrategia típica de inmunización que favorece a los nodos de alto grado [17].
- *Estrategia de inmunización selectiva (SImS):* Inmunizamos los nodos identificados por DBSCAN y K-shell como se describe en la sección III.A.

Finalmente, para analizar cómo nuestras estrategias pueden ser más o menos adecuadas según el tipo de atacante, hemos definido cuatro perfiles de adversario estandarizados. A cada uno de estos perfiles se le ha dado la capacidad de explotar un subconjunto diferente de las relaciones de AD, que corresponden a diferentes tipos de enlace en el grafo (p.ej. *AdminTo*, *HasSession*). Desde el punto de vista del modelo SI, una capacidad determinada implica que la infección puede progresar a través de los enlaces correspondientes en el grafo.

En la Tabla II se muestran los perfiles de adversario definidos en este trabajo, junto con las capacidades correspondientes asociadas y el ratio de enlaces que pueden explotar. Cabe destacar que ciertas capacidades son comunes a todos los perfiles, ya que están relacionadas con tipos de enlace que siempre son explotables por los adversarios.

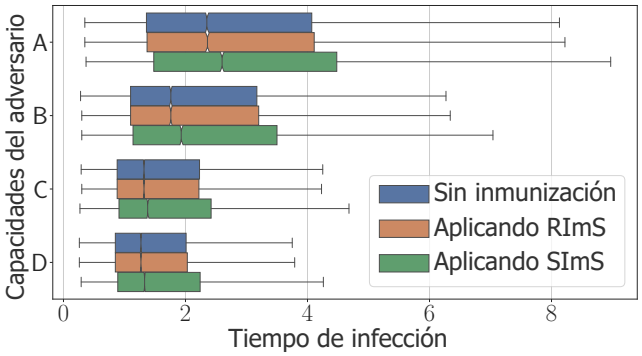


Fig. 1. Diagrama de cajas para el *Time to Half Compromise* según el perfil del adversario. Los valores medianos (MTTHC) y sus intervalos de confianza (IC-95%) están representados como muescas en el diagrama.

B. Resultados experimentales

Se ha generado un gran número de simulaciones del modelo SI con diversos parámetros. Se ejecutaron un total de 625 simulaciones en las que la infección inicial comenzó desde un conjunto de nodos seleccionados al azar. Entre estos casos, consideramos 25 tamaños diferentes para el conjunto de infección inicial ρ, seleccionados al azar y que varían entre el 5% y el 15% del total de vértices de la red. Sin pérdida de generalidad, fijamos la tasa de infección τ en el 50%. Cada simulación se repitió 50 veces. Esto resultó en un total de 31250 simulaciones del modelo SI, que se ejecutaron en paralelo para cada uno de los perfiles de adversario definidos anteriormente.

Como se discutió anteriormente, evaluamos el impacto de las contramedidas aplicadas mediante la generación de un indicador cuantitativo del impacto, y que definimos como el tiempo mediano para comprometer el 50% de la red (*Median Time to Half Compromise*, MTTHC). Los resultados de todas las simulaciones realizadas se muestran en la Fig. 1 y se resumen en la Tabla III.

El efecto de la inmunización selectiva en el movimiento lateral es claro, especialmente cuando las capacidades del oponente son menores, como se puede observar en la Tabla III. Hemos resumido este efecto mediante la relación entre

Tabla II
COBERTURA DE CAPACIDADES

Perfil de adversario	Capacidades	Ratio enlaces explotables
A	GpLink, Contains, MemberOf, GenericAl	55, 91%
B	GpLink, Contains, MemberOf, GenericAll,ForceChangePassword	77, 26%
C	GpLink, Contains, MemberOf, GenericAll, ForceChangePassword, WriteDacl, Owns	99, 61%
D	GpLink, Contains, MemberOf, GenericAll, ForceChangePassword, WriteDacl, Owns, GenericWrite, AddMembers, AllExtendedRights, HasSession, AdminTo, WriteOwner, CanRDP, DCSync, SQLAdmin	100%

Tabla III
RESULTADOS DEL MTTHC PARA LOS DISTINTOS PERFILES DE ADVERSARIO Y ESTRATEGIAS DE INMUNIZACIÓN

	Perfil de adversario							
	A		B		C		D	
	Mediana	IC	Mediana	IC	Mediana	IC	Mediana	IC
Sin inmunización	2.35	0.03	1.76	0.02	1.32	0.01	1.27	0.01
RImS	2.36	0.03	1.76	0.02	1.32	0.01	1.27	0.01
SImS	2.60	0.03	1.93	0.03	1.38	0.02	1.33	0.01
$\mathcal{G}_{\text{SImS}}$	1.1064	-	1.0966	-	1.0455	-	1.0472	-

el MTTHC para la inmunización selectiva y para cada *baseline*, mostrada como una ganancia $\mathcal{G}_{\text{SImS}}$ en la tabla. Al aplicar SImS a los enlaces que rodean solo el 0.07% de los vértices de la red, obtenemos un retraso significativo en el MTTHC. Este retraso varía desde el 4.6% en el peor caso hasta el 11% en el mejor caso. Si comparamos estos resultados con el retraso en el MTTHC obtenido al aplicar RImS [17] al mismo número de vértices, la diferencia es notable. En el caso de RImS, la inmunización no tiene ningún efecto significativo en el MTTHC. Esto demuestra que se puede lograr una reducción mucho mayor sobre el riesgo al que se enfrenta la red aplicando contramedidas según nuestra propuesta de identificación.

También hay que tener en cuenta que SImS es particularmente efectivo en casos en los que el perfil del adversario tiene menos capacidades. En estos casos, dificultarle al adversario moverse lateralmente a lo largo de las pocas rutas que puede explotar es clave para este rendimiento. Por el contrario, en casos en los que el oponente tiene mayores capacidades y, por lo tanto, alcanza el MTTHC en menos tiempo, el rendimiento se reduce porque las opciones del oponente son mayores a la hora de moverse por la red. Sin embargo, es cierto que somos capaces de identificar puntos altamente relevantes para la mayoría de estos casos, dado el retraso en el MTTHC, incluso en aquellos casos en los que la variedad de rutas de ataque que el adversario puede tomar es grande.

V. CONCLUSIONES

El movimiento lateral es una de las técnicas más extendidas en los ciberataques a redes empresariales. Por lo tanto, la capacidad de ralentizarlo, dando tiempo para activar otras contramedidas, es un factor clave para la resiliencia de la red. En este trabajo hemos explorado cómo el modelado y análisis basado en grafos de AD puede ayudar a identificar dónde implementar contramedidas preventivas para frenar el avance del movimiento lateral de un ciberataque. También hemos utilizado métricas epidemiológicas para identificar posibles ”super-propagadores” del movimiento lateral, marcando estos nodos como candidatos para recibir medidas preventivas (en nuestro caso, inmunización parcial). Hemos demostrado que mediante la inmunización selectiva de una fracción muy pequeña de la red (aproximadamente un 0.07% de los nodos), y con un impacto de inmunización muy conservador (reduciendo la probabilidad de infección solo en un 20%), podemos lograr un impacto significativo en el tiempo mediano hasta la mitad del compromiso (MTTHC) en un ataque, que oscila entre el 4.6% y el 11%, dependiendo de las capacidades del atacante.

REFERENCIAS

[1] B. A. Powell, “Role-based lateral movement detection with unsupervised learning,” *Intelligent Systems with Applications*, vol. 16, p. 200106, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667305322000448

[2] ——, “The epidemiology of lateral movement: exposures and countermeasures with network contagion models,” *Journal of Cyber Security Technology*, vol. 4, no. 2, pp. 67–105, 2020.

[3] I. Marsa-Maestre, J. M. Gimenez-Guzman, D. Orden, E. de la Hoz, and M. Klein, “React: Reactive resilience for critical infrastructures using graph-coloring techniques,” *Journal of Network and Computer Applications*, vol. 145, p. 102402, 2019.

[4] F. Zola, L. Segurola, J. L. Bruse, and M. G. Idoate, “Temporal graph-based approach for behavioural entity classification,” *Investigación en ciberseguridad: Actas de las VI Jornadas Nacionales (JNICLIVE)*, vol. 34, pp. 77–80, 2021.

[5] Y. Fang, C. Wang, Z. Fang, and C. Huang, “Lmtracker: Lateral movement path detection based on heterogeneous graph embedding,” *Neurocomputing*, vol. 474, pp. 37–47, 2022.

[6] I. J. King and H. H. Huang, “Euler: Detecting network lateral movement via scalable temporal link prediction,” in *Network and Distributed System Security Symposium*, 2022.

[7] L. Sadlek, P. Čeleda, and D. Tovarňák, “Identification of attack paths using kill chain and attack graphs,” in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*. IEEE, 2022, pp. 1–6.

[8] B. Bowman, C. Laprade, Y. Ji, and H. H. Huang, “Detecting lateral movement in enterprise computer networks with unsupervised graph ai.” in *23rd international symposium on research in attacks, intrusions and defenses (RAID 2020), USENIX Association*, 2020, pp. 257–268.

[9] J. O. Kephart and S. R. White, “Directed-graph epidemiological models of computer viruses,” *Computation: the micro and the macro view*, pp. 71–102, 1992.

[10] D. Acemoglu, A. Malekian, and A. Ozdaglar, “Network security and contagion,” *Journal of Economic Theory*, vol. 166, pp. 536–585, 2016.

[11] M. K. al., “Identification of influential spreaders in complex networks,” *Nature physics*, vol. 6, no. 11, pp. 888–893, 2010.

[12] A. Robbin, R. Vazarkar, and W. Schroeder, “Bloodhound: Six degrees of domain admin,” *BloodHound Dokumentation. URL: https://bloodhound. readthedocs. io (besucht 2020-04-23)*, 2020.

[13] M. Ester, H.-P. Kriegel, J. Sander, and X. X. al., “A density-based algorithm for discovering clusters in large spatial databases with noise,” *kdd*, vol. 96, no. 34, pp. 226–231, 1996.

[14] A. Hagberg, D. Chult, and P. Swart, “Exploring network structure, dynamics, and function using networkx,” *Proceedings of the th Python in Science Conference (SciPy2008)*, pp. 11–15, 2008.

[15] B. Pittel, J. Spencer, and N. Wormald, “Sudden emergence of a giant k-core in a random graph,” *Journal of Combinatorial Theory, Series B*, vol. 67, no. 1, pp. 111–151, 1996.

[16] I. Z. Kiss, J. C. Miller, P. L. Simon *et al.*, “Mathematics of epidemics on networks,” *Cham: Springer*, vol. 598, p. 31, 2017.

[17] R. Cohen, S. Havlin, and D. Ben-Avraham, “Efficient immunization strategies for computer networks and populations,” *Physical review letters*, vol. 91, no. 24, p. 247901, 2003.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Despliegue energéticamente eficiente de aplicaciones IoT en áreas rurales utilizando redes basadas en drones

Diego Ramos Ramos, Jaime Galán-Jiménez.
Departamento de Ingeniería de Sistemas Informáticos y Telemáticos,
Universidad de Extremadura
Av. de la Universidad, S/N, 10003, Cáceres, Extremadura, España
diegorr@unex.es, jaime@unex.es

Los índices de penetración de Internet aumentan anualmente, superando el 80% en países desarrollados. Sin embargo, cerca de dos mil millones de personas en zonas rurales y de bajos recursos carecen de conectividad, impidiendo el acceso a servicios esenciales como la atención médica remota o la educación a distancia. Para dar solución a este problema, en este trabajo se propone una arquitectura de red basada en vehículos aéreos no tripulados (*Unmanned Aerial Vehicles*, UAVs) y una solución heurística energéticamente eficiente para desplegar aplicaciones IoT de manera que la calidad de vida de la población que vive en zonas rurales se vea mejorada. Las aplicaciones IoT se descomponen en microservicios distribuidos a lo largo de la flota de UAVs para superar las limitaciones de batería y capacidad de cómputo. Las simulaciones realizadas muestran una eficacia cercana al 100% al servir las peticiones IoT de los usuarios en situaciones de baja y media de tráfico, así como una alta regresión en la carga de batería.

***Palabras Clave*—Eficiencia Energética, UAVs, IoT, microservicios, brecha digital**

I. INTRODUCCIÓN

El impacto de la tecnología en el mundo ha sido profundo, provocando cambios significativos en diversos aspectos, mejorando enormemente la eficiencia en los negocios, así como permitiendo una comunicación fluida entre empresas situadas en distintos lugares a través de videoconferencias. El rápido procesamiento de la información ha facilitado servicios en tiempo real en la sanidad electrónica a distancia, el entretenimiento y otros ámbitos. Sin embargo, a pesar de estos avances, sigue existiendo una importante brecha de conectividad entre las personas que residen en zonas urbanas y rurales [1]. Mientras que los habitantes de las ciudades tienen múltiples opciones para conectarse a Internet, como la tecnología 5G y la fibra óptica, los de las zonas rurales tienen opciones limitadas, a menudo dependiendo de la tecnología 3G o ADSL en el mejor de los casos. En consecuencia, la disponibilidad y la calidad de los sistemas de información se ven muy afectadas. Las zonas urbanas obtienen menores tiempos de respuesta de las aplicaciones y pueden desplegar aplicaciones de calidad de servicio (*Quality of Service*, QoS) en sectores como la sanidad y la industria, lo que contribuye a su desarrollo económico. Además, tecnologías avanzadas como la computación en el edge (*Fog-Edge Computing*) pueden implantarse en zonas urbanas, haciendo que aumente aún más la brecha digital.

En los últimos años, la comunidad investigadora ha realizado importantes esfuerzos para desarrollar arquitecturas tecnológicas destinadas a dar cobertura a las zonas rurales. Las redes comunitarias fueron una solución pionera en este campo y se han utilizado activamente en varios proyectos [2], [3]. Otro enfoque para ofrecer servicios digitales en las zonas rurales es la utilización de soluciones móviles oportunistas que aprovechan las técnicas de *store-carry-and-forward* [4].

Más recientemente, los esfuerzos de investigación se han centrado en la utilización de vehículos aéreos no tripulados (*Unmanned Aerial Vehicles*, UAVs) para prestar servicios a distintos niveles. El uso de vehículos aéreos no tripulados se ha extendido a numerosos ámbitos, gracias a sus características de fácil uso y costes razonables, mejorando la calidad de vida de las personas. Las aplicaciones civiles, incluida la videovigilancia a escala urbana [5], la entrega de pequeños paquetes [6] y la supervisión de la gestión de catástrofes [7], han adoptado los UAVs para aumentar los servicios terrestres tradicionales. Afortunadamente, el uso de UAVs ofrece una solución prometedora para proporcionar cobertura en zonas rurales. Cada UAV en este escenario está equipado con una pequeña estación base capaz de dar servicio a un área específica. El despliegue de un enjambre de UAV para cubrir el territorio

aporta varias ventajas, como la mejora de la cobertura, la capacidad, la fiabilidad y la eficiencia energética [8], [9].

En este trabajo, nos basamos en una arquitectura de red basada en UAV para proporcionar cobertura y servicios a zonas rurales. Debido a las capacidades limitadas de los UAV en términos de computación y batería, las aplicaciones IoT (*Internet of Things*) se descomponen en microservicios y cada UAV es capaz de desplegar y ejecutar un subconjunto específico de ellos. De este modo, se proporciona una solución heurística para gestionar las peticiones IoT de los usuarios con la mínima energía requerida del conjunto de UAVs que están prestando servicio en la zona. Los resultados de la simulación sobre un escenario realista muestran que la solución propuesta es capaz de reducir la energía requerida de la red basada en UAVs al tiempo que maximiza el número de peticiones IoT realizadas por los usuarios.

El resto del artículo se organiza de la siguiente manera: el estudio de trabajos relacionados, descrito en la Sección II. El modelo de sistema con la red basada en UAVs, el modelo de aplicación IoT basado en microservicios y el modelo de consumo de energía que se consideran se describen en la Sección III. La descripción del algoritmo propuesto para atender las peticiones de aplicaciones IoT minimizando la energía requerida se proporciona en la Sección IV. En la Sección V se presentan y discuten los resultados experimentales. Finalmente, en la Sección VI se revisan las conclusiones obtenidas.

II. TRABAJOS RELACIONADOS

El campo de la eficiencia energética es un ámbito emergente en el área de las redes basadas en UAVs. Por ejemplo, en [10], los autores proponen una estrategia de migración de tareas basada en una red Q profunda federada (*Deep Q-Network*, DQN). Esta estrategia tiene en cuenta la desviación de la carga y la desviación de la energía entre los MECS UAV. Utilizan DQN para crear un modelo local de optimización de migración para cada MECS UAV, y el aprendizaje federado genera un modelo global más eficiente al considerar características espaciales comunes entre regiones adyacentes. El rendimiento de esta estrategia se evalúa en términos de satisfacción de la restricción de retardo, desviación de carga y desviación de energía.

En otro de los estudios, [11], los autores abordan el desafío de superar las limitaciones de recursos computacionales y energéticos en los UAV utilizando *Mobile Edge Computing* (MEC). Proponen desplegar microservicios basados en contenedores en MEC para mejorar la calidad de servicio al procesar tareas descargadas de los UAV. El problema de programación de tareas se plantea considerando si se debe desplegar un nuevo servicio o utilizar un servicio existente para equilibrar la sobrecarga de transmisión de datos y el despliegue de microservicios. El objetivo es minimizar el tiempo total de finalización de la tarea. El problema se formula como una Programación Lineal Entera (*Integer Linear Programming*, ILP) y se demuestra que es NP-hard. Además, se propone un algoritmo de programación de peticiones basado en incentivos.

En el estudio descrito en [12], los investigadores proponen el uso de UAV como una red inalámbrica reorganizable energéticamente eficiente para brindar servicios a redes virtuales (*Virtual Networks*, VN). Sugieren ajustar de forma adaptativa la asignación de recursos de almacenamiento en caché, computación y comunicación, así como la topología de la red inalámbrica basada en UAV, en función de los requisitos de servicio de las VN utilizando algoritmos de inteligencia artificial. También se mencionan dos esquemas: redes de detección *UAV-IoT peer-to-peer* y redes de detección *UAV-IoT clustering*.

En [13], los autores presentan un esquema de detección inalámbrica inteligente asistida por UAV para recopilar datos de dispositivos IoT. Proponen enfoques de planificación de rutas óptimas para el despliegue de UAV con el objetivo de minimizar el tiempo de ejecución y el consumo de energía. Se consideran dos esquemas: redes de detección *UAV-IoT peer-to-peer* y redes de detección *UAV-IoT clustering*.

Para concluir, en [14], los autores proponen una arquitectura de sistema que involucra dispositivos móviles inteligentes (*System Mobile Devices*, SMDs), UAV y una estación base (*Base Station*, BS). Introducen el proceso de decisión de Markov con recompensas desconocidas (*Markov Decision Process with Unknown Rewards*, MDPUR) para considerar la distancia de migración, la energía residual de los UAV y el estado de la carga. Proponen un algoritmo de iteración de valores basado en ventajas (*Advantage-Based Value Iteration*, ABVI) para obtener una estrategia eficaz de migración de tareas, que busca equilibrar la carga y reducir el consumo total de energía del grupo de UAV, al tiempo que se mantiene la calidad del servicio al usuario.

En resumen, hasta el momento no se ha encontrado un trabajo que se centre en el despliegue proactivo de microservicios en una red basada en UAV para brindar servicios a la población rural que utilice aplicaciones IoT con el objetivo de minimizar el consumo de energía. La siguiente sección describe el modelo de sistema para el mapeo de aplicaciones IoT en la arquitectura de red basada en UAV.

III. MODELO DEL SISTEMA

En esta sección, se proporciona la descripción del modelo de sistema para la asignación de aplicaciones IoT en la arquitectura de red basada en UAV. En primer lugar, se introducen las consideraciones sobre la arquitectura de red. A continuación, se describe el modelo de aplicación IoT, que se basa en una arquitectura de microservicios. Por último, se explica el modelo de consumo de energía aplicado al enjambre de UAV.

A. Arquitectura de red basada en UAVs

Como se ha indicado en I, el escenario objetivo es una zona rural sin conectividad a internet. Por lo tanto, se considera una arquitectura de red compuesta por un enjambre de UAV capaces de comunicarse explotando la conectividad inalámbrica. Cada UAV lleva una estación base 5G encima, que es capaz de comunicarse con los

UAV vecinos dentro de un rango específico. De este modo, el área considerada queda cubierta por el enjambre de UAV y los usuarios pueden ejecutar aplicaciones IoT. Es obvio que un UAV es capaz de proporcionar cobertura a n usuarios que se encuentran dentro de los límites de un radio específico. Cada usuario puede solicitar la ejecución de una aplicación IoT específica al UAV al que está conectado, y este UAV debe devolver la información correspondiente solicitada por el usuario. Se asumen condiciones de *Line of Sight* (LoS) entre usuarios y UAVs, ya que el escenario considerado carece de grandes obstáculos (edificios, árboles, etc.) que puedan degradar la calidad de servicio (*Quality of Service*, QoS) obtenida. Las perturbaciones debidas a malas condiciones meteorológicas o a fallos de los UAVs, se dejan para futuros trabajos.

Tabla I
NOTACIÓN UTILIZADA

Símbolo	Descripción
$\mathcal{N}$	Conjunto de UAVs
$\mathcal{L}$	Conjunto de enlaces inalámbricos que conectan $\mathcal{N}$
$\mathcal{U}_n$	Usuarios conectados al UAV n
z_n	Ubicación (x_n, y_n) del UAV n
$\mathcal{M}_n$	Microservicios desplegados en el UAV n
f_n	Frecuencia de la CPU del UAV n
p_n	Coeficiente de potencia del UAV n
w_u	workflow de la aplicación IoT solicitada
m_i^u	Microservicio i-ésimo del workflow w_u
$\mathcal{E}_n^{\text{RES}}$	Batería restante del UAV n
$\mathcal{E}_{i,n}^{\text{PROC}}$	Consumo de energía para procesar el microservicio i
$\mathcal{E}_{i,n}^{\text{TRANS}}$	Consumo de energía para transferir el microservicio i
$\mathcal{E}_{w_u}$	Consumo de energía para ejecutar el workflow
L	Cantidad de datos del microservicio a transmitir
R	Velocidad de transmisión en el canal, WiFi
d	Distancia mínima entre UAV
s	Velocidad de la luz en el vacío
V	Tensión nominal de la batería [15]
t	Tiempo de transmisión
E_b	Consumo de batería del UAV

B. Cobertura de UAVs y modelo de pérdida de trayectoria

El diagrama de propagación considerado de la antena cónica colocada en la parte superior de los UAV tiene un ángulo θ. Se proyecta un área circular en el suelo (véase Fig. 1), cuyo radio, r, es proporcional a la altitud del UAV, h, según la Ecuación 1:

$$r = h \cdot \tan(\theta) \tag{1}$$

Donde θ es el ángulo de la antena cónica para el UAV y h es su altitud actual. Aunque suponemos una cobertura circular por medio del UAV, el escenario se divide en un conjunto de áreas aplicando el teselado de Voronoi como en [8], lo que puede dar lugar a un pequeño solapamiento en cuanto a la cobertura proporcionada por dos UAVs adyacentes. Si nos centramos ahora en el modelo de pérdida de trayectoria que se considera, éste se basa en el definido en [16], donde la pérdida de trayectoria media probabilística entre un UAV que se sitúa a una altitud h sobre una ubicación $z = (x, y)$ se define mediante la Ecuación 2:

$$\chi(h,k) = \frac{A}{1 + d \cdot \exp(-e \cdot (\frac{180}{\pi}) \tan^{-1}(\frac{h}{k}) - d)} + 10\log(h^2+k^2) + B \tag{2}$$

Donde $A = \eta_{\text{LoS}} - \eta_{\text{NLoS}}$, y $B = 20 \cdot \log(\frac{4\pi f_c}{c}) + \eta_{\text{NLoS}}$. Las pérdidas para las condiciones LoS y *No Line of Sight* (NLoS) están representadas por η_{LoS} y η_{NLoS}, respectivamente. Además, d y e son los parámetros de la curva S que dependen del entorno [16], f_c es la frecuencia portadora y k es la distancia entre el centro del UAV y el usuario al que se presta servicio, como en la Ecuación 3:

$$k = \sqrt{(x_k - x)^2 + (y_k - y)^2} \tag{3}$$

Destacamos que el modelo de pérdida de trayectoria definido anteriormente se considera para el tipo de escenario para el que está concebida la solución propuesta, es decir, zonas rurales abiertas en las que no existen interferencias debidas a la infraestructura existente y a los UAV vecinos.

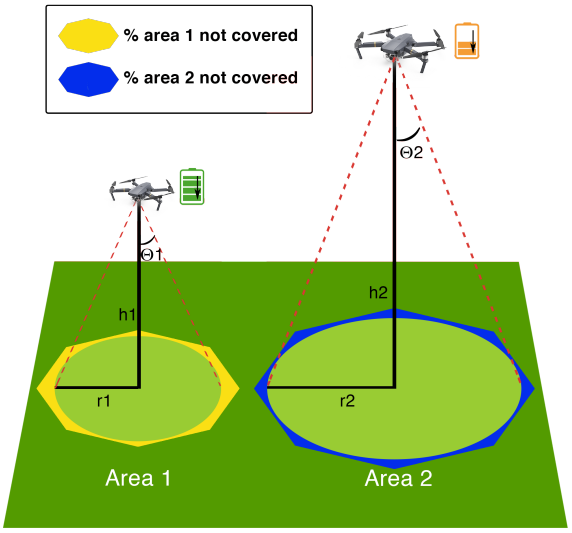


Fig. 1. Ejemplo de cobertura alcanzada por un UAV en función de la altitud.

C. Modelo de aplicaciones IoT basado en microservicios

En este trabajo, se considera que los usuarios solicitan aplicaciones IoT que siguen la Arquitectura de Microservicios (*Microservices Architecture*, MSA), es decir, las aplicaciones IoT se descomponen en un conjunto de microservicios que ejecutan una funcionalidad específica. Definimos como workflow como una instancia de una aplicación IoT solicitada por un usuario. Más formalmente, un workflow $w_u \in \mathcal{W}$ requerido por el usuario u es una secuencia ordenada de microservicios $w_u = \{m_1, m_2, ..., m_n\}, \forall m_i \in \mathcal{M}$, que se ejecutan uno tras otro para componer toda la funcionalidad de la aplicación IoT solicitada. Cada microservicio perteneciente al workflow puede desplegarse en diferentes vehículos aéreos no tripulados, de modo que el workflow puede implicar la

visita de varios vehículos aéreos no tripulados. De este modo, la colocación de los microservicios es esencial en los UAVs, para lograr una QoS aceptable (por ejemplo, el tiempo de respuesta máximo tolerable para la aplicación). Sin embargo, dado que el problema de asignación de servicios es NP-hard y computacionalmente intensivo, este trabajo en las características topológicas de la topología de la red para colocar el conjunto de microservicios en los UAVs. Cabe destacar que aspectos de QoS como el ancho de banda, latencia, confiabilidad y la priorización del tráfico, son tenidos en cuenta en trabajos actualmente en desarrollo.

Para ilustrar el comportamiento del modelo propuesto, la Fig. 2 muestra un caso de estudio que puede aplicarse a zonas rurales con carencia de conectividad. El escenario considerado (véase Fig. 2) puede dividirse en 9 áreas y cada una de ellas está cubierta por un UAV a una altitud de $h = 50$ m. La distancia mínima entre UAVs es de 900 m., que es una distancia lo suficientemente grande como para evitar interferencias y permitir un enlace directo LoS con UAVs vecinos [8]. En el caso de estudio se consideran dos usuarios: u_1 y u_2. El usuario u_1 está conectado a un UAV y requiere la ejecución de una aplicación IoT de electrocardiograma (ECG). Por otra parte, el usuario u_2 tiene un smartwatch y su objetivo es controlar su presión arterial. Por lo tanto, se solicita una aplicación IoT para el análisis de la presión arterial (PA). Aplicaciones similares para la detección de enfermedades cardiovasculares se utilizan comúnmente en el campo de la inteligencia en el edge [17].

Dado que ambas aplicaciones siguen el paradigma MSA, consideramos el uso de 4 microservicios diferentes y atómicos: m_1, un monitor de ECG; m_2, un microservicio de compresión de datos; m_3, un monitor de BP; y m_4, un microservicio de cifrado. Así, las dos aplicaciones IoT (ECG y BP) pueden definirse como dos flujos de trabajo: $w_1 = \{m_1, m_2, m_4\}$ y $w_2 = \{m_3, m_4\}$. La aplicación ECG está definida por w_1, donde el microservicio de monitorización ECG es ejecutado en primer lugar por el wearable del usuario. Dado que el tamaño del ECG obtenido es grande, se comprime antes de ser cifrado. La aplicación de PA, por su parte, sólo requiere la invocación del monitor de PA y del microservicio de cifrado, que combinados forman el workflow w_2. Estas dos aplicaciones de *Internet of Medical Things* (IoMT) se representan en la Fig. 2, y los microservicios asociados se colocan en UAVs específicos. Como se puede observar, varias réplicas del mismo microservicio pueden ser desplegadas en diferentes UAVs permitiendo a diferentes usuarios ejecutarlos desde diferentes partes del escenario.

Observando el ejemplo representado en la Fig. 2, el usuario u_1 solicita un análisis de ECG y la ruta para satisfacer w_1 está representada por flechas negras. Por otra parte, el camino seguido por el análisis de PA solicitado por el usuario u_2 se muestra mediante flechas amarillas. Obsérvese que el microservicio 4 está replicado en varios nodos y es solicitado por ambas aplicaciones.

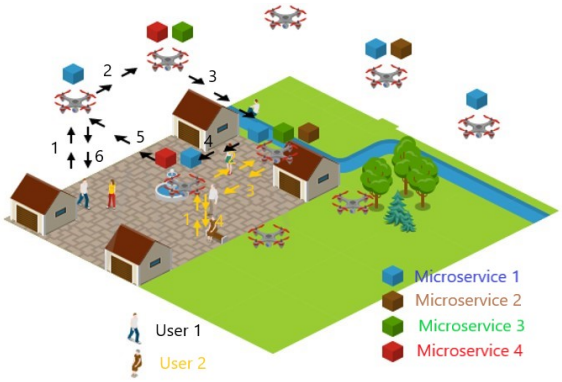


Fig. 2. Ejemplo de despliegue de aplicaciones IoMT.

D. Modelo de consumo de energía

En esta sección se describe el modelo de consumo de energía considerado. El consumo de energía para procesar un microservicio m_i^n por el UAV n considera el consumo de energía requerido por la ejecución del propio microservicio más el consumo de energía derivado de la transmisión de los datos emitidos al siguiente UAV del workflow. Así, si denotamos con $\mathcal{E}_{i,n}^{\text{PROC}} = \sum_{k=1}^{c_i} p_n \cdot f_n$ la energía requerida por el UAV n para procesar y ejecutar el microservicio i, y la energía requerida por el UAV n para enviar la información al siguiente UAV de la secuencia (si lo hay) o al usuario final si es el último del workflow como $\mathcal{E}_{i,n}^{\text{TRANS}}$, entonces el consumo de energía necesario para ejecutar un workflow w_u solicitado por el usuario u, el cual requiere la ejecución de K microservicios en un subconjunto de UAVs $\mathcal{N}' \subseteq \mathcal{N}$, puede evaluarse aplicando la Ecuación 4.

$$\mathcal{E}_{w_u} = \sum_{i=1}^{K} \mathcal{E}_{i,n}^{\text{PROC}} + \sum_{i=1}^{K} \mathcal{E}_{i,n}^{\text{TRANS}} \qquad \forall n \in \mathcal{N}' \tag{4}$$

La definición de la energía necesaria para la transmisión para un UAV se define según la Ecuación 5:

$$\mathcal{E}_i^{\text{TRANS}} = E_b * (\frac{1000}{V}) \tag{5}$$

Los componentes restantes de la ecuación se explican de la siguiente manera según las Ecuaciones 6 y 7:

$$t = \frac{L}{R} + \frac{d}{s} \tag{6}$$

$$E_b = p_n * t \tag{7}$$

donde E_b representa el consumo de batería del UAV, V representa el voltaje nominal de la batería, t representa el tiempo de transmisión, L representa el tamaño de entrada del microservicio a transmitir, R representa la velocidad de transmisión en el canal, d representa la distancia mínima entre UAVs y s representa la velocidad de la luz en el vacío. En Tabla I se muestra la notación utilizada en el modelo del sistema. De esta forma, se definirían cada una de las ecuaciones que intervienen en el coste de

transferencia de un microservicio concreto de un UAV a otro en la red.

Nótese que puede ocurrir que dos microservicios consecutivos en el workflow puedan ser desplegados en el mismo UAV. En este caso, $\mathcal{E}_{i,n}^{\text{TRANS}} = 0$, ya que no hay necesidad de enviar los datos hacia un UAV diferente.

E. Formulación del Problema

Una vez definidos los modelos considerados, en esta sección se describe la formulación del problema. Consideremos un grafo $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ compuesto por un conjunto de UAVs $n \in \mathcal{N}$ que están situados a una altitud h sobre una localización $z_n = (x_n, y_n)$. Cada UAV n está conectado a un subconjunto de vecinos $\mathcal{N}^n \subset \mathcal{N}$ con tecnología WiFi. Un conjunto de usuarios $\mathcal{U}_n$ está conectado a cada UAV $n \in \mathcal{N}$ y requiere ejecutar peticiones para diferentes aplicaciones IoT. Además, cada UAV tiene la capacidad de proporcionar recursos de computación, desplegando e invocando así un conjunto de microservicios $\mathcal{M}_n \subset \mathcal{M}$.

Cada microservicio $m_i \in \mathcal{M}$ se define como $m_i = (c_i, d_i)$, donde c_i representa los ciclos de CPU requeridos por el microservicio y d_i son los datos asociados al mismo. Así, un UAV n se define mediante una 5-tupla de tipo $n = \{z_n, \mathcal{M}_n, f_n, p_n, \mathcal{E}_n^{\text{RES}}\}$, donde z_n es la posición del UAV, M_n indica el conjunto de microservicios desplegados, f_n representa la frecuencia de la CPU, p_n es el coeficiente de potencia y $\mathcal{E}_n^{\text{RES}}$ se refiere a la batería restante del UAV. La función objetivo, que pretende minimizar el consumo de energía necesario para servir todas las peticiones IoT compuestas por flujos de trabajo, y requeridas por todos los usuarios del escenario, se define según la Ecuación 8:

$$min \sum \mathcal{E}_{w_u} \qquad \forall u \in \mathcal{U} \tag{8}$$

Tras la definición de los modelos considerados en la arquitectura y la descripción del problema, el algoritmo prupuesto en la siguiente sección, plantea una localización de microservicios en los UAVs con el objetivo de minimizar la energía consumida para la ejecución de aplicaciones IoT en escenarios rurales.

IV. ALGORITMO PROPUESTO

En esta sección, se aborda el desafío de la colocación eficiente de microservicios en UAVs. Para ello, se propone una solución innovadora denominada UAV-GreenMS, un algoritmo diseñado para maximizar la escalabilidad y optimizar el despliegue de microservicios en redes basadas en UAV. Esta solución tiene como objetivo satisfacer las peticiones de los usuarios, minimizando el consumo de energía de la red.

UAV-GreenMS se basa en la técnica del camino más corto para encontrar la mejor ubicación del siguiente microservicio en el workflow de la aplicación. Si hay varias opciones para seleccionar el siguiente microservicio, el algoritmo calculará el coste en dos situaciones posibles: i) el coste (en términos de consumo de energía) de desplegar el microservicio y sus dependencias (si las tiene) en el UAV; y ii) el coste de transmitir el microservicio solicitado a la red de UAV, es decir, a otro UAV donde el microservicio ya esté desplegado. Una vez calculados ambos costes, el algoritmo seleccionará la situación en la que el coste sea menor. Evidentemente, para esta decisión deben cumplirse varias restricciones: i) el UAV tiene el nivel de batería por encima de un umbral predefinido y batería suficiente para transmitir un microservicio tanto al usuario como dentro de la red, o desplegar el microservicio solicitado en sí mismo; y ii) el UAV tiene suficiente capacidad de procesamiento para ejecutar el microservicio.

En UAV-GreenMS, se establece un tiempo máximo de despliegue para cada microservicio ya desplegado en un UAV. Así, si un microservicio desplegado no ha recibido ninguna solicitud durante cierto tiempo, puede ser eliminado del conjunto de microservicios disponibles en el UAV: i) si el microservicio alcanza ese tiempo de espera, debe ser eliminado del conjunto de microservicios desplegados en el UAV; y ii) si el microservicio es solicitado en el UAV antes de que expire ese tiempo de espera, se restablece.

Una vez seleccionado el UAV disponible para transmitir o desplegar el microservicio solicitado, el algoritmo pasará al siguiente microservicio del workflow. Si una solicitud no puede satisfacerse debido a limitaciones de batería o computación, la solicitud será rechazada.

El Algoritmo 1 muestra el pseudocódigo de UAV-GreenMS. En primer lugar, se realiza una fase de inicialización. Después se comprueba si el microservicio no está desplegado en el nodo n actual (línea 4) o si ya está desplegado (línea 29).

En el primer caso, cuando el microservicio no está desplegado en el nodo actual, se calcula el coste de desplegar el microservicio en el nodo actual y el coste de transmitirlo desde otro UAV de la red que haya desplegado dicho microservicio. A continuación, se comprueba si el microservicio solicitado m está disponible en alguno de los nodos de la red, si está disponible (línea 7), se compara el coste de desplegarlo en el nodo actual o de transmitirlo desde la red (línea 9). En caso de que el microservicio no esté disponible, se comprueban las restricciones de batería y recursos necesarias para su despliegue (líneas 22 - 23). Si se cumplen las restricciones, el microservicio m se desplegará en el nodo n (línea 24); de lo contrario, se rechazará la solicitud de microservicio. Como se ha mencionado anteriormente, si el coste de desplegar el microservicio m en el nodo n es inferior al coste de transmitirlo desde la red y el nodo n cumple las restricciones de despliegue (línea 9), se desplegará el microservicio m (línea 10). Por otro lado, si el coste de transmitir el microservicio desde la red es inferior al coste de desplegarlo sobre el nodo n, o si el nodo n no cumple las restricciones de despliegue, se comprueba si dicho nodo puede satisfacer la petición solicitando este microservicio a la red (líneas 12 - 13).

Si se cumplen las restricciones, la transmisión se ejecutaría a lo largo de la ruta más corta (s_path) calculada desde el UAV en el que está desplegado el microservicio solicitado, hasta el UAV en el que se solicita dicho microservicio. Además, se actualizaría la batería de los UAV implicados en la transmisión, así como el tiempo

Algorithm 1 UAV-GreenMS pseudo-código.

Require: Topología de red $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, nodo fuente: s, el workflow asociado a la solicitud del usuario u: w_u

1: $n \leftarrow s$ ▷ UAV Actual
2: $s\_path \leftarrow \emptyset$
3: **for all** $m \in w$ **do** ▷ Gestionar microservicio
4: **if** $m \notin microservicios(n) \vee micorservicios(n) \in \emptyset$ **then**
5: *costeEnUAV* $\leftarrow$ calcularCosteEnUAV(n, m)
6: *costeEnRed* $\leftarrow$ calcularCosteEnRed$(\mathcal{G}, m, s_path)$
7: **if** *costeEnRed* $\neq 0$ **then**
8: *desplegar* $\leftarrow$ restriccionesDespliegue(n, m)
9: **if** *costeEnUAV* $\leq$ *costeEnRed* $\wedge$ *desplegar* **then**
10: desplegarMicroservicio(n, m)
11: **else**
12: *transmitir* $\leftarrow$ restrTransmision$(\mathcal{G}, m, s_path)$
13: **if** *transmitir* **then**
14: ejecutarTransmision$(\mathcal{G}, m, s_path)$
15: actBateriaUAVsEnPath(m, s_path)
16: actTiempoMicroservEnPath(m, s_path)
17: **else**
18: mostrarMensajeDenegacion() ▷ UAV no puede dar servicio
19: **end if**
20: **end if**
21: **else**
22: *desplegar* $\leftarrow$ restriccionesDespliegue(n, m)
23: **if** *desplegar* **then**
24: desplegarMicroservicio(n, m)
25: **else**
26: mostrarMensajeDenegacion()
27: **end if**
28: **end if**
29: **else**
30: *suficienteBateria* $\leftarrow$ restriccionesBateria(n, m)
31: **if** *suficienteBateria* **then**
32: actualizarBateriaUAV(n, m)
33: actTiempoDespliegueMicroservicios(n, m)
34: **else**
35: mostrarMensajeDenegacion()
36: **end if**
37: **end if**
38: actualizarTiempoMicroserviciosEnRed$(\mathcal{G}, n)$ ▷ Update microservices time deployed in the network
39: **end for**

de espera del microservicio solicitado m en cada uno de los UAV implicados en el s_path (líneas 14 - 16). Por el contrario, si el nodo n no cumple las restricciones para transmitir la solicitud, ésta se rechaza (línea 18).

En el segundo caso, cuando el microservicio solicitado m ya está desplegado en el nodo actual n (línea 29), se comprueba si el nodo actual tiene batería suficiente para seguir manteniendo dicho microservicio desplegado (línea 30). En caso afirmativo, se actualiza la batería del nodo y el tiempo de espera de los microservicios desplegados (líneas 32 - 33). Por otro lado, si el nodo no puede mantener el microservicio desplegado, la petición será rechazada. Al final de la ejecución, se actualizan los tiempos de espera de todos los microservicios desplegados en la red (línea 38).

V. RESULTADOS

En esta sección se realiza un análisis para evaluar el algoritmo UAV-GreenMS mediante simulaciones en un escenario rural realista. En primer lugar, se describe la configuración de la simulación, incluyendo el escenario considerado y la parametrización. A continuación, se evalúa el número de peticiones IoT satisfechas en función de la densidad de usuarios. Además, se lleva a cabo un análisis del rendimiento de la batería restante media por UAV, teniendo en cuenta la variación en la densidad de usuarios y el porcentaje de microservicios desplegados. Por último, también se evalúa el impacto de la distribución de a lo largo del terreno.

A. Configuración de la simulación

UAV-GreenMS ha sido evaluado sobre un escenario rural realista localizado en el Valle del Jerte, Cáceres (España), el cual está compuesto por 5 pueblos (Casas del Castañar, Cabrero, Barrado, Piornal y Valdastillas) conectados por caminos rurales como puede verse en la Fig. 3, a escala de 2 kilómetros.

Fig. 3. Despliegue de la red de UAVs en el escenario considerado.

Con el fin de proporcionar cobertura y servicios a la población de dichas localidades, se despliega sobre el escenario una arquitectura de red basada en UAVs, recogida en la descripción de la Sección III colocando un conjunto de UAVs a una altitud de 50 m. sobre el terreno, con una distancia máxima entre UAVs de 900 m. Esta configuración, permite minimizar las interferencias en la comunicación inter-UAV [8]. Además, se asumen condiciones LoS entre usuarios y UAVs, ya que el escenario considerado carece de grandes obstáculos (edificios, árboles, etc.) que puedan degradar la QoS obtenida. Las perturbaciones debidas a malas condiciones meteorológicas o a fallos de los UAVs se dejan para futuros trabajos. Así, la red considerada se compone de 57 UAVs que dan cobertura a las zonas residenciales y a las carreteras que las conectan. La Fig. 3 muestra el despliegue de UAVs sobre el escenario rural considerado. Cada UAV del escenario está equipado con una Raspberry Pi de 4 GB de capacidad RAM, una CPU de 1GHz y una batería de 1000 W/h. Para monitorizar los parámetros de salud de las personas mayores que viven en el escenario, se han considerado dos aplicaciones de

Internet of Medical Things (IoMT): i) una aplicación IoT de electrocardiograma (ECG) que requiere la ejecución de un workflow compuesto por tres microservicios, $w_1 = \{m_1, m_2, m_4\}$, y ii) una aplicación IoT para la monitorización de la presión arterial (PA), que solo requiere la ejecución de dos microservicios, $w_2 = \{m_3, m_4\}$. La Tabla II muestra los requerimientos de RAM y los ciclos de batería por microservicio, según los valores reportados en [18]–[20]. En cuanto al consumo de energía derivado del uso de microservicios, cada ciclo requiere una cantidad de $3.5 \cdot 10^{-9}$ W, mientras que se requieren $2 \cdot 10^{-4}$ W por cada segundo empleado en una petición [21].

Tabla II
ESPECIFICACIÓN DE MICROSERVICIOS

Microservicio	Descripción	RAM	Batería	Tamaño de Entrada
m_1	Monitor ECG	393 MB	24.4 GCycles	3.93 MB
m_2	Compresión	136 MB	9.9 GCycles	1.36 MB
m_3	Presión sanguínea	393 MB	24.4 GCycles	3.93 MB
m_4	Encriptado	79 MB	6.1 GCycles	0.79 MB

B. Evaluación del algoritmo

En esta sección, se realiza una evaluación del rendimiento del algoritmo UAV-GreenMS con el foco en cuatro métricas principales: i) el porcentaje de batería restante tras servir el conjunto de peticiones de los usuarios, ii) el número de veces que un UAV participa en la transmisión de un microservicio a/desde otro UAV, iii) el porcentaje de peticiones satisfechas y no satisfechas en cada escenario; y iv) el porcentaje de peticiones transmitidas y desplegadas en cada escenario. En este caso, el área objeto de evaluación es una submatriz de 3x3 que abarca el pueblo Casas del Castañar, es decir, los nodos 1 a 9 en la Fig. 3. El número de peticiones de los usuarios oscila entre 1.000 y 20.000, donde en cada simulación, las peticiones se distribuyen aleatoriamente entre el conjunto de UAV del escenario.

En primer lugar, la Fig. 4 informa del porcentaje mínimo, máximo y medio de batería restante en el conjunto de UAVs en función del número de peticiones en el escenario. Puede observarse que, a medida que se incrementa el número de peticiones, el promedio de batería restante en los UAVs comienza a disminuir, con una estabilización tras atender 5.000 peticiones.

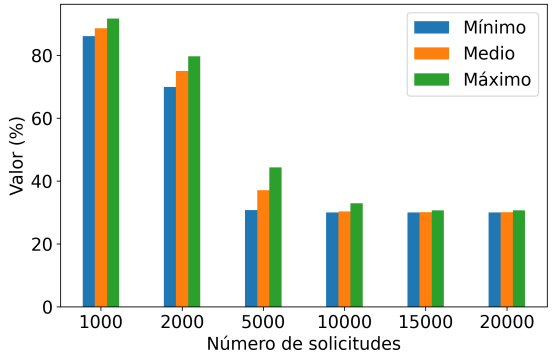


Fig. 4. Porcentaje del nivel de batería en los UAVs.

Si trasladamos ahora nuestra atención al porcentaje de peticiones satisfechas, la Fig. 5 muestra que a partir de 5.000 peticiones la red empieza a saturarse con un ligero porcentaje de peticiones rechazadas.

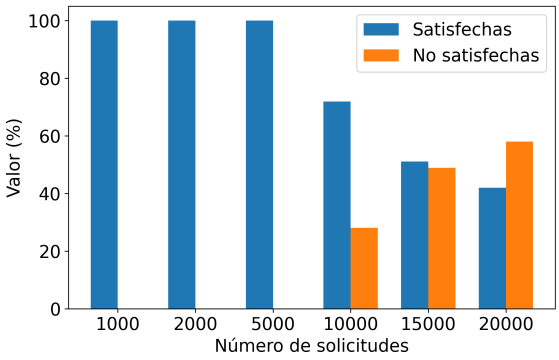


Fig. 5. Peticiones satisfechas y no satisfechas en cada uno de los escenarios.

A medida que aumenta el número de peticiones, la batería media del UAV se reduce (Fig. 4), satisfaciendo casi el 100% de las peticiones cuando hay 5.000 peticiones. También se puede observar que, para un mayor número de solicitudes, el enjambre de UAVs no es capaz de satisfacer todas estas. Dado que los escenarios se generan aleatoriamente, la distribución de microservicios puede afectar al consumo de batería, ya que influye en el tiempo en el que los microservicios se despliegan o no en un determinado UAV.

En la Fig. 6 se muestran los resultados extraídos tras evaluar el porcentaje de peticiones transmitidas y desplegadas en cada uno de los escenarios.

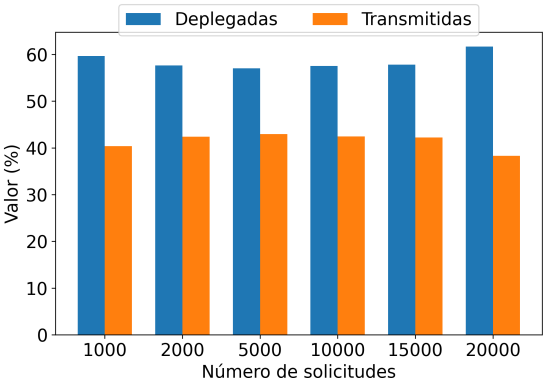


Fig. 6. Solicitudes transmitidas y desplegadas en cada uno de los escenarios.

Se puede observar que tanto el porcentaje de peticiones desplegadas directamente en los UAVs (barras azules) como las transmitidas desde los UAVs a otros UAVs (barras naranjas), se mantienen estables con una variación del 5%, a pesar del incremento en el número de peticiones. Un último punto a considerar es la evaluación del número de saltos en los que se han visto involucrados los UAVs de la red.

Por último, en la Fig. 7 se representa el número mínimo, máximo y medio de saltos en los que se han visto

involucrados los UAVs de la red en cada uno de los escenarios. Se puede observar que el número de saltos en los que se han visto involucrados los UAVs de la red comienza a aumentar gradualmente hasta estabilizarse en aquéllas pruebas en las que los usuarios solicitan más de 10.000 peticiones. Una vez alcanzado este valor el número de saltos se mantiene constante, debido a que aunque el número de solicitudes aumenta, los drones ya no cuentan con el nivel necesario de batería que les permita seguir transmitiendo.

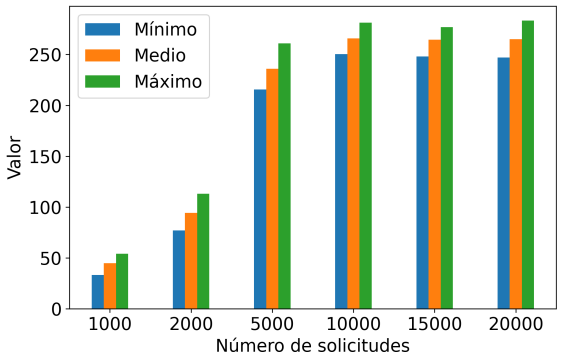


Fig. 7. Número de saltos para la transmisión en cada uno de los escenarios.

VI. CONCLUSIÓN

Las zonas rurales son menos atractivas para los operadores de red, ya que la baja densidad de población no justifica el despliegue de la infraestructura necesaria para proporcionar acceso a Internet de banda ancha. Con el fin de acercar los servicios a las personas que viven en zonas rurales, este artículo explota las capacidades de las redes basadas en UAV para proponer una solución energéticamente eficiente que permita desplegar aplicaciones IoT basadas en microservicios y así mejorar la calidad de vida de la población rural.

Los resultados de las simulaciones realizadas muestran la efectividad de la solución propuesta, evaluando el porcentaje de peticiones IoT que se sirven a los usuarios en un escenario realista y reduciendo el consumo de energía requerido por los UAVs al gestionar dichas peticiones.

Como futuros pasos de investigación, se planea evaluar la solución propuesta en un escenario rural real.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por MCIN/AEI/10. 13039/501100011033 y por la Unión Europea ”Next GenerationEU /PRTR”, por el Ministerio de Ciencia, Innovación y Universidades (proyectos TED2021-130913B-I00, PDC2022-133465-I00), por el proyecto PID2021-124054OB-C31 y la subvención CAS21/00057 (MCI/AEI/FEDER, UE), y por la Consejería de Economía, Ciencia y Agenda Digital de la Junta de Extremadura (GR21133).

REFERENCIAS

[1] A. Delaporte and K. Bahia, “The state of mobile internet connectivity 2021,” GSMA Connected Society, Tech. Rep., 2021.
[2] “Fon is the global WiFi network with millions of hotspots,” [Online]. Available: https://fon.com/. Accessed: 2023-05-31.
[3] D. Talbot, K. Hessekiel, and D. Kehl, “Community-owned fiber networks: Value leaders in america,” Berkman Klein Center for Internet & Society Research Publication, Tech. Rep., 2017.
[4] M. Jesus-Azabal, J. Berrocal, V. N. Soares, J. García-Alonso, and J. Galán-Jiménez, “A self-sustainable opportunistic solution for emergency detection in ageing people living in rural areas,” *Wireless Networks*, pp. 1–18, 2023.
[5] A. Trotta, F. D. Andreagiovanni, M. Di Felice, E. Natalizio, and K. R. Chowdhury, “When uavs ride a bus: Towards energy-efficient city-scale video surveillance,” in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 1043–1051.
[6] B. D. Song, K. Park, and J. Kim, “Persistent uav delivery logistics: Milp formulation and efficient heuristic,” *Computers & Industrial Engineering*, vol. 120, pp. 418–428, 2018.
[7] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, “Help from the sky: Leveraging uavs for disaster management,” *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 24–32, 2017.
[8] 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.
[9] L. Amorosi, L. Chiaraviglio, and J. Galán-Jiménez, “Optimal energy management of uav-based cellular networks powered by solar panels and batteries: Formulation and solutions,” *IEEE Access*, vol. 7, pp. 53 698–53 717, 2019.
[10] A. Shin and Y. Lim, “Federated-learning-based energy-efficient load balancing for uav-enabled mec system in vehicular networks,” *Energies*, vol. 16, no. 5, 2023.
[11] S. Huang, D. Zeng, and Z. Qu, “Toward performance efficient uav task scheduling in cloud native edge,” in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 4517–4522.
[12] S. Fu, L. Yin, C. Jiang, and A. Jamalipour, “An energy-efficient intelligent framework of uav-enhanced vehicular networks: Open problems and a case study,” *IEEE Vehicular Technology Magazine*, vol. 17, no. 2, pp. 94–102, 2022.
[13] D. Van Huynh, T. Do-Duy, L. D. Nguyen, M.-T. Le, N.-S. Vo, and T. Q. Duong, “Real-time optimised path planning and energy consumption for data collection in uav-aided intelligent wireless sensing,” *IEEE Transactions on Industrial Informatics*, 2021.
[14] W. Ouyang, Z. Chen, J. Wu, G. Yu, and H. Zhang, “Dynamic task migration combining energy efficiency and load balancing optimization in three-tier uav-enabled mobile edge computing system,” *Electronics*, vol. 10, no. 2, 2021.
[15] Y. Chu, C. Ho, Y. Lee, and B. Li, “Development of a solar-powered unmanned aerial vehicle for extended flight endurance,” *Drones*, vol. 5, no. 2, p. 44, 2021.
[16] A. Al-Hourani, S. Kandeepan, and S. Lardner, “Optimal lap altitude for maximum coverage,” *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.
[17] V. Hayyolalam, M. Aloqaily, O. Ozkasap, and M. Guizani, “Edge intelligence for empowering iot-based healthcare systems,” *arXiv preprint arXiv:2103.12144*, 2021.
[18] A. Limaye and T. Adegbija, “A workload characterization for the internet of medical things (iomt),” in *2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2017, pp. 302–307.
[19] “ARM Cortex-A53 MPCore Processor Technical Reference Manual r0p3,” Available: https://developer.arm.com/documentation/ddi0500/e/level-1-memory-system/about-the-l1-memory-system. Accessed: 2023-05-31.
[20] U. Jayasankar, V. Thirumal, and D. Ponnurangam, “A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications,” *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 2, pp. 119–140, 2021.
[21] M. Milosevic, A. Dzhagaryan, E. Jovanov, and A. Milenković, “An environment for automated power measurements on mobile computing platforms,” in *Proceedings of the 51st ACM Southeast Conference*. New York, NY, USA: Association for Computing Machinery, 2013.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Red de Sensores para el Control de la Salinización del Agua de Riego en Agricultura Vertical

Ali Ahmad, Francisco Javier Diaz, Sandra Viciano-Tudela, Sandra Sendra, Jaime Lloret
Instituto de investigación para la gestión integrada de zonas costeras, Escola Politècnica Superior de Gandia.
Universitat Politècnica de València.
Carrer del Paranimf, 1, 46730 Grao de Gandia, Valencia
aahmad1@upv.es, fjdiabla@doctor.upv.es, svictud@upv.es, sansenco@upv.es, jlloret@dcom.upv.es.

El uso de la agricultura vertical está aumentando su importancia en la sociedad actual debido a que permite reducir el espacio necesario para llevar a cabo un cultivo. Se sabe, además, que el sector primario requiere de la incorporación de tecnología para mejorar la sostenibilidad de la actividad. Por ello, en este artículo se presenta el diseño y despliegue de una red inalámbrica de sensores encargada de monitorizar la calidad del riego, en términos de salinización, para agricultura vertical. El artículo presenta 2 prototipos de sensores de salinidad basadas en una única bobina capaces de determinar los niveles de salinidad del líquido donde son introducidas. Ambos prototipos son probados y comparados con equipamiento profesional para finalmente someterlas a un proceso de validación. Por último, se analiza el tráfico de red que generaría una instalación de este tipo para diferentes números de nodos y los resultados arrojan que el ancho de banda consumido por un noto tipo ESP32 es extremadamente bajo obteniendo valores medios de ancho de banda de unos 700 Bps.

Palabras Clave- **Agricultura vertical, sensores, sostenibilidad, Internet de las cosas (IoT), redes inalámbricas de sensores.**

I. INTRODUCCIÓN

En el último siglo, se han desarrollado nuevas técnicas agrícolas para satisfacer la creciente demanda de alimentos y productos agrícolas [1]. Sin embargo, esta continua demanda, ejerce una elevada presión sobre los recursos naturales limitados de nuestro planeta. El desafío radica en encontrar formas de satisfacer las necesidades alimentarias, sin agotar los recursos del planeta [2]. A medida que aumenta la conciencia sobre los efectos negativos de la agricultura en el medio ambiente [3], es necesario desarrollar enfoques y técnicas que minimicen la huella ambiental de la agricultura. Estos enfoques deben ser capaces de cumplir con la demanda futura de alimentos de manera sostenible [4]. El desarrollo de nuevas tecnologías, como geoespaciales, "Internet of Things" (IoT), análisis "Big Data" e Inteligencia Artificial (IA), podría llegar a utilizarse para tomar decisiones de gestión informadas con el fin de aumentar la producción de cultivos [5] [6].

La Agricultura de Precisión (AP), consiste en una estrategia de gestión que utiliza un conjunto de técnicas avanzadas de información, comunicación y análisis de datos en el proceso de toma de decisiones. Esto permite optimizar y aumentar las aportaciones agrícolas y su producción, reduciendo así las pérdidas [7].

El uso de tecnologías de teledetección para la AP ha aumentado rápidamente en las últimas décadas. Los sistemas de teledetección, que utilizan tecnologías de información y comunicación, suelen generar un gran volumen de datos espectrales con una alta resolución, necesarios para su aplicación en la AP [8]. Algunas nuevas tecnologías de procesamiento de datos, como el análisis "Big Data", las IA y "Machine Learning" (ML), se utilizan para extraer información útil de ese gran volumen de datos [9]. Además, los sistemas de computación basados en la nube se utilizan para almacenar, procesar y distribuir o utilizar esos datos y aplicarlos a la AP [10]. Todas estas técnicas avanzadas de obtención y tratamiento de datos se ejecutan para facilitar el proceso de toma de decisiones durante la monitorización, gestión de riego, enfermedades y plagas, aplicación de nutrientes, y predicción del rendimiento de los campos de cultivo [5] [6].

El principal impulsor de la AP son las Redes de Sensores Inalámbricos (WSN). Estas se basan en una red de múltiples nodos inalámbricos conectados entre sí, que monitorizan parámetros físicos del entorno. Cada nodo consta de un transceptor de radio, un microcontrolador, sensores, antena y otros circuitos. Todos estos componentes permiten que se comunique con una pasarela o "gateway", que transmite la información recogida por los

sensores [11] que miden los parámetros físicos y envían la información recogida al controlador, que a su vez transmite esta información a la nube o a un dispositivo portátil.

El sector agrícola tiene requisitos específicos basados en las propiedades del suelo, naturaleza de los cultivos, clima, fertilizantes y agua. Los sensores controlan estos parámetros y gracias a los avances de las tecnologías WSN, el tamaño y coste de los sensores se ha reducido considerablemente, lo que hace factible su implantación en muchos sectores, como la agricultura [12]. Ya que el principal objetivo de la AP busca generar excedentes, optimizando el uso de recursos como agua, pesticidas y fertilizantes, el uso de sensores permite a los agricultores cuantificar los recursos necesarios para mantener el crecimiento de los cultivos.

La escasez de agua es un problema grave que afecta a nuestro planeta, disminuyendo sus niveles de agua superficial y subterránea (AS). Los recursos hídricos subterráneos están sometidos a una creciente presión, lo que acelera su agotamiento en términos de cantidad y calidad. Este impacto es especialmente preocupante en las zonas costeras, donde las condiciones socioeconómicas y ambientales agravan la situación [13]. En este tipo de áreas, la salinización de los recursos hídricos, tanto de forma natural como causada por actividades humanas, es un fenómeno extendido y alarmante [14]. Debido a la creciente demanda de agua dulce, para consumo humano y agricultura, ha sido necesario realizar estudios sobre la salinización de AS, ya que presenta una alta resistencia a la contaminación y por su gran capacidad de almacenamiento [15].

Las prácticas agrícolas pueden causar la destrucción de la vegetación natural, el deterioro del suelo y la contaminación de cuerpos de agua y acuíferos. El riego agrícola representa una amenaza potencial, especialmente en zonas con alta tasa de evapotranspiración y variabilidad de precipitaciones [16]. Este riego incrementa la cantidad de agua en el suelo, lo que aumenta la recarga de las AS [17], generando repercusiones negativas en su calidad. Asimismo, la extracción constante de AS para riego, ha llevado al agotamiento y deterioro de los acuíferos en todo el mundo [18]. Otro problema asociado, y que agrava la salinización de las aguas subterráneas, es el riego en períodos de sequía. Además, la conexión entre acuíferos puede provocar que un acuífero de buena calidad se vea afectado por el drenaje de agua de un acuífero salino, deteriorando así la calidad del AS [19].

El objetivo de este trabajo consiste en desarrollar una red de sensores que permita establecer unos valores de conductividad asociados a la cantidad de sal disuelta en agua de riego. Para ello, se estudiaron diferentes eluciones extraídas tras añadir a una porción de suelo una cantidad de agua a una concentración de sal determinada. La principal novedad de este estudio consiste en desarrollar sensores "low-cost" que permitan obtener esos valores y que realiza una serie de toma de decisiones para evaluar su respuesta.

El resto del estudio se estructura de la siguiente manera; en la Sección 2 se describen los trabajos relacionados. El sistema propuesto se describe en detalle en la Sección 3. A continuación, la Sección 4 detalla el "Test bench". Los resultados se discuten en la Sección 5. Por último, la Sección 6 consta de un resumen de las conclusiones y trabajos futuros.

II. Trabajos Relacionados

En esta sección detallamos los trabajos anteriores sobre AP e implementación de sensores con respecto a la regulación del riego. Un estudio reciente de Cobbenhagen et al. [20] destacó la necesidad de utilizar AP para aumentar la producción de los cultivos por área disponible, al tiempo que se reduce significativamente el uso de recursos como el agua, los pesticidas y los herbicidas para minimizar el impacto ambiental y aumentar la sostenibilidad de la cadena de producción de alimentos.

En otro estudio, Sanjeevi et al. [21] propusieron una arquitectura escalable de WSN utilizando IoT para monitoreo y control en agricultura y cultivo en áreas remotas. Los resultados experimentales demostraron que este enfoque supera a los sistemas convencionales basados en IoT, mejorando el rendimiento, la latencia, la relación señal-ruido, el error cuadrático medio y el área de cobertura en agricultura de precisión y cultivo. Del mismo modo, Khan et al. [22] implementaron un sistema para monitorear los parámetros de calidad del agua y el uso de plantas de tratamiento de efluentes utilizando WSN, dispositivos IoT y un módulo GSM para la transmisión de datos. El sistema propuesto recopiló datos de sensores de temperatura, turbidez y pH a través de un microcontrolador Arduino Uno R3, que luego se enviaron a un servidor en la nube basado en IoT. En caso de emergencias, los supervisores podían recibir alertas por SMS. Se calculó el índice de calidad del agua (WQI), mostrando un rendimiento excelente, y se comparó con el estado en la nube para garantizar un informe preciso de la calidad del agua.

Más recientemente, Siregar et al. [23] informaron sobre los tipos de tecnologías digitales utilizadas en la agricultura vertical de relaciones agrícolas innovadoras, el nivel de desarrollo y los modelos inteligentes adoptados, y los desafíos y barreras para aprovechar las oportunidades de implementación de sistemas de agricultura vertical inteligentes. Los autores identificaron el uso de IoT en la agricultura vertical, incluida la tecnología de sensores, los sistemas de monitorización y las aplicaciones de LED, con un enfoque en modelos de investigación de suelos (28%), hidroponía (18%) e iluminación y riego (22%). El estudio destaca la necesidad de explorar e implementar modelos y tecnologías inteligentes en la agricultura vertical para superar los desafíos relacionados con los costos y mejorar la sostenibilidad y la seguridad alimentaria.

Mientras que Singh y Sobti, [24] discutieron los desafíos de la urbanización y la necesidad de soluciones avanzadas en el desarrollo sostenible. Los autores destacaron la importancia de la agricultura urbana en las ciudades inteligentes y se resaltaron sus diversos elementos. Este estudio también se centró en la importancia de la utilización eficiente de los recursos

agrícolas, particularmente el agua de riego, y presentó un diseño para el riego de precisión para la monitorización del campo. El diseño ofreció una solución de largo alcance, en tiempo real y escalable para monitorear los requisitos de riego en función de las condiciones del suelo y el clima, abordando los desafíos relacionados con la cobertura del rango de nodos de sensor, la escalabilidad, la grabación de datos, la energía y el costo. También se discutió el uso de IoT para la monitorización del suelo y el clima, así como el desarrollo de una estación meteorológica de bajo costo con nodos de sensor de suelo de largo alcance.

Siddiqi y Al-Mulla, [25] investigaron las variaciones en el flujo de savia y el contenido volumétrico de agua en palmeras datileras bajo sistemas tradicionales de riego por inundación y modernos sistemas de riego por goteo. Informaron el uso de sensores basados en suelo y plantas para recopilar datos en tiempo real para el riego de precisión. Según los autores, los sensores recopilaron datos de temperatura, radiación solar neta, déficit de presión de vapor y velocidad del viento, y resultaron efectivos para determinar la eficiencia real del uso del agua.

Aunque se ha informado que las tecnologías novedosas como IoT, WSN, IA, etc. contribuyen notablemente a la agricultura de precisión, persiste la necesidad de desarrollar soluciones de bajo costo, rápidas, confiables y económicas. La complejidad y los desafíos de implementación de la tecnología desarrollada representan un gran desafío en el campo de la agricultura de precisión. Hasta donde tenemos conocimiento, solo se han informado algunos estudios para abordar los desafíos mencionados anteriormente. Por ejemplo, Sendra et al. [26] abordaron el problema de la salinización en fuentes de agua utilizadas para el riego en áreas mediterráneas. Los autores desarrollaron una red de sensores inalámbricos de bajo coste basada en LoRa que monitoreaba la concentración de salinidad en las fuentes de agua para la agricultura inteligente. Hallazgos similares fueron informados por Diaz et al. [27], donde informaron sobre un modelo de sensor eficaz y económico para la estimación de la concentración de fertilizantes en el agua de riego. Sin embargo, estos estudios no proporcionaron información detallada sobre la precisión y calibración del sensor o no proporcionaron información sobre la escalabilidad del sistema propuesto.

Cabe destacar que en un trabajo previo presentado en 2014 [28], se empezó a trabajar con la hipótesis de poder medir los niveles de salinidad del agua mediante bobinas. Sin embargo, en este nuevo trabajo, además de medir los niveles de concentración de sal, se pretende hacer una reutilización del agua usada con el objetivo de reducir el uso de recursos naturales.

En resumen, el estado del arte revela que aunque algunos estudios abordan desafíos relacionados con costos y sostenibilidad, no se enfocan específicamente en la aplicación del sistema propuesto para la monitorización del agua de riego y su reutilización en la agricultura vertical. Además, los sistemas desarrollados en algunos estudios previos se caracterizan por ser costosos y no se proporciona una adecuada información sobre su escalabilidad. Otro aspecto a destacar es que algunos estudios que se centran en la eficiencia real del uso del agua para el riego de precisión no son aplicables a los sistemas de agricultura vertical. En este sentido, nuestra propuesta se enfoca en el desarrollo y aplicación de sensores de bajo costo en la agricultura vertical y el riego sostenible, respondiendo a una necesidad crítica de la agricultura moderna. El potencial impacto de nuestra investigación en la conservación del agua y en prácticas agrícolas eficientes es significativo, y contribuye directamente a enfrentar los desafíos de la seguridad alimentaria a nivel mundial. Por lo tanto, el presente estudio se realizó sobre el desarrollo y la aplicación de sensores de bajo costo en el contexto de la agricultura vertical y el riego sostenible.

III. PROPUESTA

Esta sección presenta el sistema completo. Primeramente, se realiza una descripción general del sistema desarrollado. Así mismo se presentará el sensor diseñado para la medición de la salinidad del agua de riego. Por último, se presentará el algoritmo de control del sistema de riego.

A. Descripción general

Un sistema de agricultura vertical es una técnica innovadora que busca maximizar la producción de alimentos en espacios reducidos y de forma vertical, utilizando estructuras apiladas o en capas. En lugar de cultivar plantas en campos abiertos, se utilizan estructuras como torres, estantes o estanterías equipadas con sistemas de iluminación artificial, riego automatizado y control ambiental. En algunos casos, es posible reutilizar el agua de riego sobrante. No obstante, resulta vital determinar la idoneidad de dicha agua, especialmente en términos de cantidades de sales disueltas en el agua, ya que un riego continuo con agua no apta puede suponer la pérdida del cultivo.

En nuestra propuesta, se pretende instalar sensores de conductividad basados en una bobina en configuración LR en las distintas canaletas donde se crían los cultivos. Usando nodos tipo Arduino o similar con una interfaz inalámbrica, los datos serán enviados a un servidor local quien se encargará de decidir si el agua sobrante es apta o no para ser reutilizada en este cultivo o si, por el contrario, debería ser usada en otros cultivos como los tomates que se desarrollar con bastante éxito en aguas salmallas. Finalmente se enviarán los datos a un almacenamiento en la nube para futuros usos, análisis o aplicaciones. La Fig. 1 muestra la descripción general del sistema propuesto.

B. Sensor de salinidad desarrollado

Para la medida de la salinidad en el agua se desarrolla una bobina tipo solenoide de pequeño tamaño que se usa junto a una resistencia formando un circuito LR.

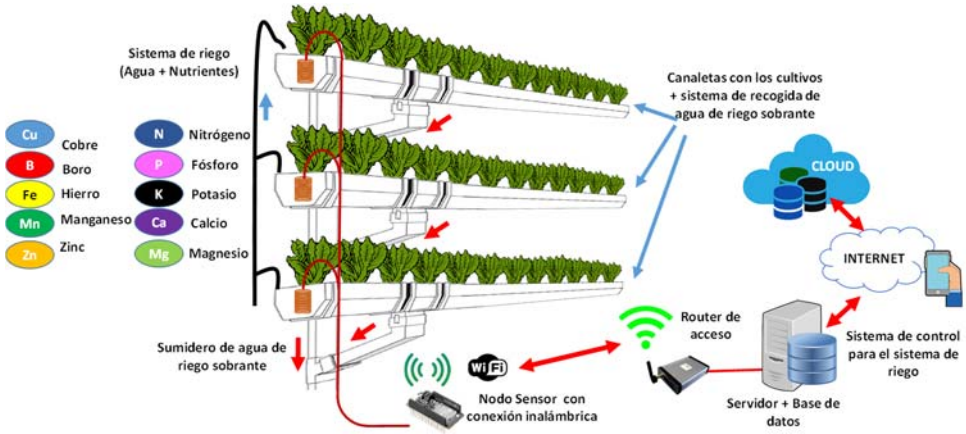


Fig. 1. Descripción general del sistema propuesto.

Tabla I
CARACTERÍSTICAS DE LAS BOBINAS USADAS

Modelo	Nº espiras	Diámetro de la bobina (cm)	Longitud de la bobina (cm)	Calibre del cable (mm)	Inductancia (µH)
1	75	1,4	1,7	0,2	52.10
2	36	1,4	1,8	0,6	9.15

Un circuito LR es un tipo de circuito eléctrico que combina una bobina o inductor (L) y una resistencia (R). Estos circuitos presentan una respuesta transitoria inductiva y una variación de la impedancia con la frecuencia de la señal de entrada. Este tipo de circuitos tienen varias aplicaciones en electrónica y comunicaciones, como filtros pasivos, circuitos de temporización y estabilización de corriente. También se utilizan en campos como la ingeniería de audio, sistemas de control y electrónica de potencia. La Tabla 1 muestra las características de las bobinas diseñadas mientas que la Fig. 2 muestra una imagen real de las bobinas. Estas bobinas son construidas con hilo de cobre esmaltado de diferentes calibres, como el usado en la fabricación de altavoces. Finalmente, los bobinados son pintados con laca transparente que únicamente les da robustez y resistencia. Como primer paso, antes de empezar a utilizar el prototipo desarrollado, se precisa definir su frecuencia máxima de resonancia. Para ello, el circuito LR es alimentado con una señal senoidal de 6 Vpp y diferentes frecuencias. Para el caso de la bobina del prototipo 1, se realiza un barrido en frecuencia desde 0,25 MHz hasta 4 MHz, mientras que para el prototipo 2, el barrido en frecuencia se realiza en el rango 3 MHz y 8 MHz. Para ambos casos, se recoge el valor de amplitud registrada en cada caso. La Fig. 3 muestra la frecuencia de funcionamiento para la bobina 1, mientras que la Fig. 4 muestra la frecuencia de funcionamiento para la bobina 2. Como podemos observar la mayor amplitud para la bobina 1, se registra para 2,5 MHz, mientras que la frecuencia de funcionamiento para la bobina 2 se sitúa en los 7MHz. Cabe destacar que se trata de valores bastante altos comparados con otros prototipos ya desarrollados en trabajos previos. Sin embargo, el pequeño tamaño de las mismas, hace que tengan estas frecuencias tan altas.

Para el uso del circuito LR como parte del sistema de senado se requiere el uso de una pequeña placa generadora de señal basado en el integrado AD9850 [29] capaz de generar señales de hasta 40 MHz.

Por otra parte, la tensión de salida del circuito, medida directamente en bornes de la resistencia nos dará un valor de tensión proporcional a la cantidad de sal contenida en el agua. Este valor es recogido con un nodo tipo Arduino Mega o similar como una señal analógica de la cual únicamente necesitamos registrar el valor de tensión máximo. Este valor será identificado con la canaleta o cultivo que monitoriza y almacenado en una base de datos local para posteriormente tomar las decisiones oportunas.

Finalmente, nos restará extraer el modelo de comportamiento de nuestras bobinas y estudiar su rango óptimo de trabajo, comparando los resultados medidos con los obtenidos con un conductímetro profesional. La Fig. 5 muestra cómo quedarían los sensores instalados dentro de las canaletas, así como la conexión del circuito con el nodo sensor.

C. Algoritmo de control para el riego sostenible

Una vez que se ha desarrollado el sensor, es esencial contar con un algoritmo que tenga la responsabilidad de supervisar tanto los niveles de salinidad del agua como las necesidades hídricas de las plantas.

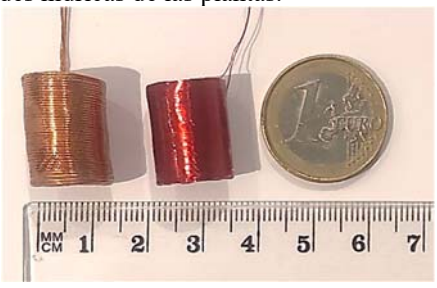


Fig. 2. Bobinas usadas.

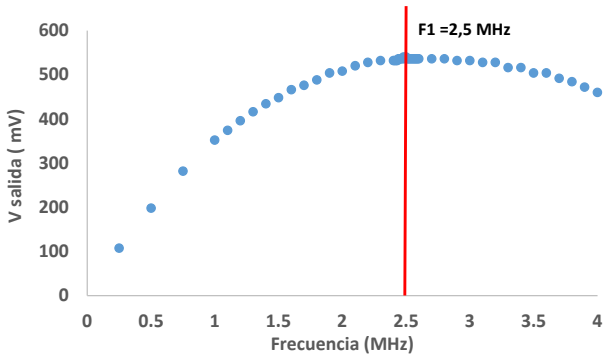


Fig. 3. Frecuencia de funcionamiento de la bobina 1.

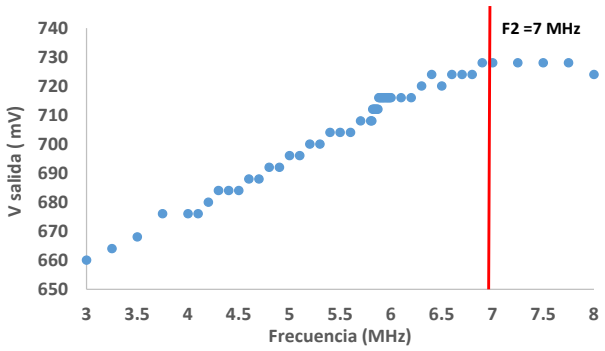


Fig. 4. Frecuencia de funcionamiento de la bobina 2.

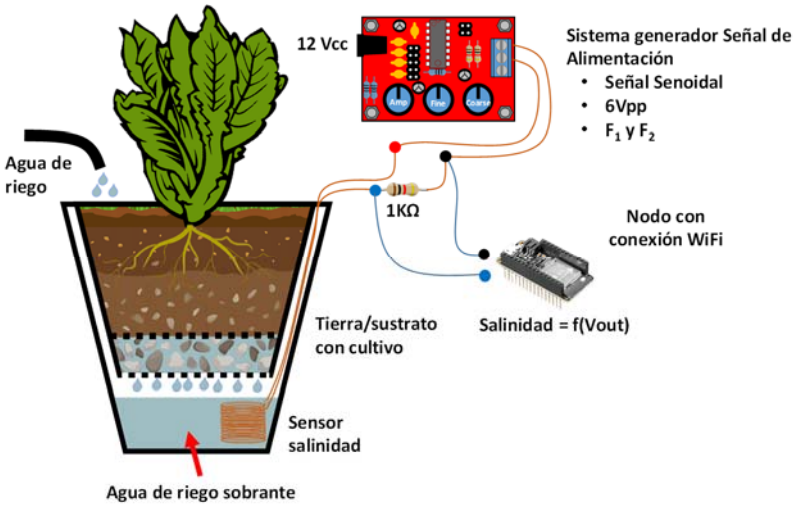


Fig. 5. Sistema de sensado instalado en cada canaleta.

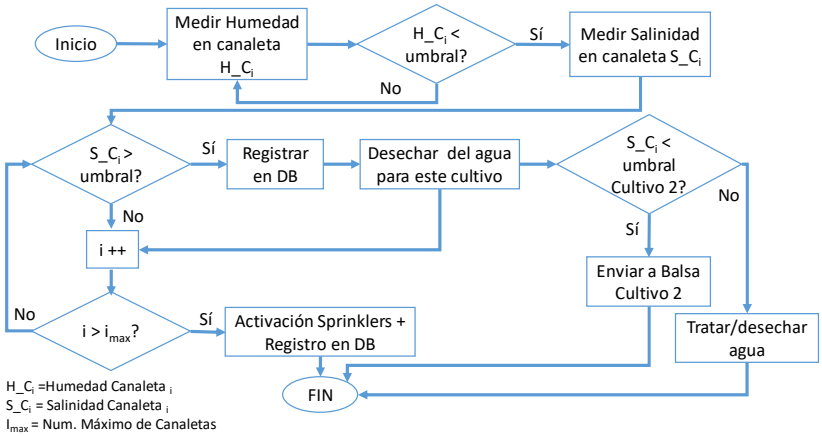


Fig. 6. Algoritmo de toma de decisiones para determinar el uso del agua.

El algoritmo también decidirá las acciones a tomar en función de los valores registrados.

En la Fig. 6 se presenta el algoritmo completo de control de riego del sistema. En un primer paso, el sistema realiza una verificación de cada canal de riego para determinar si es necesario aplicar riego en ese momento. Si no se requiere riego, el sistema entra en un estado de espera hasta el próximo ciclo de comprobación de la humedad. En caso de que un canal necesite ser regado, el siguiente paso consiste en verificar si el agua almacenada en ese canal tiene niveles de salinidad adecuados. Es importante recordar que la salinidad del agua puede deberse a la presencia de sales disueltas en el agua misma, a los nutrientes presentes o a los fertilizantes utilizados. Si el canal necesita riego y el agua en él es adecuada en términos de salinidad, se procede a recircular el agua para aprovecharla al máximo. Sin embargo, si el nivel de salinidad del agua es excesivo en ese canal, se lleva a cabo

una evaluación adicional. Se verifica si el agua podría ser utilizada para el riego de otros cultivos, como los tomates. Algunos estudios han sugerido que ciertos cultivos, como los tomates, pueden prosperar en aguas con un cierto grado de salinidad. Si se determina que el agua es adecuada para otros cultivos, se redirige hacia ellos. En caso contrario, se debe considerar la necesidad de filtrar el agua o destinarla a otros usos.

IV. Material y Métodos

En esta sección se describen, por un lado, los materiales que han sido utilizados para el desarrollo de la parte experimental para llevar a cabo este estudio. Por otro lado, se describe la metodología, así como el protocolo que se ha utilizado para la preparación y obtención de los datos.

A. Material utilizado

Para la preparación de las muestras se ha utilizado, sustrato de turba, agua dulce de grifo y NaCl (sal común).

En cuanto al material de laboratorio: báscula de precisión, donde se ha pesado la sal y el sustrato.

Además, para medir el volumen de agua el uso de una probeta de vidrio ha sido necesaria. Se han utilizado vasos de precipitado para la recogida de la muestra y la homogenización de ésta. Para el filtrado, se ha usado papel de filtro y embudo de cristal. Además, el líquido resultante a testear, se ha recogido con pipetas Pasteur y se ha depositado en tubos de vidrio. Finalmente, para la toma de medidas se ha utilizado el conductímetro comercial.

B. Preparación de las muestras

Para el procesamiento de las muestras, en primer lugar, se pesan 50 mg de sustrato. A continuación, mediante el uso de una probeta medimos 50 ml de agua dulce de grifo. Al agua utilizada se le añade la cantidad correspondiente de NaCl. El NaCl, es disuelta en el agua. La disolución de agua de grifo junto con el NaCl, es vertida en un vaso de precipitado junto con el sustrato y se homogeneiza. A continuación, se mide la humedad del suelo mediante un sistema de electrodos. La muestra es vertida al embudo de cristal el cual contiene el papel de filtro que permitirá el filtraje de la muestra.

Fig. 7. Muestras de sustrato con diferentes concentraciones de NaCl durante el filtraje del agua.

Las muestras son dejadas en reposo durante 24 horas como se muestra en la Fig. 7. Pasado este tiempo, el agua filtrada de la muestra es recogida en un tubo de ensayo para su posterior procesamiento con el conductímetro comercial. Se debe tener en cuenta que se necesita un mínimo de 5 ml para poder medir la conductividad de la muestra. De esta forma se consigue cubrir por completo el electrodo. En la tabla II, se muestra los patrones preparados para la realización de la recta patrón. Para todos los puntos los gramos de sustrato y el volumen de agua se mantienen constante. Cabe destacar, que para cada uno de los patrones se han realizado tres replicas, permitiendo así minimizar los errores.

C. Procesamiento de las muestras

Para la obtención de los resultados, se ha utilizado el conductímetro comercial citado anteriormente. Los resultados del conductímetro han sido comparados con las bobinas desarrolladas, permitiendo validar el funcionamiento de las bobinas.

En cuanto al procesamiento de las muestras mediante la implementación de las 2 bobinas testeadas, se utilizan 5 ml del lavado obtenido mediante la filtración del sustrato con las diferentes concentraciones de sal. La bobina es sumergida por completo en el interior del líquido. Es necesario una fuente de alimentación que nos permite establecer la frecuencia de trabajo, así como un osciloscopio para poder obtener los datos de voltaje pico-pico.

Tabla II
Características de las Muestras usadas

Muestra	Cantidad sustrato (g)s	Volumen de agua (ml)	NaCl (mg/50ml)	Concentración NaCl (g/l)
P1	50	50	0	0
P2	50	50	25	0.5
P3	50	50	50	1
P4	50	50	75	1.5
P5	50	50	100	2
P6	50	50	125	2.5
P7	50	50	150	3
P8	50	50	175	3.5

V. Resultados

Esta sección muestra los resultados de las pruebas realizadas con ambos prototipos de bobinas, así como el proceso de validación de los modelos matemáticos extraídos. Por último, se muestra el tráfico de red que se generaría en una instalación agrícola de estas características, en función de la cantidad de nodos presentes en la red.

A. Resultados del sensor de salinidad

En la Fig. 8, se muestran los valores en el eje X de las concentraciones de sal utilizada en g/l frente al voltaje en mV. En azul se representan los valores de la bobina 1, mientras que en naranja se muestra los datos de la bobina 2. Se observa que ambas disminuyen el voltaje cuando mayor es la concentración de sal en agua de riego. Sin embargo, cabe destacar, que la bobina dos presenta mayores diferencias que la bobina 1.

En la Fig. 9, se muestran los valores de voltaje (mV) para la bobina 1 frente a la conductividad (mS/cm). La

línea de tendencia muestra una disminución de los valores de conductividad para los voltajes con mayor valor.

La Fig. 10 muestra los valores del voltaje (mV) frente a la conductividad (mS/cm) para la bobina 2. Al igual que en caso de la bobina 1, la pendiente de la línea de tendencia es negativa. Cuanto mayor es el valor de voltaje, menor es la conductividad de las muestras.

Teniendo en cuenta los datos obtenidos (Fig. 9 y Fig. 10), y los valores de R^2, la bobina 2 presenta mejores resultados que la 1, siendo su R^2 de 0.8329 frente al R^2 de la bobina 1 que es 0.6493.

B. Validación del funcionamiento del sensor

Por otro lado, la Tabla III, presenta las muestras preparadas para la validación del sistema. En este caso, igual que para las muestras de la recta patrón se mantienen los gramos de sustrato y el volumen de agua contantes. Como podemos observar, la bobina que menor error comete en la estimación de la salinidad del agua es la bobina 2, con un error máximo de un 10%.

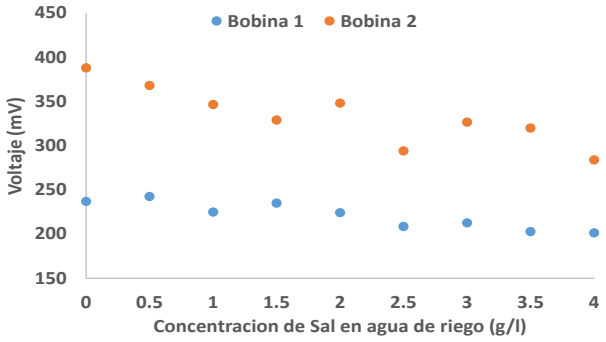


Fig. 8. Tensión (mV) de salida medida en función de la concentración de sal (g/l) para ambas bobinas.

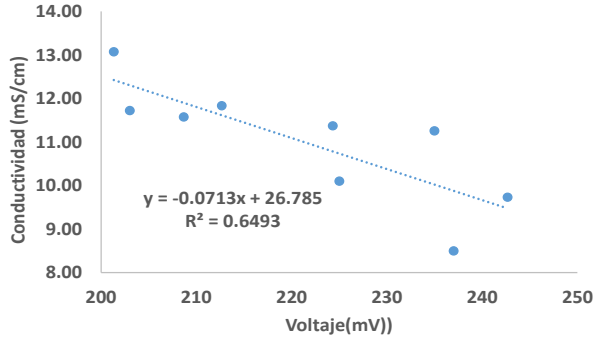


Fig. 9. Conductividad (mS/cm) en función de la tensión medida (mV) para la bobina 1.

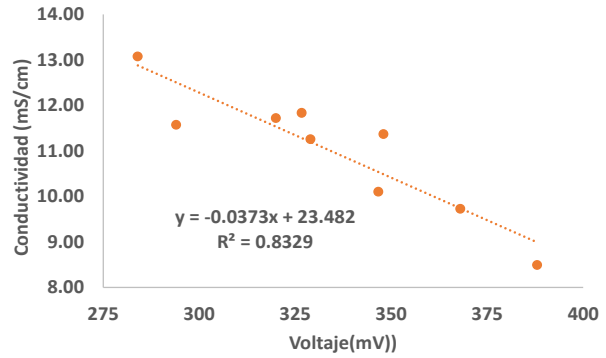


Fig. 10. Conductividad (mS/cm) en función de la tensión medida (mV) para la bobina 2.

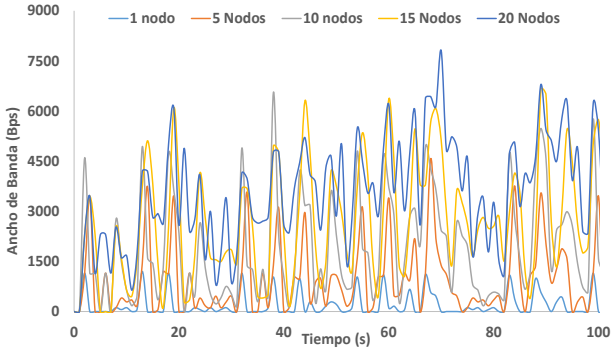


Fig. 11. Ancho de banda en bps para configuraciones de redes con distinto número de nodos inalámbricos.

C. Tráfico de red en función del número de dispositivos sensores.

Finalmente, se mide el tráfico de red en términos de ancho de banda consumido para diferentes números de nodos con el fin de determinar parámetros de diseño de la red. Para ello, se han empleado módulos Adafruit Huzzah ESP32. La Fig. 11 muestra el ancho de banda consumido para redes con diferentes números de red, como observamos el consumo medio de un nodo transmitiendo son aproximadamente unos 650 -700Bps, usando una conexión IEEE 802.11n y considerando que cada nodo transmite el dato medido cada 5 segundos.

VI. Conclusiones y Trabajos Futuros

La agricultura de precisión y en especial la agricultura vertical está en auge por la posibilidad de reducir el espacio necesario para llevar a cabo un cultivo. Sin embargo, podemos seguir aplicando tecnología en estas instalaciones desde el punto de vista de la reducción del uso de recursos naturales.

Tabla III
Validación con las muestras desconocidas

Muestra desconocida	Sustrato (g)	Volumen agua (ml)	NaCl (mg)	Valor real conductímetro (mS/cm)	Salinidad medida con Bobina 1 (mV)	Error con la Bobina 1 (%)	Salinidad con Bobina 2 (mV)	Error con la Bobina 2 (%)
M1	50	50	60	10.72	14.4231	34.50%	10.8746	1.41%
M2	50	50	110	12.21	17.3464	42.11%	11.919	2.36%
M3	50	50	112.5	11.84	17.59595	48.57%	11.6206	1.88%
M4	50	50	160	13.48	16.3482	21.28%	12.0682	10.47%
M5	50	50	162.5	13.3	17.7742	33.64%	12.5158	5.90%

Por esta razón, en este artículo se ha presentado una red inalámbrica de sensores encargada de monitorizar la salinidad del agua de los cultivos con el objetivo de analizar y determinar si el agua sobrante del riego de las mismas podría ser reutilizada en el mismo cultivo o bien por su degradación debía ser empleada en otro cultivo o destinada a otros usos. Para ello, se han probado 2 prototipos de sensores de salinidad basadas en una única bobina. Tras llevar a cabo las diferentes pruebas se ha determinado que el modelo de la bobina 2, ofrece resultados bastante buenos, con errores medios entorno al 5%. Así mismo, se han probado diferentes combinaciones de nodos y se ha medido el ancho de banda consumido, con el objetivo de determinar el diseño de la misma. Se ha observado que como término medio un nodo inalámbrico tipo ESP32 tiene un consumo medio de aproximadamente 700 Bps cuando únicamente envía datos medidos, por lo que podemos afirmar que su ocupación de red es bastante baja.

Como futuros trabajos, se quiere integrar algoritmos de IA sobre redes de entornos colaborativos que permitan la combinación de parámetros procedentes de diferentes tipos de nodos sensores y escenarios que permitan automatizar el proceso de riego y tener en cuenta otros parámetros como la procedencia del agua y la detección de pesticidas y fertilizantes en la misma.

AGRADECIMIENTOS

Este trabajo está parcialmente financiado por el Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2017–2020, proyecto PID2020-114467RR-C33/AEI/10.13039/501100011033, y el Plan Estatal de Investigación Científica, Técnica y de Innovación para el período 2021-2023, proyecto TED2021-131040B-C31).

REFERENCIAS

[1] P.L. Pingali, "Green revolution: Impacts, limits, and the path ahead." Proc. Natl. Acad. Sci. USA, vol. 109, pp. 12302–12308, 2012.

[2] R. Mumtaz, S. Baig, and I. Fatima, "Analysis of meteorological variations on wheat yield and its estimation using remotely sensed data. A case study of selected districts of Punjab Province, Pakistan (2001–14)," Ital. J. Agron., vol. 12, 2017.

[3] G. S. Hendricks, S. Shukla, F. M. Roka, R. P. Sishodia, T. A. Obreza, G. J. Hochmuth, and J. Colee, "Economic and environmental consequences of overfertilization under extreme weather conditions," J. Soil Water Conserv., vol. 74, pp. 160–171, 2019.

[4] J. Delgado, N. M. Short, D. P. Roberts, and B. Vandenberg, "Big data analysis for sustainable agriculture," FSUFS, vol. 3, p. 54, 2019.

[5] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," Artif. Intell. Agric., vol. 2, pp. 1–12, 2019.

[6] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. N. Hindia, "An overview of Internet of Things (IoT) and data analytics in agriculture: Benefits and challenges," IEEE Internet Things, vol. 5, pp. 3758–3773, 2018.

[7] C. Hedley, "The role of precision agriculture for improved nutrient management on farms," J. Sci. Food Agric., vol. 95, pp. 12–19, 2014.

[8] Y. Huang, Z. Chen, T. Yu, X. Huang, and X. Gu, "Agricultural remote sensing big data: Management and applications," J. Integr. Agric., vol. 7, pp. 1915–1931, 2018.

[9] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," Comput. Electron. Agric., vol. 143, pp. 23–37, 2017.

[10] N. Pavón-Pulido, J. A. López-Riquelme, R. Torres, R. Morais, and J. A. Pastor, "New trends in precision agriculture: A novel cloud-based system for enabling data storage and agricultural task planning and automation," Precis. Agric., vol. 18, pp. 1038–1068, 2017.

[11] K. Tzerakis, G. Psarras, and N. N. Kourgialas, "Developing an Open-Source IoT Platform for Optimal Irrigation Scheduling and Decision-Making: Implementation at Olive Grove Parcels," Water, vol. 15, no. 9, p. 1739, Apr. 2023.

[12] M. Garcia, D. Bri, S. Sendra, and J. Lloret, "Practical deployments of wireless sensor networks: a survey," Int. J. Adv. Networks Services, vol. 3, pp. 170–185, 2010.

[13] S. Basack, M. K. Loganathan, G. Goswami, and H. Khabbaz, "Saltwater intrusion into coastal aquifers and associated risk management: Critical review and research directives," J. Coastal Res., vol. 38, no. 3, pp. 654–672, 2022.

[14] A.D. Werner, M. Bakker, V.E.A. Post, A. Vandenbohede, C. Lu, B. Ataie-Ashtiani, C.T. Simmons, and D. Barry, "Seawater intrusion processes, investigation and management: Recent advances and future challenges," Adv. Water Resour., vol. 51, pp. 3–26, 2013.

[15] A. Singh, "Salinization and drainage problems of agricultural land," Irrig. Drain., vol. 69, pp. 844–853, 2020.

[16] D. A. Tran, M. Tsujimura, L. P. Vo, V. T. Nguyen, D. Kambuku, and T. D. Dang, "Hydrogeochemical characteristics of a multi-layered coastal aquifer system in the Mekong Delta, Vietnam," Environ. Geochem. Health, vol. 42, no. 2, pp. 661–680, 2020.

[17] S. S. D. Foster and C. J. Perry, "Improving groundwater resource accounting in irrigated areas: a prerequisite for promoting sustainable use," Hydrogeol. J., vol. 18, pp. 291–294, 2010.

[18] C. C. Faunt, M. Sneed, J. Traum, and J. T. Brandt, "Water availability and land subsidence in the Central Valley, California, USA," Hydrogeol. J., vol. 24, pp. 675–684, 2016.

[19] I. Cartwright, T. R. Weaver, C. T. Simmons, L. K. Fifield, C. R. Lawrence, R. Chisari, and S. Varley, "Physical hydrogeology and environmental isotopes to constrain the age, origins, and stability of a low-salinity groundwater lens formed by periodic river recharge: Murray Basin, Australia," J. Hydrol., vol. 380, pp. 203–221, 2010.

[20] A. Cobbenhagen, D. J. Antunes, M. van de Molengraft, and W. Heemels, "Opportunities for control engineering in arable precision agriculture," Annual Reviews in Control, vol. 51, pp. 47-55, 2021.

[21] P. Sanjeevi, S. Prasanna, B. Siva Kumar, G. Gunasekaran, I. Alagiri, and R. Vijay Anand, "Precision agriculture and farming using Internet of Things based on wireless sensor network," Transactions on Emerging Telecommunications Technologies, vol. 31, no. 12, p. e3978, 2020.

[22] M. S. I. Khan, A. Rahman, S. Islam, M. K. Nasir, S. S. Band, and A. Mosavi, "IoT and wireless sensor networking-based effluent treatment plant monitoring system," acta polytechnica hungarica, vol. 18, no. 10, pp. 205-224, 2021.

[23] R. R. A. Siregar, K. B. Seminar, S. Wahjuni, and E. Santosa, "Vertical Farming Perspectives in Support of Precision Agriculture Using Artificial Intelligence: A Review," Computers, vol. 11, no. 9, p. 135, 2022.

[24] D. K. Singh and R. Sobti, "Long-range real-time monitoring strategy for Precision Irrigation in urban and rural farming in society 5.0," Computers & Industrial Engineering, vol. 167, p. 107997, 2022.

[25] S. A. Siddiqi and Y. Al-Mulla, "Wireless Sensor Network System for Precision Irrigation using Soil and Plant Based Near-Real Time Monitoring Sensors," Procedia Computer Science, vol. 203, pp. 407-412, 2022.

[26] S. Sendra, S. Viciano-Tudela, A. Ivars-Palomares, and J. Lloret, "Low-Cost Water Conductivity Sensor Based on a Parallel Plate Capacitor for Precision Agriculture," in International Conference on Advanced Intelligent Systems for Sustainable Development, 2022: Springer, pp. 500-514.

[27] F. J. Diaz, A. Ahmad, S. Viciano-Tudela, L. Parra, S. Sendra, and J. Lloret, "Development of a Low-Cost Sensor to Optimise the Use of Fertilisers in Irrigation Systems," presented at the ICNS 2023: The Nineteenth International Conference on Networking and Services, Barcelona, Spain, pp. 34-39, 2023.

[28] L. Parra, S. Sendra, J. Lloret, J.J. Rodrigues, "Low cost wireless sensor network for salinity monitoring in mangrove forests". In Proc. Of 2014 IEE SENSORS, Valencia, Spain, November 2-5, 2014 (pp. 126-129).

[29] Características del integrado AD9850. Disponible en: https://www.analog.com/media/en/technical-documentation/data-sheets/AD9850.pdf [Acceso: 14 de junio de 2023]

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Diseño de una Plataforma para el Almacenamiento y Gestión de los datos en Redes IoT

Miguel Zaragoza-Esquerdo, Alberto Ivars-Palomares, Sandra Sendra, Jaime Lloret.
Instituto de Investigación para la Gestión Integrada de Zonas Costeras (IGIC), Escola Politècnica Superior de Gandia, Universitat Politècnica de València.
Paraninf 1, 46730 Grao de Gandia, España.
mizaes2@epsg.upv.es, aivapal@epsg.upv.es, sansenco@upv.es, jlloret@dcom.upv.es.

Con el auge de las redes de sensores y la preocupación por cómo aprovechar los datos procedentes de las redes desplegadas en distintos entornos, se hace necesario el uso de plataformas de almacenamiento y gestión de los mismo. Actualmente, existen soluciones comerciales (algunas de ellas gratuitas, pero con limitaciones importantes). Sin embargo, este hecho te obliga a depender siempre de esta plataforma para el correcto desempeño del servicio o aplicación. Por este motivo, en este artículo se presenta un sistema jerárquico novedoso que permite el diseño y despliegue de una plataforma para redes de monitorización e Internet de las cosas (IoT) personalizable para la recolección, almacenamiento y visualización de los datos en tiempo real. A lo largo del artículo se mostrará la arquitectura de red, así como los protocolos, tareas y servicios desplegados en cada capa. Finalmente, la plataforma será probada en términos de funcionamiento de los diferentes servicios y del tráfico de red registrado en función del número de nodos que conformarían la red de monitorización.

***Palabras Clave*- Plataforma, Internet de las cosas (IoT), Redes Inalámbricas de sensores, monitorización, almacenamiento de datos, gestión de datos**

I. INTRODUCCIÓN

El Internet de las cosas (IoT) tiene como función interconectar los objetos físicos a través del Internet. Estos objetos pueden estar conectados a sensores capaces de recopilar la información del entorno que les rodea [1][2]. La información recopilada puede ser utilizada para diferentes propósitos [3]. Por lo tanto, IoT es de vital importancia para la sociedad actual, ya que saber integrarlo de forma correcta con la ciencia y la tecnología, permitiría una gran conectividad a bienes y servicios [4].

Para las empresas es esencial monitorizar los entornos externos e internos con la finalidad de realizar metodologías acordes a los cambios del entorno [5]. Esto les facilita tener una mejor visión de los incidentes que puedan afectar a la empresa y así poder realizar una estrategia acorde a estos [6]. Por otro lado, también es indispensable el monitoreo del medio ambiente, pues monitorizar el medio ambiente puede ayudar a detectar emisiones que puedan dañarle y se pueden utilizar para la implementación de sistemas que ayuden a mitigarlas. Concretamente se puede utilizar en la detección de gases de efecto invernadero, que son una de las principales causas del cambio climático [7].

Para el almacenamiento de los datos obtenidos por dispositivos IoT es necesaria la implementación de una base de datos robusta y estable que permita almacenar grandes cantidades de datos. La principal problemática respecto a las bases de datos es que son altamente vulnerables ante ataques cibernéticos. Al no implementarse las políticas de seguridad adecuadas, las bases de datos muestran brechas que aprovechan los atacantes cibernéticos y esto puede producir que realicen acciones que puedan dañar los datos o quebrantar la privacidad de estos [8].

Existen múltiples opciones de bases de datos disponibles online que permiten almacenar datos en la nube en tiempo real [9]. Uno de los ejemplos más comunes es el servicio de google llamado Firebase. Esta base de datos es accesible libremente con ciertas limitaciones, pero el servicio de Google da unas reglas a seguir y permite que exista autentificación directamente desde el código [10].

La Tabla I muestra un resumen de algunas de las principales plataformas IoT para la gestión de datos y sus características. En ella, se incluyen opciones tanto gratuitas como de pago [11][12][13].

Otra opción es la instalación de una base de datos gestionable por el propio usuario. Uno de los ejemplos más comunes para datos IoT es la base de datos MongoDB, donde se envían los datos con el formato JSON [14].

Tabla I
PLATAFORMAS IOT PARA LA GESTION DE DATOS

Plataforma IoT	Modelo de acceso	Formato de datos Soportado	Lenguaje de programación soportado	Protocolos	Precio	Tecnología usada
AWS IoT Platform	*PaaS, IaaS*	*JSON*	*Java, C, NodeJS, Javascript, Python, SDK for Arduino, iOS,*	*HTTP, MQTT, Websockets*	*Pago cuando se ejecuta las funciones propias escritas*	*Todos los servicios de Amazon*
Microsoft Azure IoT Hub	*IaaS*	*JSON*	*.NET, UWP, Java, C, NodeJS, Ruby, Android, iOS*	*HTTP, AMQP, MQTT*	*Pago de acuerdo el número de dispositivos y mensajes por día*	*Azure Cosmos DB, Azure Tables, SQL database*
IBM Watson IoT platform	*PaaS, IaaS*	*JSON,CSV*	*C#, C, Python, Java, NodeJS*	*MQTT*	*Pago de acuerdo el número de dispositivos y mensajes por día*	*Cloudant NoSQL DB*
Google IoT Platform	*PaaS, IaaS*	*JSON*	*Go, Java, NET, Node.js, php, Python, Ruby*	*MQTT, HTPP*	*Precio por Mbyte*	*Servicios de Google*
Kaa IoT Platform	*IaaS*	*JSON*	*Java, C, C++*	*MQTT, CoAP, XMPP, TCP, HTTP*	*gratuito*	*NoSQL, MangoDB, analisis en tiepo real y visualizaciñon con Kaa*
ThingSpeak	*PaaS*	*JSON,XML*	*Matlab*	*MQTT API and REST*	*Gratuito*	*Matlab, dashboard and Matlab analytics, MySQL*
Carriots	*PaaS*	*JSON,XML*	*Java*	*MQTT*	*Pago por servicios*	*NoSQL Big- Database*
Temboo	*PaaS*	*Excel, CSV, XML,JSON*	*C, Java, Python, iOS, Android, javascript*	*HTTP, MQTT, CoAP*	*Gratuito para los primeros 100 dispositivos, despues pago por dispositivo*	*Microsoft Power BI, Google BigQuery*
Thinger.io	*PaaS*	*JSON*	---	*HTTP, MQTT*	---	*MongoDB*
Sentilo	*PaaS*	*JSON*	*C, Java*	*HTTP*	*Gratuito*	*Redis, Apache, PubSub, MongoDB, ElasticSearch*
Cosmos IoT	*PaaS*	*JSON*	*Java, Ruby*	*HTTP (RESTfull API)*	*Propietario*	*Azure Cosmos DB, Azure Tables, SQL database*
SmartThings	*PaaS*	*JSON*	*SmartThings SDK*	*HTTP (RESTfull API)*	*Propietario*	---
SenseWeb	*PaaS*	*HTML, Scalar number, XML,. Multimedia*	---	*HTTP, Web Service Interface*	*Gratuito*	---
IFTTT	*PaaS*	---	*No Programable*	*HTTP*	*Propietario*	---
FIWARE	*PaaS*	*JSON, CSV, Excel o XML*	*C, ++, Python*	*HTTP, MQTT, CoAP, Ultralight 2.0*	*Open Source, pero con pago por uso a conjuntos de datos*	*SQL, SGBD MongoDB*
FIREBASE	*PaaS*	*JSON*	*iOS, Android, C++, Web y Unity, Flutter, Node.js*	*HTTP (RESTfull API)*	*Pago por servicios*	*Servicios de google, No SQL,*

Esta segunda opción destaca porque utilizan bases de datos relacionales con alta facilidad para conectarse desde cualquier nodo mediante el uso de una conexión a Internet y directamente desde el código. Todo ello brinda una estructura muy cómoda para trabajar con un volumen de datos bajo como puede ser los datos obtenidos por los sensores [15].

Para trabajar con redes de sensores donde el acceso a Internet es limitado y la problemática de la arquitectura debe de ser muy concreta se requiere de una estructura más personalizable [16]. El hecho de trabajar la base de datos

en local donde toda la arquitectura está localizada nos da más accesibilidad en el medio y por tanto más portabilidad [17]. Al tener localizado donde va a llegar todo el flujo de datos reducimos el tráfico de red, en este caso un servidor web. A parte de ello podemos monitorizar más fácil los datos tanto en un portal web como exportarlo en formato CSV de forma rápida y sencilla.

En este artículo está estructurado como sigue. En la sección 2 se muestran los trabajos relacionados. En la sección 3 se muestra la arquitectura hecha para resolver esta problemática y sus tecnologías, así como una descripción de los protocolos utilizados. En la sección 4 se mostrará el rendimiento del sistema implementado. La sección 5 incluye las conclusiones obtenidas en esta investigación y los trabajos futuros.

II. Trabajos Relacionados

En esta sección analizamos varios artículos donde se ha empleado una metodología similar a la que se va a emplear en la presente investigación. Consiste en una red de nodos conectados que envían datos a través de distintos protocolos de enlace como puede ser WIFI o ETHERNET a un servidor web mediante peticiones HTTP.

En primer lugar, el protocolo escogido para transmitir los datos del nodo IoT a la base de datos es IEEE 802.11b/g/n por su alta facilidad de implementación y su gran operabilidad. Un ejemplo se puede encontrar [18]. En esta investigación muestra un sistema domótico para ayudar en las tareas domésticas a personas con movilidad reducida. Entre estas tareas se encontraba encender las luces, poder controlar la apertura y el cierre de cortinas. Este sistema utiliza un servidor que se conecta a una tarjeta de red inalámbrica que utiliza el protocolo IEEE 802.11b/g/n y a su vez utiliza una tarjeta relé que permite la conexión del mundo análogo y el digital. Por otro lado, generalmente el servidor con la base de datos se encuentra conectado en la red cableada, utilizando el protocolo IEEE 802.3u, pues permite mayor disponibilidad y ancho de banda para recibir los datos enviados por los sensores. Aunque también se encuentran casos en los que el propio dispositivo IoT está conectado a la red cableada como en [19]. En este artículo se muestra el caso de un Arduino que está configurado como un servidor y a su vez es el responsable de recibir la señal del sensor y gracias a estar conectado puede subir los datos a la red.

En segundo lugar, para la transmisión de datos del servidor a la página web se hace mediante peticiones HTTP. En la web se emplea, en la parte de cliente, el framework de javascript REACT y Bootstrap. En [20] se muestra la implementación de aplicaciones isomórficas con Javascript. En este artículo se creó una aplicación que permite crear y listar evaluaciones para proyectos de alumnos. Donde los requerimientos funcionales eran listar las evaluaciones creadas, seleccionar una evaluación y registrar una nota para cada uno de los criterios de calificación.

Por último, para la base de datos se requiere la utilización de diferentes tipos de servicios. Por una parte un servidor express de node js para la transmisión del servidor a la base de datos, donde se emplea mysql usando como interfaz phpmyAdmin. En [21] se puede encontrar un ejemplo de aplicación de administración con Angular, Node y Express para una aplicación Django donde su objetivo era la implementación de una aplicación web, realizada con Node.js y Express para la parte del segundo servidor, y AngularJS y Bootstrap para la parte del cliente. El sistema implementado permite gestionar la parte de administración de una aplicación web Django, donde se podrán crear, editar, eliminar y consultar las distintas entidades que forman la base de dicha aplicación.

III. Propuesta

Esta sección presenta la descripción general de la arquitectura de red propuesta para el despliegue de soluciones IoT. Así mismo se describen diferentes modos de implementación con sus ventajas y desventajas.

A. Descripción general

Cuando una red transporta grandes cantidades de datos, debe haber una arquitectura adecuada encargada de controlar el flujo de datos, enviar y recibir señales, realizar cálculos y manejar el almacenamiento.

Cuando hablamos de entornos IoT, no existe una arquitectura específica en la que se esté de acuerdo universalmente [22]. Sin embargo, en todas ellas, debe haber presentes, al menos, tres capas fundamentales, es decir, una capa que incluya los sistemas de captación de variables, una capa que se encargue de encaminar la información y una capa de gestión de las aplicaciones que facilitan el acceso a la información medida. La arquitectura aceptada debe garantizar que admita varios tipos de protocolos y aplicaciones de red. Además, la arquitectura debe verificar los requisitos estándar como seguridad, portabilidad, confiabilidad y estructura para garantizar la protección de los datos recibidos.

Para nuestro sistema, proponemos una arquitectura de red de 4 capas como la que muestra la Fig. 1. Cada una de las capas desempeña una serie de funciones bien definidas que sirven para proveer de datos a las capas superiores. Las capas de nuestra arquitectura se encargan de realizar las siguientes tareas:

- **Capa de percepción:** la capa de percepción que contiene dispositivos como sensores, actuadores y máquinas que tienen la capacidad de detectar, calcular y conectarse con otros dispositivos. Los sensores detectan cambios físicos en el entorno y recopilan información. Esta capa se puede llamar el lado del cliente, ya que se adapta a la ubicación o dirección del cliente. La información recopilada se transmite a la capa de agregación o la puerta de enlace de IoT. Esta capa combina y procesa los datos. Los operadores controlan esta capa y participan servidores.

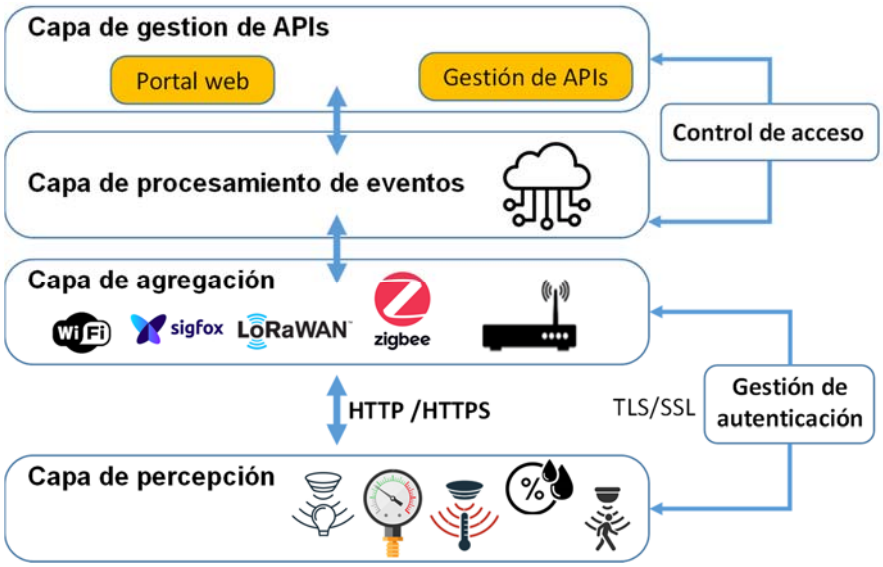


Fig. 1. Arquitectura de red para sistema IoT

- **Capa de procesamiento de eventos:** La siguiente capa se establece en la nube y se denomina capa de procesamiento de eventos. Aunque podríamos trabajar con un cloud local, en lugar de uno remoto. Esta capa también puede contener varios algoritmos y scripts para el manejo de la información obtenida de la capa de sensores. Finalmente, la capa de agregación es responsable de recopilar todos los datos e información provenientes de varios dispositivos de IoT.
- **Capa de conectividad o de agregación:** esta capa transporta los datos físicos desde los sensores y dispositivos de IoT hacia los servidores. Luego, transporta las respuestas generadas de vuelta hacia los dispositivos. El transporte se lleva a cabo a través de una red basada en tecnologías como IEEE 802.11, Ethernet, Zigbee, Bluetooth, LoRa y redes celulares.
- **Capa de gestión de APIs:** La capa final es la capa de gestión de API, la cual se comporta como la interfaz entre las aplicaciones de terceros y la infraestructura o los usuarios. En esta capa también se encarga de administrar sus dispositivos, la identidad y acceso, que brindan seguridad y protección, respaldan todo el sistema en cada etapa.

Esta arquitectura no es única, aunque sí da cabida a un gran número de protocolos, en función del tipo de servicio que se pretenda desarrollar.

B. Cómo montar la red.

Para poder establecer una conexión entre los nodos presentes en la red y la base de datos, podemos hacerlo de modo directo (Ver Fig. 2), donde el nodo sensor conecta directamente con el servidor MySQL e interactua con el servidor MySQLusando el protocolo MySQL [23] o bien, de modo indirecto (Fig. 3), donde el nodo se conecta indirectamente con el servidor MySQL a través de un servidor web y usando el protocolo HTTL/HTTPS.

Interactuar directamente con MySQL puede parecer una solución simple y viable. Sin embargo, tiene diversos problemas que deben ser analizados, en especial aquellos relacionados con la seguridad [24]:

- Por una parte, debemos dar acceso remoto a una cuenta o usuario MySQL. Esto tiene un problema de seguridad importante, incluso en el caso de limitar los privilegios a dicha cuenta.
- El nodo necesita guardar y enviar peticiones MySQL al servidor MySQL. Esto implica que el código programado y por tanto ejecutado en el nodo sensor es más largo y complejo, lo que implicará mayor consumo de CPU y memoria de almacenamiento.
- El servidor MySQL debe procesar datos en bruto, lo que implica mayor complejidad del código MySQL.
- Finalmente, muchas de las librerías MySQL usadas para nodos tipo Arduino, como ESP32, Leonardo, Wemos, etc. No soportan SSL/TLS lo que implicará que el nombre de usuario y la contraseña, serán enviadas sin encriptación.

Por otra parte, la interacción indirecta con MySQL, a través de peticiones HTTP/HTTPS, soluciona muchos de estos problemas. Su modo de trabajo puede simplificarse en 3 pasos:

- **Paso 1:** el nodo toma los datos, los incluye en una petición HTTP/HTTPS y los envia al servidor web.
- **Paso 2:** el servidor web ejecuta un script de PHP que gestiona las peticiones llegadas desde el nodo sensor.
- **Paso 3:** el script de php extrae los datos recibidos en la petición http, los procesa e interactúa con la base de dato MySQL

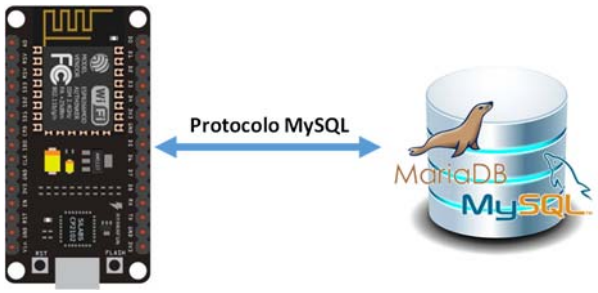


Fig. 2. Conexión directa de un nodo con MySQL

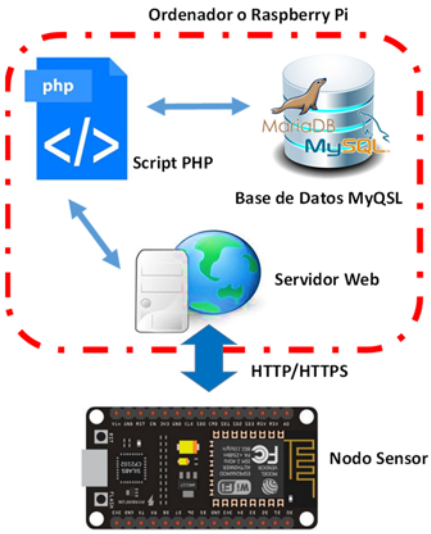


Fig. 3. Conexión indirecta de un nodo con una base de datos MySQL

Existen varias ventajas al utilizar este enfoque en comparación con el método directo. Estas son algunas de ellas:

- **Mayor seguridad:** Almacenar la cuenta de usuario de MySQL (nombre de usuario y contraseña) en el servidor proporciona una capa adicional de seguridad para los datos. Esto garantiza que estén protegidos. Además, es posible limitar el acceso a la cuenta de MySQL solo desde localhost.
- **Reducción de la carga en el nodo sensor y MySQL:** Los datos se procesan en un archivo de script PHP, lo que permite reducir la complejidad del servicio en el nodo y en en la base de datos o servidor MySQL. Además, resulta más sencillo procesar los datos en el archivo PHP. Para evitar problemas de memoria en el nodo sensor, el archivo de script PHP envía únicamente los datos relevantes después de procesarlos. Además, generar una solicitud HTTP o HTTPS (si se usa la versión segura del protocolo) utilizando un nodo sensor es muy sencillo.
- **Encriptación de datos:** Los datos se encriptan en las solicitudes HTTP, lo cual proporciona una capa adicional de seguridad. Esto garantiza que los datos transmitidos estén protegidos contra posibles accesos no autorizados.

Por tanto, y aunque como se ha visto, es posible desplegar la red de distintas formas, en la arquitectura propuesta, se opta por el uso de una conexión de los nodos sensores con la base de datos MySQL de forma indirecta.

C. Intercambio de paquetes dentro de la arquitectura IoT propuesta

El intercambio de mensajes entre los diferentes elementos presentes en la red, se muestra en la Fig. 4. Como observamos, cuando un nodo registra un nuevo dato (o bien de manera periódica) el nodo enviará una petición HTTP mediante un POST al servidor web, quien puede además establecer un sistema de autenticación para su acceso, por ejemplo, mediante usuario y contraseña. El siguiente paso será publicar los datos en la base de datos a través del script Pub_Datos.php. En nuestro caso, forzamos a que sea el propio servidor que hospede todos los servicios quien proporcione la hora y día en el que el dato es registrado, por ello, la base de datos MySQL almacenará cada dato, con su marca temporal.

Finalmente, si se solicita acceso a la web para visualizar los datos, los scripts Graf_Dato.php y Nodo_Dato.php solicitarán acceso a la base de datos, la cual le devolverá todos los valores almacenados y los mostrará en formato numérico o gráfico.

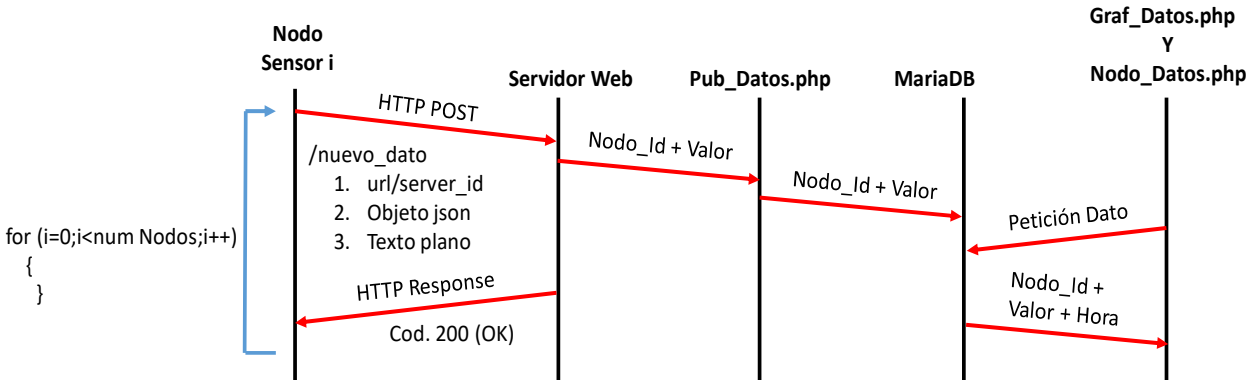


Fig. 4. Intercambio de mensajes realizado desde que un nodo registra un nuevo dato hasta que se almacena en la base de datos

Este intercambio de mensajes se repetirá tantas veces como nodos haya presentes en la red, teniendo en cuenta que, con toda probabilidad, no estén enviando los datos justo en el mismo momento.

IV. Implementación de la Plataforma IoT y Resultados de Tráfico

Esta sección describe brevemente el modo en el que se ha implementado la plataforma IoT para el almacenamiento y gestión de los datos. Finalmente, la red entera es testeada con redes de diferentes tamaños. Con ello, se pretende cuantificar el ancho de banda útil de la red así como el año de banda consumido por el tráfico de paquetes erróneos.

A. Implementación práctica de la plataforma.

Para llevar a cabo nuestro experimento y desplegar nuestra plataforma optamos por usar un ordenador HP con un procesador i7 de 11º generación, en el que instalamos un servidor web Apache 2 y un servidor y cliente MySQL, en concreto se decide usar Maria BD [25]. Por otra parte, se emplean diversos nodos sensores inalámbricos tipo ESP 32 [26] con capacidad inalámbrica. Se establece que cada nodo haga envío de sus datos cada 20 segundos. Tanto el ordenador como todos los nodos sensores son conectados a una red local creada mediante un punto de acceso IEEE 802.11n.

Dentro del servidor MySQL, se crea una base de datos con una tabla por cada nodo incluido en la red, aunque sería posible trabajar con una única tabla, identificando cada nodo que envía los datos.

El script en php llamado *Pub_Datos.php* es el encargado de recibir las diferentes peticiones http recibidas de cada nodo y enviarlos a la base de datos con la estructura establecida inicialmente. Tras almacenar los datos en las diferentes tablas, los ficheros *Nodo_Datos.php* y *Graf_Datos.php* harán consultas cada minuto, los cuales se mostrarán en un navegador web convencional.

Finalmente, si se pretende consultar los datos medidos desde internet, deberíamos conocer el nombre de dominio o la IP publica donde se ha desplegado la red y acceder desde cualquier navegador. La Fig. 5 muestra la red montada.

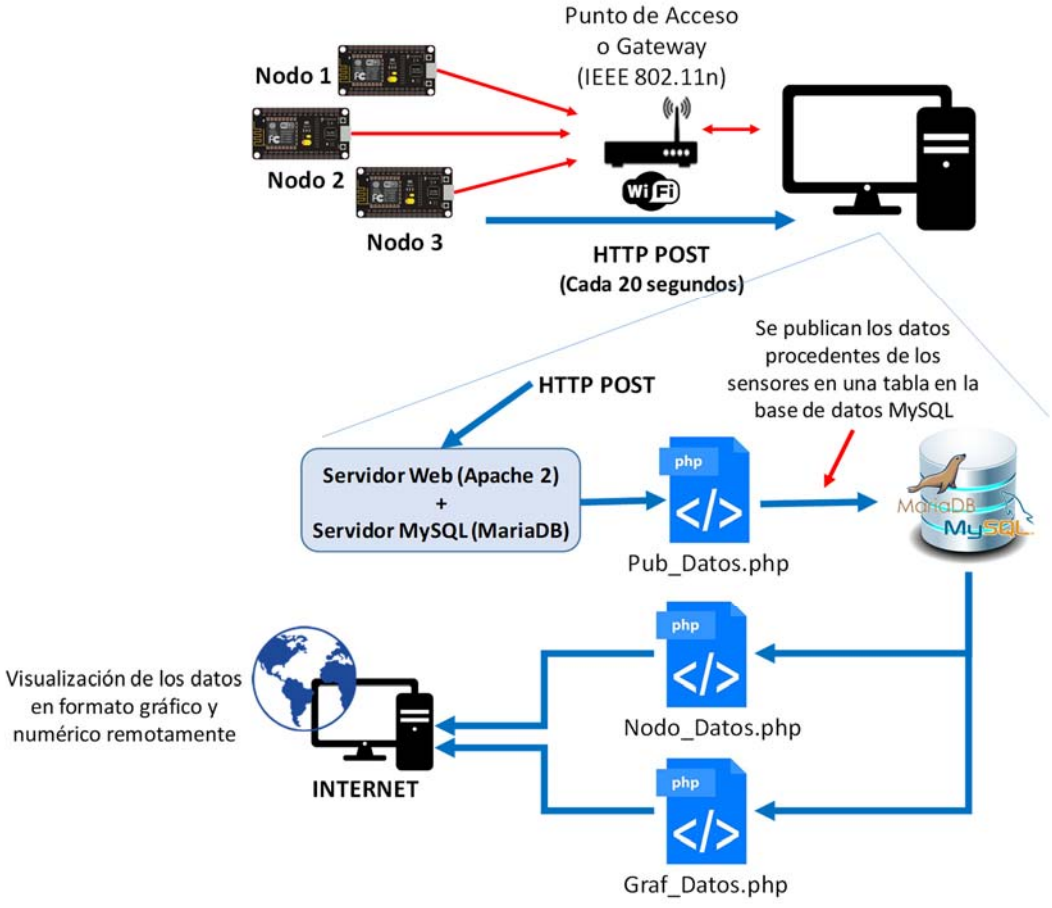


Fig. 5. Sistema desplegado

B. Tráfico de red en función del número de dispositivos sensores.

Con el objetivo de determinar el funcionamiento de la red, se mide el tráfico generado en la misma, en términos de ancho de banda consumido para diferentes números de nodos y el ancho de banda consumido debido a la presencia de errores en las transmisiones por la presencia de demasiados nodos transmitiendo simultáneamente en la red. Para ello, se han empleado diferentes módulos Adafruit Huzzah ESP32. La Fig. 5 muestra el ancho de banda consumido para redes con diferentes números de red, como observamos el consumo medio de un nodo transmitiendo son aproximadamente unos 650 -700 Bps,

usando una conexión IEEE 802.11n y considerando que cada nodo transmite el dato medido cada 20 segundos.

En la Fig, 7 se muestra el ancho de banda medio para cada tamaño de red. Como se observa, el ancho de banda crece de manera prácticamente lineal, habiendo un ancho de banda medio de unos 250 Bps por nodo. Finalmente, se comprueba el ancho de banda medio consumido debido al envío de bytes erróneos (Ver Fig. 8), debido a interferencias o simplemente errores de transmisión. De nuevo, se observa que, a mayor número de nodos, mayor cantidad de bytes erróneos. Así mismo, cabe destacar que la cantidad de paquetes erróneos es bastante grande (aproximadamente 1/5 del tráfico total de la red). Creemos que este valor es tan grande porque el período entre medidas es bastante pequeño y la presencia de tantos nodos en la red inalámbrica transmitiendo casi simultáneamente genera muchas interferencias.

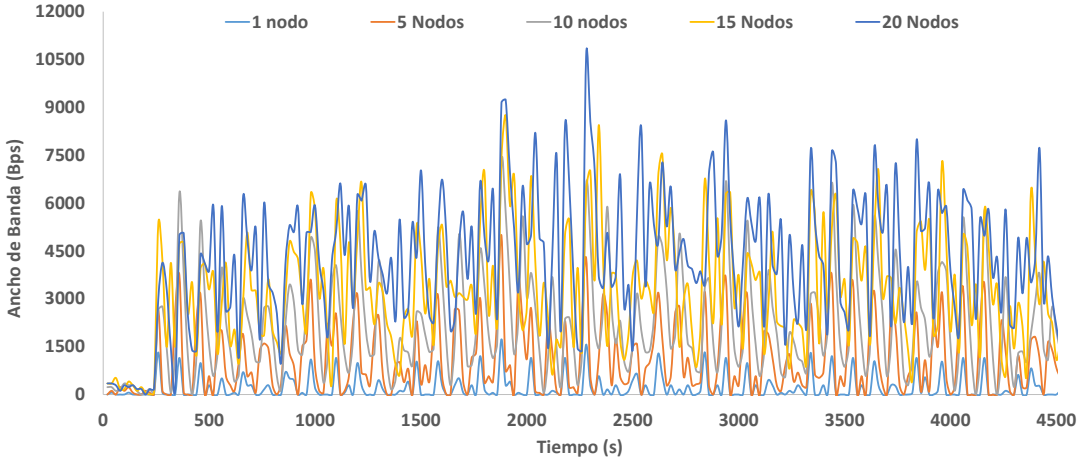


Fig. 6. Ancho de banda consumido por envío de datos

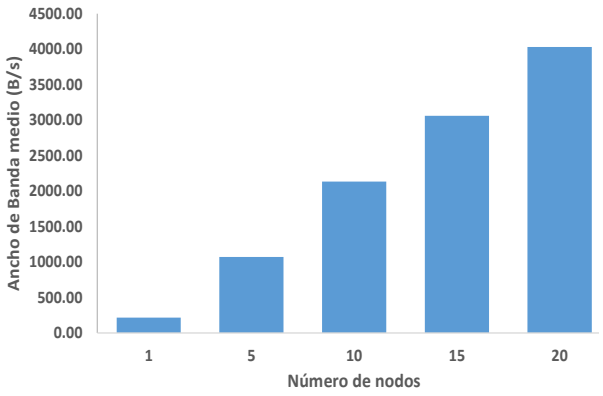


Fig. 7. Ancho de banda medio

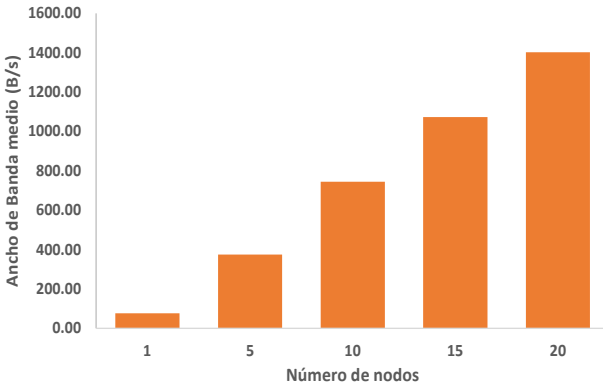


Fig. 8. Ancho de banda medio debido a errores

V. Conclusiones y Trabajos Futuros

En los últimos años, el concepto del Internet de las cosas y el desarrollo de aplicaciones y servicios bajo este nombre ha experimentado un crecimiento extremadamente grande y rápido. Con la creación de este tipo de servicios, se hace necesario la creación de plataformas que permitan la gestión y almacenamiento de los datos. Por ello, en este artículo hemos presentado la propuesta de una arquitectura de red para soluciones IoT fácilmente personalizable y adaptable a las necesidades de la aplicación. Se ha propuesto una arquitectura de 4 capas que incluye todas las tareas necesarias para la correcta comunicación de los dispositivos. Finalmente se ha implementado y probado su correcto funcionamiento para finalmente comprobar el rendimiento de la red. Se han probado diferentes redes con diferente número de nodos y se ha medido el ancho de banda consumido, el ancho de banda medio y el ancho de banda medio consumido por la existencia de paquetes erróneos. Se ha observado que el ancho de banda de un nodo inalámbrico tipo ESP32 tiene un consumo medio de aproximadamente 250 Bps cuando únicamente envía datos medidos, sin embargo, se ha observado que la cantidad de tráfico de paquetes erróneos es aproximadamente 1/5 del tráfico total de la red.

Como futuros trabajos, se quiere incluir un módulo de aprendizaje máquina e inteligencia artificial que permita configurarse y adaptarse a la aplicación o servicio desplegado con el objetivo de mejorar las decisiones

futuras y acciones que se puedan desempeñar en la propia red. Por otra parte, se quiere estudiar la posibilidad de trabajar con redes multisalto para aplicaciones en escenarios extensos. Así mismo, cabe destacar que esta propuesta forma parte de del desarrollo de proyectos de investigación sobre agricultura de precisión, que están actualmente en marcha. Tras desplegar los nodos en las diferentes ubicaciones, los datos serán recogidos en un observatorio, haciendo accesibles los mismos de manera pública.

AGRADECIMIENTOS

Este trabajo está parcialmente financiado por el Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2017–2020, proyecto PID2020-114467RR-C33/AEI/10.13039/501100011033, el Plan Estatal de Investigación Científica, Técnica y de Innovación para el período 2021-2023, proyecto TED2021-131040B-C31) y la Generalitat Valenciana a través del programa Investigo (Proyecto INVEST/2022/467)

REFERENCIAS

[1] M Garcia, D Bri, S Sendra, J Lloret, “Practical deployments of wireless sensor networks: a survey”, International Journal on Advances in Networks and Services 2010, vol.3, no. 12, pp.170-185.

[2] D Bri, M Garcia, J Lloret, P Dini, Real deployments of wireless sensor networks, The Third International Conference on Sensor Technologies and Applications. SENSORCOMM 2009. June 18-23, 2009 - Athens/Glyfada, Greece

[3] I. Bonilla-Fabela, A. Tavizon-Salazar, M. Morales-Escobar, L.T. Guajardo-Muñoz, C.I. Laines-Alamina, “IoT, el internet de las cosas y la innovación de sus aplicaciones”. Vinculatégica efan, 2016, vol. 2, no.1, pp.2313-2340.

[4] R. del C. Valadez Hegler, "La importancia del internet de las cosas en una publicación científica". Management Review 2019, vol. 3, no. 3.

[5] A Rego, A Canovas, JM Jiménez, J Lloret, “An intelligent system for video surveillance in IoT environments”, IEEE Access 2018, vol. 6, pp. 31580-31598.

[6] L.F. Velásquez Franco, “Metodología para monitorear riesgos estratégicos provenientes del entorno externo” (Methodology to monitor strategic risks) (Doctoral dissertation, Universidad EAFIT). 2014

[7] N. Mayhew, “El OIEA y el cambio climático: adaptación, monitorización y mitigación”. Boletín del OIEA, 5. 2018

[8] O.A. Escalante Quimis, “Prototipo de sistema de seguridad de base de datos en organizaciones públicas para mitigar ataques cibernéticos en Latinoamérica”, Univ. Politécnica Salesiana (Ecuador). 2021. Disponible en: http://dspace.ups.edu.ec/handle/123456789/20576 (Últ. acceso: 15/06/2023)..

[9] O. Mejıa, “Computación en la nube”. ContactoS 2011, vol. 80, pp. 45-52.

[10] D. Martínez Martínez, “Diseño e implementación de un sistema de control y monitorización remoto de bajo coste mediante Raspberry Pi y gestión en la nube con Firebase”. Universitat Politècnica de València. 2022. Disponible en: http://hdl.handle.net/10251/189520 (Últ. acceso: 15/06/2023).

[11] A. Zaballos, A. Briones, A. Massa, P. Centelles, V. Caballero, “A smart campus’ digital twin for sustainable comfort monitoring”. Sustainability 2020, vol.12, no. 21, pp. 9196.

[12] S. Traboulsi, S. Knauth, Stefan. “Towards implementation of an IoT analysis system for buildings environmental data and workplace well-being with an IoT open software”, Procedia Computer Science 2020, vol. 170, pp. 341-346.

[13] T. Domínguez-Bolaño, O. Campos, V. Barral, C. J. Escudero, J. A. García-Naya, "An overview of IoT architectures, technologies, and existing open-source projects", Internet of Things 2022, vol. 20, pp. 100626

[14] M. J. Gonzales Reinoso, “Estudio comparativo entre los servicios web utilizando formato de transferencia XNL y JSON para la optimización del tiempo de repuesta al consumir grandes volúmenes de datos desde dispositivos móviles”, Universidad Técnica Estatal de Quevedo (Ecuador). 2015. Disponible en: http://repositorio.uteq.edu.ec/handle/43000/4045 (Últ. acceso: 15/06/2023).

[15] E. M. Rubio Díaz, “Plataforma para persistencia de sensores IoT en bases de datos NoSQL”, Universida de Jaen, 2021. Disponible en: https://hdl.handle.net/10953.1/18948 (Últ. acceso: 15/06/2023).

[16] D. M. Berbes Villalón, M. E. Díaz Aguirre, T. Delgado Fernández, L. Sánchez Jiménez, “API para el desarrollo de aplicaciones IoT personalizadas usando FIWARE API for the development of custom IoT applications using FIWARE”. Revista cubana de transformación digital 2021. Disponible en: http://portal.amelica.org/ameli/journal/389/3893118005/3893118005.pdf (Últ. acceso: 15/06/2023).

[17] J. M. Lozano Banqueri, “Creación y gestión de una base de datos con MySQL y phpMyAdmin”, Universidad de Jaen. 2018. Disponible en: https://hdl.handle.net/10953.1/9445 (Últ. acceso: 15/06/2023).

[18] A. D. Espinoza Altamirano, “Desarrollo de un prototipo de Iot basado en el módulo wifi y servidor web local en la plataforma raspberry-pi enfocado para personas con capacidades especiales de movilidad”, Escuela Politécnica Nacional de Quito (Ecuador). 2021. Disponible en: http://bibdigital.epn.edu.ec/handle/15000/21774 (Últ. acceso: 15/06/2023).

[19] E. M. S. Cruz F. A. S. Cruz, C.R. Huerta, F. G. C. Julián. Monitoreo “TCP/IP En Labview© De Una Señal Ultrasónica Mediante Un Sistema IoT, Utilizando El Sistema Arduino En Modo Servidor Con El Shield De Ethernet”, Pistas Educativas 2018, vo. 40, no.130.

[20] A. H. Quintana-Cruz, “Implementación de aplicaciones isomórficas con Javascript”. Interfases 2015, vol. 8, pp.143-161.

[21] R. Carreño Villalba, (2016). Aplicación de administración con Angular, Node y Express para una aplicación Django. Universidad de Málaga. Disponible en: http://hdl.handle.net/10630/11484 (Últ. acceso: 15/06/2023).

[22] P. P. Ray, “A survey on Internet of Things architectures”. Journal of King Saud University-Computer and Information Sciences 2018, vol. 30, no. 3, pp. 291-319.

[23] Protocolo MySQL. Disponible en: https://dev.mysql.com/doc/dev/mysql-server/latest/PAGE_PROTOCOL.html (Últ. acceso: 15/06/2023).

[24] N. Bahbouh, A. Basahel, S. Sendra, A. Sen, A. Ahmed, “Tokens Shuffling Approach for Privacy, Security, and Reliability in IoHT under a Pandemic. Applied Sciences 2023, vol, 13, no.1, pp.114.

[25] Caracteristicas de la Base de datos MariaDB. Disponible en: https://mariadb.org/ (Últ. acceso: 15/06/2023).

[26] Caracteristicas del módulo Huzzah ESP32. Disponhttps://www.adafruit.com/product/3405 (Últ. acceso: 15/06/2023).

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Exploring Cybernetic Solutions for Health Monitoring: Integrating RPL Network Protocol and Intrabody Communications

Anna Gutiérrez Melenchon, Alan Briones Delgado, Agustín Zaballos
Smart Society Research Group – Blended Network Architectures (BNA) Research Line
La Salle - Ramon Llull University
08022
anna.gutierrez@students.salle.url.edu , alan.briones@salle.url.edu, agustin.zaballos@salle.url.edu

Cybernetic solutions have revolutionized health monitoring by incorporating principles that explore feedback, control, and communication in complex systems. These solutions enhance patient care, diagnostics, and treatment outcomes by seamlessly integrating sensors, devices, and data analysis techniques. This paper explores the potential of cybernetics by combining the RPL (Routing Protocol for Low Power and Lossy Networks) network protocol for efficient communication and data transmission with Intrabody Communications, specifically Galvanic Coupling. The aim is to develop a holistic architecture for healthcare applications, revolutionizing health monitoring with network sensors embedded in the human body that collect data from multiple health conditions. Through extensive experimentation, we seek to validate the feasibility and effectiveness of this integrated approach, with the goal of improving healthcare applications through real-time analysis, early detection of complications, and personalized treatment plans.

Keywords **- health monitoring, communication architecture, sensor integration, real-time monitoring, patient care, medical research.**

I. State of Art

In this section, we delve into the state of the art, beginning with an exploration of cybernetics. This field, rooted in the study of communication and control in natural and artificial systems, has profoundly influenced various disciplines. It serves as a foundational element in our quest to understand and integrate complex systems into healthcare and other domains.

A. Cybernetics

Cybernetics is an interdisciplinary field that focuses on the study of communication and control in animals and machines. It originated from the Greek word "kybernetes," meaning "steersman" or "governor." Cybernetics has had a significant impact on various disciplines, including biology, engineering, computer science, and psychology.

The history of cybernetics can be divided into three main periods: the classical era, the modern era, and the postmodern era. The classical era began in the 1940s with the work of Norbert Wiener, who proposed that the principles of communication and control observed in living systems can be applied to the design of machines [1]. One of the key concepts introduced by Wiener was the "feedback loop," which is essential for the functioning of both living and non-living systems.

In the modern era, cybernetics saw significant advancements with the development of computers and the internet. This led to the emergence of artificial intelligence (AI), which has been applied in various fields, including language translation, image recognition, and self-driving cars.

The postmodern era of cybernetics started in the 21st century with the proliferation of digital technologies like smartphones and the Internet of Things (IoT). These technologies have brought about new forms of communication and control, raising ethical and social implications [2].

B. Healthcare Applications

Cybernetics has made significant contributions to healthcare through various applications. One of these is the use of robots in surgery, known as robotic surgery. Robotic systems allow for precise and minimally invasive procedures, resulting in less pain, faster recovery, and fewer complications [3].

Wearable devices, such as fitness trackers and smartwatches, have also utilized cybernetic principles in healthcare. These devices monitor vital signs and activities, providing real-time feedback to patients and facilitating remote monitoring by healthcare providers [4].

Telemedicine, the use of technology for remote medical care, has greatly benefited from cybernetics. It enables diagnosis and treatment for patients in remote or underserved areas, as well as post-hospital care. Videoconferencing, email, and phone calls are common telemedicine methods [5].

Cybernetics has contributed to the development of medical devices like pacemakers and insulin pumps, which use feedback loops to adjust their function based on patient needs. Brain-Computer Interfaces (BCIs) have also been developed, allowing individuals to control devices using their thoughts [6].

While these technologies have numerous benefits, they raise ethical and social concerns. Surgeon roles in robotic surgery, privacy issues with wearable devices and telemedicine, increased healthcare costs, and disparities in access to care are among the considerations that need attention [7].

C. Intrabody Area Networks

Intrabody Area Networks (IBANs) are networks of devices embedded within the human body that facilitate communication and data exchange. These networks have gained significant attention due to their potential applications in various fields, particularly in the medical domain [8].

One of the key technologies employed in IBANs is Ultra WideBand (UWB) communication. UWB offers several advantages, including low power consumption, high data rates, and the ability to operate effectively in noisy environments. This makes it an ideal choice for establishing communication links between devices within the body. The efficient utilization of UWB technology ensures reliable and efficient transmission of data in IBANs [8].

To regulate device communication in IBANs, the Intrabody Communication (IBC) protocol is employed. This protocol governs the rules and procedures for data exchange between devices within the network. By standardizing communication protocols, IBC ensures interoperability and seamless integration of devices in an IBAN [9].

One prominent application of IBANs is in the field of medical monitoring. Devices embedded within the body can collect various health-related data, such as heart rate, blood pressure, and glucose levels. This data is transmitted wirelessly to central monitoring devices in real-time, allowing healthcare professionals to monitor patients' health remotely and take appropriate actions when necessary [8].

IBANs also hold promise in the field of drug delivery. By incorporating drug delivery devices within the body, medications can be released precisely and targeted to specific areas. This targeted drug delivery approach has the potential to revolutionize the field of medicine by enhancing treatment efficacy and minimizing side effects [9].

Moreover, IBANs can play a significant role in rehabilitation. By utilizing devices that stimulate muscles or nerves, IBANs can aid in the rehabilitation process by improving motor functions and facilitating movement in patients. This technology has the potential to significantly enhance the effectiveness of rehabilitation programs and improve patients' quality of life [8].

Different radio frequency methods are utilized in IBANs, including narrowband (NB), UWB, and millimeter wave (mmWave) channels. Each method has its own advantages and trade-offs, making it suitable for specific scenarios and requirements. Additionally, ultrasound is another method employed in IBANs, offering advantages such as power efficiency and deeper tissue penetration [10].

In terms of intrabody communication, two primary methods are commonly used: capacitive coupling and galvanic coupling. Capacitive coupling involves the transmission of data through the variation of electric fields, while galvanic coupling relies on the use of electrical currents. Both methods have their own advantages and challenges, but galvanic coupling shows promise in terms of power consumption, tissue safety, security, and transceiver complexity [10] [11].

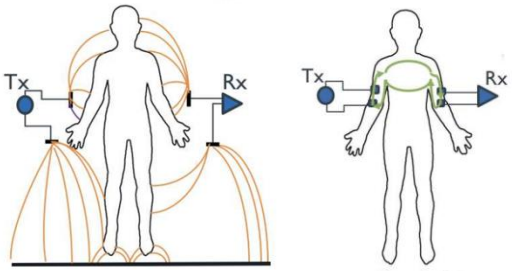


Fig. 1. Capacitive Coupling and Galvanic Coupling [11]

Overall, IBANs have immense potential in revolutionizing healthcare and other fields by enabling seamless communication and data exchange within the human body. It is necessary to consider the limitations of the hardware for the implementation of the solution and the effect of minituarization of the components, which can affect the communication. With further advancements in technology and research, IBANs are likely to play an increasingly significant role in improving medical treatments, rehabilitation processes, and overall well-being [11].

The paper is structured as follows. The Section 2 presents the background related to intra-body communications. Section 3 describes the Use Case, the deployment and results of the Proof of Concept (PoC) for the RPL deployment in an intra-body simulated environment. Finally, Section 4 highlights the future work related to this simulated PoC and the Section 5 provides the findings and conclusions..

II. BACKGROUND

This section provides essential background knowledge. We begin with Galvanic Coupling, a technique crucial for intra-body communication, highlighting its non-invasiveness and portability. We also address tissue safety concerns and the complexities of optimizing intra-body

network topology, especially in Galvanic Coupling Intra-Body Networks (GC-IBNs).

Next, we delve into the Routing Protocol for Low Power and Lossy Networks (RPL). RPL is designed for Low Power and Lossy Networks (LLNs), addressing challenges like energy efficiency and packet loss. We uncover its features, including support for various network topologies, auto-configuration, self-healing, and more. RPL's use of Directed Acyclic Graphs (DAGs) and unique identifiers ensures efficient routing in LLNs, making it vital for our exploration of advanced health monitoring solutions.

A. Galvanic Coupling

Galvanic coupling is a form of electrical coupling that enables the transmission of electrical signals between conductive materials in close proximity. It is particularly useful in intra-body communication, where it allows for the exchange of information between implanted or placed devices inside the human body. This technique relies on electromagnetic induction, where an induced current flows in the second conductor due to the electromagnetic field created by the first conductor. Factors such as distance, orientation, and conductivity influence the strength of the induced current. Galvanic coupling offers advantages such as non-invasiveness and portability, eliminating the need for external wires or cables in body-worn or implanted devices [12].

In transmitting signals through human tissues, ensuring tissue safety is of utmost importance. High-frequency and high-power signal transmission can lead to elevated body temperatures, which has been extensively studied in medical procedures such as diathermy, RF ablation, MRI, and ultrasound. However, the effects of low-power galvanic coupling with multiple concurrent transmitters on tissue heating remain less explored. Strategies to control medium access and avoid tissue heating in such scenarios are needed [13].

Optimizing the topology of intra-body networks presents challenges, as the placement of relay nodes for dynamic clustering is a computationally difficult problem. The constraints of tissue safety and the limitation of transmission power levels set by regulatory guidelines further complicate the clustering process. In addition, the unique characteristics of Galvanic Coupling Intra-Body Networks (GC-IBNs), such as bidirectional traffic and non-redundant implants, require specialized design considerations [14].

Accurately modeling the tissue channel is crucial for understanding the propagation of signals within an intra-body network. Existing in-vivo tissue experiments and commercial phantoms do not fully capture the heterogeneity and electrical propagation characteristics of human tissue. Therefore, developing an analytical model that accurately represents the communication channels along different tissue paths is essential. The use of equivalent circuit models and impedance measurements aids in studying the channel responses and enables the assessment of wireless communication through the body using galvanic coupling [15].

Furthermore, an example of a 2-Port Model demonstrates the feasibility of galvanic coupling for intra-body communication. Through experimentation, the model evaluates parameters such as bit error rate, gain, and the maximum number of nodes, providing insights into the performance and potential of galvanic coupling as an effective communication method within the body [16].

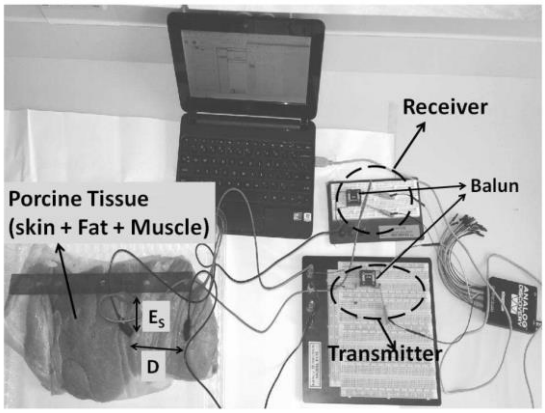


Fig. 2. Galvanic Coupling Communication through Porcine Tissue [16]

Building upon the successful demonstration of galvanic coupling, the exploration of the Routing Protocol for Low-Power and Lossy Networks (RPL) becomes relevant. RPL provides efficient and reliable communications among the sensors embedded in the human body. It offers features such as communication optimization, power consumption minimization, and adaptability to dynamic network conditions. Understanding RPL's design principles and application enhances its significance in healthcare monitoring and personal area network applications.

B. Routing Protocol for Low Power and Lossy Networks (RPL)

RPL (IPv6 Routing Protocol for Low-Power and Lossy Networks) is a routing protocol specifically designed for Low Power and Lossy Networks (LLNs). It addresses the challenges of these networks by adapting the network behavior to parameters such as convergence time, energy, packet loss, and packet delay. RPL achieves this by providing robustness to changes and offering alternate routes when default routes become inaccessible [17].

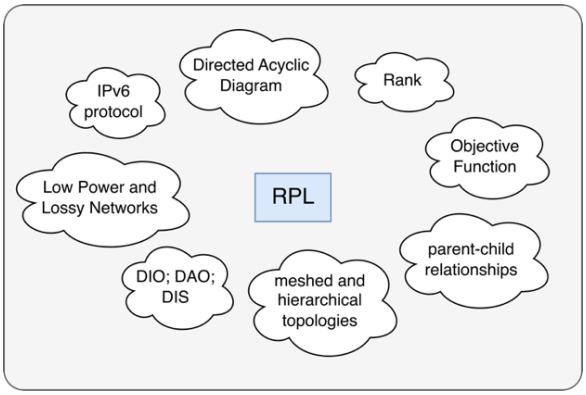


Fig. 3. RPL Basics

The design of RPL takes into consideration two main characteristics of LLNs: low data-rate (less than 250 kbps) and high error rates. These factors result in a low data throughput and the need for a highly adaptive routing protocol. RPL uses the concept of Directed Acyclic Graphs (DAGs) to organize the network topology. Each node in the network is connected to its parent through a directed link, forming a tree-like structure. The root node acts as the gateway to other non RPL networks [17].

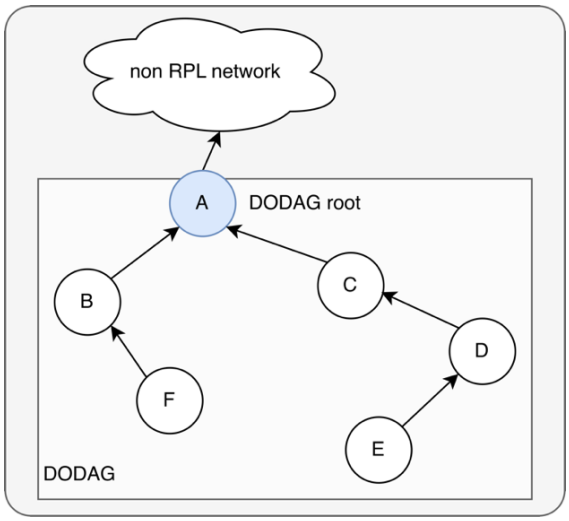


Fig. 4. RPL Structure

One of the key features of RPL is its support for both meshed and hierarchical topologies. It allows routing through siblings, providing flexibility in routing and topology management. RPL also introduces several terminologies to describe its operation, such as DAG, DAG root, DODAG (Destination-Oriented DAG), DODAG root, and more [17].

The protocol aims to achieve specific objectives, including auto-configuration, self-healing, loop avoidance and detection, independence and transparency, and support for multiple edge routers. RPL is designed to operate on top of multiple link layer mechanisms, such as IEEE 802.15.4 PHY and MAC layers, without relying on any specific link-layer features [17].

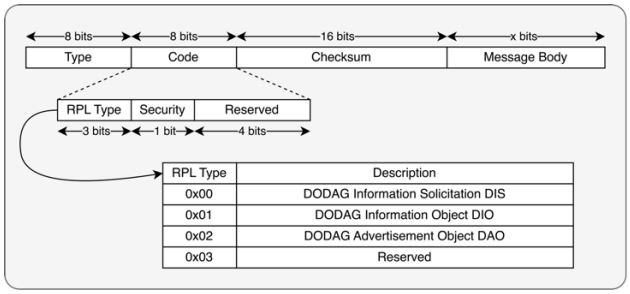


Fig. 5. RPL Control Message

RPL operates by creating a DODAG rooted at a specific node, the DODAG root. Each node in the DODAG is assigned a rank, indicating its position relative to others with respect to the DODAG root. The construction process involves nodes selecting their preferred parent based on rank calculation. Nodes periodically transmit their rank and update their parent if a better route is available [17].

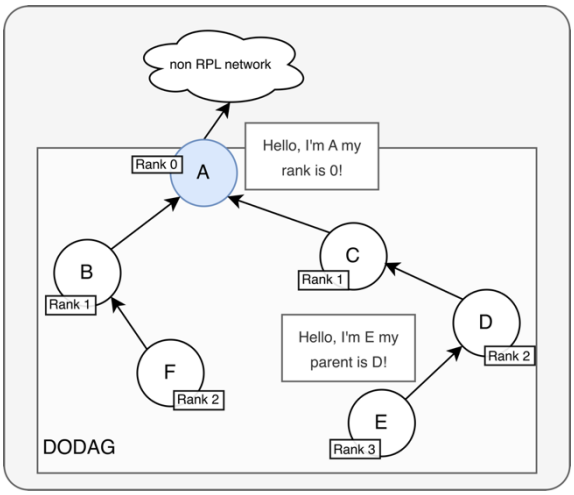


Fig. 6. DODAG Creation

The topology information is stored at the DODAG root, which maintains the parent-child relationships. RPL uses identifiers such as RPLInstanceID, DODAGID, and DODAG Version to uniquely identify and manage the topology. The protocol ensures efficient routing and adaptation to changes in the LLN environment [17].
Overall, RPL is a robust and adaptive routing protocol designed to meet the challenges of Low Power and Lossy Networks. It provides features like auto-configuration, self-healing, loop avoidance, and supports both meshed and hierarchical topologies. By leveraging Directed Acyclic Graphs, RPL enables efficient routing and ensures network connectivity in LLN environments [18].

III. Use Case: Protocol Design

Advancements in healthcare technology have led to the integration of multiple sensors in the human body for real-time health monitoring. This study focuses on designing a communication architecture that enables seamless data transmission among these sensors.

The proposed model involves strategically placed sensors collecting vital signs data, which is transmitted to a central node. This architecture has applications in real-time patient monitoring and medical research, enabling prompt decision-making and insights into various health conditions. By harnessing technology and data, this approach has the potential to revolutionize healthcare and improve patient care.

A. Media Access Technology

The research indicates that Galvanic Coupling (GC) is the most suitable protocol for sensor communication. GC offers several advantages over other peer technologies:

- **Low Tissue Absorption**: GC's low-frequency signal is safe and minimally affects tissue temperature compared to RF signals.
- **Energy Efficiency**: GC-IBN provides long battery life, crucial for wearable devices.
- **Coexistence with Other Wireless Applications**: GC can operate alongside RF technologies without interference.
- **Short Range and Secured Transmission**: GC links offer short-range, secured transmission confined to the body.
- **Acceptable Bandwidth**: GC ensures sufficient bandwidth for real-time monitoring applications.
- **Cost-Effective Devices and Network Deployments**: GC requires simple protocols and circuitry, enabling cost-effective deployments.
- **Continuous Unobtrusive Real-Time Connectivity**: GC allows patients to lead normal lives while being monitored.

In summary, GC advantages make it an attractive option for communication between sensors embedded in the human body for health monitoring.

B. Requirements for the Network Architecture

A network protocol is crucial for the communication infrastructure in a sensor network monitoring body health conditions using Galvanic Coupling. Key requirements for the protocol include:

- **Routing**: The protocol should support efficient data transmission and reduce power consumption through effective routing.
- **Small Distance between Sensors**: Given the limited distance of Galvanic Coupling, the protocol must enable communication among sensors within a short range.
- **Sensor Location**: The protocol should address sensor identification and addressing, such as assigning unique IP addresses, to ensure proper data transmission and flexibility for adding or removing sensors.
- **Autodiscovery of Nodes**: The protocol should automatically discover and integrate new nodes, facilitating sensor additions or updates.
- **Structure**: The network protocol's structure should be flexible and scalable, allowing easy node addition or removal, potentially utilizing a cluster mode with designated leaders.
- **Battery**: Battery efficiency is crucial, and the protocol should minimize power consumption, adapt to varying battery capacities, and maximize battery life.

These requirements ensure reliable and efficient monitoring of body health conditions using Galvanic Coupling and contribute to the development of an effective network architecture.

C. Network Architecture

The proposed model for the network involves the use of multiple sensors placed around the human body to collect data related to human constants such as heart rate, blood pressure, and temperature. The collected data is transmitted through communication links to a central node for processing and analysis. The model utilizes the RPL routing protocol, designed for low-power and lossy networks, to manage the communication links and establish efficient routes based on network topology. The use of RPL enables dynamic adaptation to changes in the network and supports energy-efficient routing. The proposed model is scalable, energy-efficient, and reliable, making it suitable for healthcare settings. It has the potential to improve patient outcomes by providing real-time data and automated responses to alarms. The network structure consists of clusters of sensors organized in a neuron architecture, inspired by the biological neuron structure. The central node acts as a border router, facilitating communication between the RPL network and the Internet. Overall, the proposed model enhances health monitoring, leading to improved patient outcomes and reduced healthcare costs.

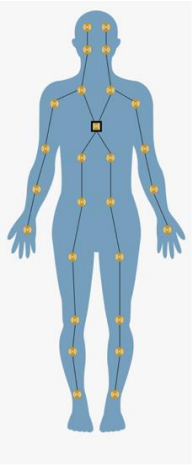


Fig. 7. Sensor's Network Neuron Architecture

D. Prove Of Concept

The study proposes a simulation model for an RPL network in Contiki OS using the COOJA emulator to simulate communication between nodes placed in the human body. COOJA is a powerful tool for testing and debugging wireless sensor network applications. The simulation aims to analyze the performance of the RPL protocol in terms of packet delivery ratio, energy consumption, and end-to-end delay. The simulation parameters, such as the radio medium, mote type, transmission ratio, transmission range, interference range, simulation period, squared area, and topology, are detailed. These parameters impact the performance of the sensor network application. The COOJA simulator, along with a step-by-step manual provided in the study, facilitates the development and testing of WSN applications, reducing time and cost. A validation test is proposed to ensure that the network protocol meets the specified requirements, addressing aspects like efficient routing, problem-solving for small distances, sensor identification and addressing,

automatic discovery and addition of nodes, flexibility and scalability of the protocol structure, and power consumption minimization. The validation test aims to verify the reliability and efficiency of monitoring body health conditions using the RPL Galvanic Coupling method as the physical communication method. The test questions include:

1. Can the network protocol efficiently route data to the necessary nodes?
2. Does the network protocol solve the problem of small distances between sensors?
3. Can the network protocol properly identify and address each sensor?
4. Does the network protocol have the ability to automatically discover and add new nodes to the network?
5. Is the network protocol structure flexible and scalable for easy addition or removal of nodes?
6. Does the network protocol minimize power consumption and maximize battery life?

E. Implemented Scenario

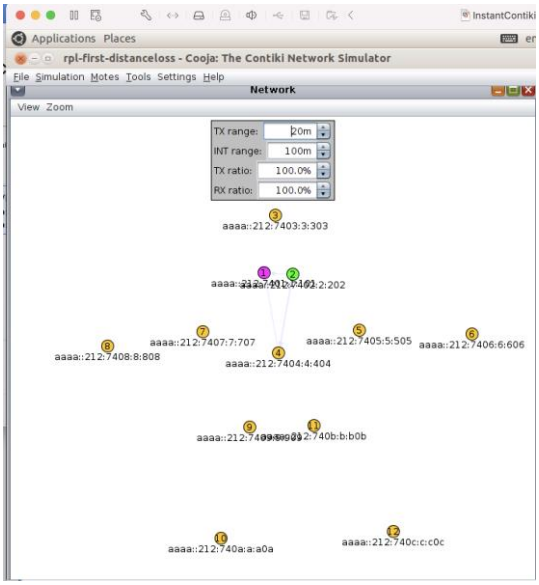


Fig. 8. Contiki Network Simulation

Fig. 8 illustrates the configuration of the simulation, where node 1 is the border router, node 2 is the sink, and the other 10 nodes are sender nodes. Once the simulation starts, nodes engage in data transmission and reception. Each node selects its best parent based on distance-loss and establishes parent-child relationships. With all communications established, all nodes can communicate with each other. The simulator includes a sensor data collector that evaluates node communication and generates graphics and tables for validation. These include a sensor's map depicting parent-child relationships and calculated distances, a table with node information such as hops to the central node, CPU power, ETX (Expected Transmission Count), and received data, and a table displaying average power consumption for different components.

Communication with the outside is validated by establishing communication from the PC to the border router and individual nodes. The border router publishes its routing table on a website, allowing for consultation of discovered and accessible nodes.

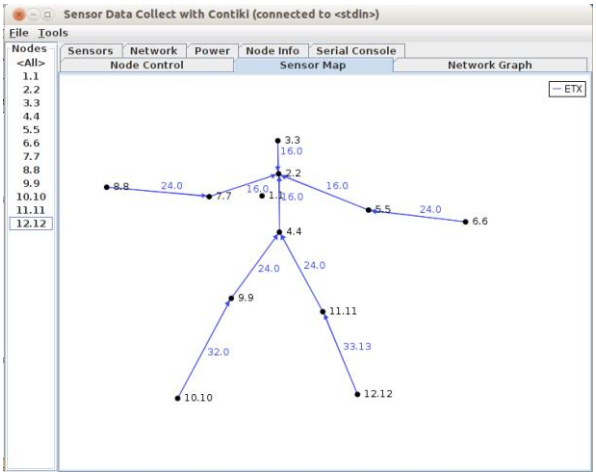


Fig. 9. Sensors Map

F. Validation Test Results

The validation tests conducted on the network protocol for a sensor network used in monitoring body health conditions yielded the following insights, answering to the six questions presented:

1. **Efficient routing**: The network protocol includes routing capabilities that enable efficient transmission of data to the necessary nodes. This ensures effective communication among all nodes in the network given the creation of a hierarchical structure, ensuring the path to the gateway node.
2. **Overcoming small distances**: The network protocol effectively addresses the challenge of small distances between sensors. It ensures optimal performance for communication by utilizing routing capabilities and forwarding information to the next-hop node keeping the hierarchy stablished. This is particularly important when using Galvanic Coupling as the physical communication method.
3. **Sensor identification**: The network protocol provides a reliable method for identifying and addressing each sensor by assigning a unique ID to each node of the network. This enables accurate communication and data exchange.
4. **Automatic node discovery and integration**: The network protocol has the capability to automatically discover and add new nodes to the network. Nodes constantly monitor received data, allowing them to detect and establish connections with newly appearing nodes. This feature is crucial for adding new sensors or updating existing ones without manual intervention, reconfiguring it.

5. **Flexible and scalable structure**: The network protocol is designed with a flexible and scalable structure, facilitating easy addition or removal of nodes. Nodes actively monitor the status of their connected nodes and adapt the network structure accordingly. This ensures efficient communication and allows the network to adapt to changing conditions.
6. **Power consumption and battery life optimization**: The simulation did not specifically consider battery life as a parameter for generating parent-child relationships. Therefore, it remains uncertain whether the network protocol effectively minimizes power consumption and maximizes battery life. Further research and optimization are required in this aspect to enhance the energy efficiency of the protocol.

In summary, although the simulation outcomes offer valuable insights, it is crucial to tackle these limitations to guarantee the real-world reliability and precision of the network protocol.

G. Discussion and Limitations

The discussion focuses on the results of the validation tests conducted on the network protocol design for monitoring body health conditions. The tests confirm that the protocol efficiently routes data to the necessary nodes, solves the problem of small distances between sensors, properly identifies and addresses each sensor, and has a flexible and scalable structure. However, the protocol lacks the ability to adapt to battery capacity, which is an area for future improvement.

Regarding limitations, firstly, the simulation environment does not fully replicate the physical layer of the sensor network, as it uses a wireless channel instead of the Galvanic Coupling method employed in real-world scenarios. Minituarization of the sensors and the implementation of the protocol must be considered. This may affect the network's behavior and performance when deployed in practical settings. Additionally, the simulation does not account for all the effects of noise and interference, and it assumes a static network topology, disregarding factors such as mobility. Future studies should address these limitations and incorporate them into the research.

Furthermore, the impact of battery life on the network's performance is not considered in the simulation. This oversight can lead to suboptimal performance, particularly for devices with limited battery capacity. Incorporating battery life as a factor in the network protocol design should be a focus of future research. Overall, while the simulation results provide valuable insights, addressing these limitations is essential to ensure the reliability and accuracy of the network protocol in real-world applications.

IV. Future Work

The future work chapter presents potential research directions and areas for improvement in the studied system. It serves as a roadmap for researchers and practitioners to explore new avenues for advancing the field.

One area of future work is the development of new sensors to enhance the monitoring capabilities of the network. This includes sensors for measuring biomarkers and environmental factors. Improving the efficiency and accuracy of existing sensors through advanced signal processing techniques, such as machine learning, is also important.

Integration with healthcare systems is another avenue for enhancement. This involves integrating the collected data with electronic medical records, telemedicine platforms, healthcare analytics, and decision support systems. These integrations can improve patient care and provide valuable insights to healthcare providers.

The incorporation of machine learning can significantly enhance the network's performance. Using machine learning algorithms, patterns in sensor data can be identified for accurate prediction of changes in the body's condition. Machine learning can also optimize routing algorithms and improve overall network efficiency.

Optimizing the RPL network protocol is crucial for reliable and efficient communication. Techniques such as objective functions, multi-parent routing, traffic engineering, and security mechanisms can be applied to enhance performance, reliability, and security of the network.

Cybersecurity is a critical consideration for the network. Implementing authentication, encryption, access control, firewalling, network segmentation, and regular monitoring and updates are essential to ensure the security of sensitive data transmitted within the network.

In summary, the future work chapter highlights the potential for advancements in sensor development, integration with healthcare systems, machine learning applications, RPL network optimization, and cybersecurity measures. These areas offer opportunities for further research, innovation, and improvement in the field.

V. Conclusions

The proposed research on intrabody communication and the RPL network protocol for healthcare and personal area network applications holds immense importance and offers potential benefits in various fields. By applying principles of cybernetics to healthcare, researchers aim to enhance patient care, improve diagnostics, and optimize treatment outcomes.

The architecture provides a viable and cost-effective solution for healthcare monitoring, leading to improved patient outcomes and a better quality of life. Through real-time monitoring and analysis of various health parameters using multiple sensors, early detection of health complications becomes possible, enabling timely intervention and proactive healthcare management. This

can result in reduced hospitalization rates and enhanced overall quality of life for patients.

One key aspect of the research is the exploration of intrabody communication, particularly galvanic coupling. This method utilizes the human body's conductive properties to establish secure and reliable data transmission between implanted or wearable devices. Galvanic coupling has shown promise in enabling seamless data exchange between sensors, offering low power requirements and the ability to overcome signal interference. Its potential to revolutionize healthcare monitoring and improve patient outcomes has been demonstrated through extensive experimentation and analysis.

Additionally, the use of the RPL network protocol enables efficient communication and data transmission among sensors, reducing the reliance on external infrastructure and minimizing power consumption. This scalability and flexibility of the network protocol allow for easy integration and removal of nodes without disrupting the overall network structure. Such advantages lead to significant cost savings and increased convenience for users.

Moreover, the development of new sensors and their integration with healthcare systems hold potential for further advancements in the field. Machine learning algorithms can aid in the analysis and interpretation of large amounts of data, providing better insights and predictive models. This can lead to early disease detection and more personalized treatment plans for patients.

In conclusion, the proposed research has far-reaching implications for healthcare and personal area network applications. It has the potential to improve patient outcomes, reduce healthcare costs, and advance medical research. The simulation results validate the viability of the proposed architecture, and further testing in simulated body channel conditions can provide additional insights and improvements. However, it is important to note that the current simulation does not fully simulate the physical layer using Galvanic Coupling, which is a limitation. Addressing this limitation and gaining a comprehensive understanding of the proposed architecture's performance in real-world scenarios requires further research.

ACKNOWLEDGEMENTS

This research has also received founding from the framework of the Advanced Training in Health Innovation Knowledge Alliance (ATHIKA) project, funded by the European Commission Erasmus+ Programme – KA2 cooperation for innovation and the exchange of good practices – Knowledge Alliances (601106-EPP-1-2018-1-ES-EPPKA2-KA).

REFERENCES

[1] Wiener, N. (1948). Cybernetics: Or Control and Communication in the Animal and the Machine. MIT Press.

[2] Hayles, N. K. (1999). How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. University of Chicago Press.

[3] Rovira, E., & Lambrecht, J. T. (2016). Robotic Surgery: Current Applications and Future Directions in Urology. Archives of Esp-Urologia, 69(5), 220-228.

[4] Bonato, P. (2019). Wearable Sensors and Systems. In Textbook of Neural Repair and Rehabilitation (Vol. 2, pp. 364-380). Cambridge University Press.

[5] Dorsey, E. R., Topol, E. J., & Telemedicine Study Group. (2016). State of Telehealth. New England Journal of Medicine, 375(14), 1400-1401.

[6] Duffy, V. G. (2018). Designing with Care: Technological Mediation in Healthcare. MIT Press.

[7] Sharkey, N. (2019). Should Robots Be Used in Surgery?: A Conversation with Dr. Peter Kim. Science and Engineering Ethics, 25(4), 1189-1203.

[8] Karedal, J., & Tufvesson, F. (2017). Ultra-Wideband for Wireless Body Area Networks. In Wireless Body Area Networks: Technology, Implementation, and Applications (pp. 41-68). CRC Press.

[9] Latre, B., Braem, B., Moerman, I., Blondia, C., & Demeester, P. (2009). A Survey on Wireless Body Area Networks. Wireless Networks, 17(1), 1-18.

[10] Yu, L., Zheng, G., Zhang, Y., Wang, C., Liu, J., & Huang, H. (2021). Ultra-Wideband Intra-Body Communication Channel Modeling and Its Application in Body Area Networks. IEEE Access, 9, 130181-130196.

[11] Tomlinson, W. J., Banou, S., Yu, C., Stojanovic, M., & Chowdhury, K. R. (2018). Comprehensive survey of galvanic coupling and alternative intra-body communication technologies. IEEE Communications Surveys & Tutorials, 21(2), 1145-1164.

[12] Kumar, S., & Tripathi, S. (2020). Intra-Body Communication Techniques: A Comprehensive Review. In 2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3) (pp. 1-6). IEEE.

[13] Zahedi, M., & Zheng, Y. R. (2019). Non-Invasive Intra-Body Communication for Implantable Biomedical Devices: Modeling, Simulation, and Experimental Investigation. IEEE Access, 7, 180092-180106.

[14] Kibret, B., & Lai, D. T. H. (2015). Galvanic Coupling Intra-Body Sensor Networks: A Comprehensive Survey. IEEE Communications Surveys & Tutorials, 17(3), 1537-1559.

[15] Behbahani, F. M., Behbahani, M. F., & Behbahani, S. M. R. (2018). A Novel Analytical Model for Galvanic Coupling Intra-Body Communication Channels Based on Conformal Mapping. IEEE Sensors Journal, 18(1), 274-281.

[16] Swaminathan, M. (2017). Wireless Intra-Body Communication for Implantable and Wearable Body Devices using Galvanic Coupling (Doctoral dissertation, Northeastern University).

[17] Paragraph 6 and 7: Winter, T., Thubert, P., Brandt, A., Hui, J., Kelsey, R., Levis, P., ... & Struik, R. (2012). RPL: IPv6 routing protocol for low-power and lossy networks. Request for Comments (RFC) 6550.

[18] Paragraph 8: Palattella, M. R., Accettura, N., Vilajosana, X., Watteyne, T., Grieco, L. A., Boggia, G., & Dohler, M. (2013). Standardized protocol stack for the Internet of (important) Things. IEEE Communications Surveys & Tutorials, 15(3), 1389-1406.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Eco4rupa: buscando tu ruta

S. Felici-Castell*, JJ. Perez-Solano*, J. Segura-Garcia*, A. Soriano-Asensi*, R. Fayos,†, J. Lopez-Ballester*, G. Mas‡

* ETSE, Universitat de València,† UWS (UK),‡ BSG Ingenieros

felici@uv.es, jjperezs@uv.es, jsegura@uv.es, soan@uv.es

rafael.fayos@uws.ac.uk, jesus.lopez-ballester@uv.es, gemma@bsg.es

Resumen

Los sensores de bajo coste para calidad del aire nos ofrecen la oportunidad de integrarlos en redes inalámbricas con el objetivo de permitir una vigilancia en tiempo real y mayor densidad espacial de muestreo de la contaminación en las ciudades. Teniendo en cuenta que gran parte de la población vive en estas y el aumento de los problemas respiratorios/alérgicos de la ciudadanía, resulta de gran interés ofrecer servicios y aplicaciones para mejorar su calidad de vida. Estos planteamientos se llevan a cabo dentro del proyecto *Eco4rupa* y *Greenish*, donde una red de sensores de bajo coste, asistida con redes oficiales de monitorización de calidad del aire, con el apoyo de comunicaciones 5G y a través de técnicas de inteligencia artificial e interpolación espacial, permiten planificar rutas *adhoc* acorde con el perfil y necesidades del ciudadano. En los resultados se presenta dicho planificador que permite reducir en media la exposición a la contaminación en un 17.82 %, a cambio de un aumento de la distancia recorrida en un 9.8 %.

***Palabras Clave*—redes dedicadas, gases contaminantes, planificador de rutas, vigilancia ambiental**

I. INTRODUCCIÓN

Los ciudadanos se enfrentan constantemente a niveles de contaminación del aire que violan los umbrales de seguridad para la salud humana definidos por la Organización Mundial de la Salud (OMS) [1]. Según Eurostat [2], 441.831 residentes de la UE fallecieron solo en 2015 debido a enfermedades respiratorias en la UE-28. Además, el problema se complica si consideramos que gran parte de la población presenta o puede presentar algún tipo de alergia, problemas respiratorios y cutáneos [3].

En este escenario, las redes inalámbricas de sensores (en inglés *Wireless Sensor Networks*, WSN) para monitorización basadas en Internet de las Cosas (*Internet of Things*, IoT) con soporte de las tecnologías 5G, junto con técnicas de Inteligencia Artificial (IA), pueden ayudar al ciudadano en su día a día a través de un sistema que vele por la salud en sus desplazamientos, especialmente cuando presenta problemas respiratorios y/o alérgicos.

En este artículo se presentan las líneas de trabajo que se están llevando a cabo dentro del proyecto *Eco4rupa* y *Greenish*, donde se pretende abordar esta necesidad con un planificación de rutas, atendiendo a los perfiles de usuario (que recogen las patologías e historial clínico) asistido con una red de monitorización en tiempo real de la calidad del aire. Además, esta red de monitorización se complementa con los datos públicos disponibles de las estaciones de vigilancia de gases contaminantes (por ejemplo, la red de estaciones de la Generalitat Valenciana [4]), junto con técnicas estadísticas de inferencia espacial.

El resto del trabajo, en la Sección II, mostramos sensores disponibles para calidad del aire y trabajos relacionados con nuestra propuesta. En la Sección III, analizamos las diferentes alternativas de diseño a utilizar en la red de monitorización. En la Sección IV, discutimos las opciones para poder integrar y fusionar la información obtenida con el planificador de rutas junto con la información geográfica. Finalmente, en la Sección V, resumimos las principales conclusiones.

II. ESTADO DEL ARTE

El reciente auge de los sensores de calidad del aire de bajo coste, debido a la facilidad de instalación y bajo consumo, hace que se utilicen cada vez más y son interesantes para integrarlos en WSN. Estos sensores pueden medir contaminantes atmosféricos como NO, NO2, SO2, CO, CO2, O3, así como compuestos orgánicos volátiles (Volatile Organic Compounds, VOC, normalmente medidos en totales, TVOC), metales pesados (Pb, As, Cd) y partículas en suspensión en el aire (Particulate Matter, PM), además de Temperatura (T), Presión Atmosférica (PA) y Humedad Relativa (HR).

Tabla I
COMPARATIVA DE SENSORES PARA CALIDAD DEL AIRE

Módulo	Gases y medidas	Conexión
SDS011 [5]	PM,T,HR,PA	UART
DL-LP8P [6]	CO2,T,HR,PA	LoRAWAN
MiCS-6814 [7]	CO,NO2,C2H5OH, NH3, CH4	I2C,SPI
ZPHS01B [8]	PM,CO,O3,NO2,TVOC,T,HR	UART

Los fabricantes también integran varios de estos sensores en el mismo módulo lo que facilita su uso. En la

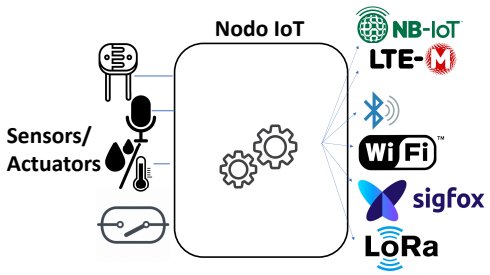


Figura 1. Esquema genérico de nodo IoT y comunicaciones.

Tabla I, se muestra una relación de este tipo de sensores (o modulo de sensores) y sus características principales. De todos ellos, el que hemos considerado con mejores prestaciones, mayor número de gases y mejor calidad/precio es ZPHS01B [8].

Sin embargo, en referencia a los márgenes y calidad de medida de los sensores de bajo coste, la norma CEN/TS 17660-1:2021 ha fijado los criterios establecidos por la Directiva 2008/50/CE para la equivalencia de los sistemas de sensores utilizados en exteriores con los instrumentos de mediciones indicativas y estimaciones objetivas. En este escenario, estos sensores presentan muchas limitaciones, por lo que no pueden utilizarse como sustituto de una referencia medida absoluta fiable [9], pero sí para tener un orden de magnitud y/o concienciación de la calidad del aire, ajustándose con técnicas de calibración basadas con algoritmos de IA [10]. Estos ajustes quedan fuera del ámbito del presente artículo.

Además en esta línea, existen iniciativas comerciales [11][12], aunque con un propósito limitado sólo a la monitorización de gases contaminantes, sin pretender cubrir y analizar la problemática planteada a nivel urbano en su conjunto. Hay también trabajos con un planteamiento similar al nuestro, como el realizado en el proyecto europeo Hope [13], que se centra en dar soluciones para la calidad del aire y ayuda al ciudadano en 2 distritos de Finlandia. Sin embargo en este proyecto, no se considera el perfil del usuario para poder analizar y medir la exposición o impacto en su salud en el cálculo de las rutas.

Aunque hemos dejado fuera gran parte del estado del arte por cuestiones de espacio, destacamos que el planteamiento propuesto de rutas saludables presenta un gran potencial e interés para la sociedad.

III. ALTERNATIVAS DE DISEÑO Y TÉCNICAS A UTILIZAR EN LA RED DE MONITORIZACIÓN Y SU ARQUITECTURA

En el citado proyecto, la red de monitorización está formada por nodos IoT basados en un microcontrolador al cual conectamos sensores de bajo coste, con posibilidad de diferentes alternativas de comunicación, como se muestra en la Figura 1. El objetivo de esta red, dado que por sus características tal como hemos comentado anteriormente

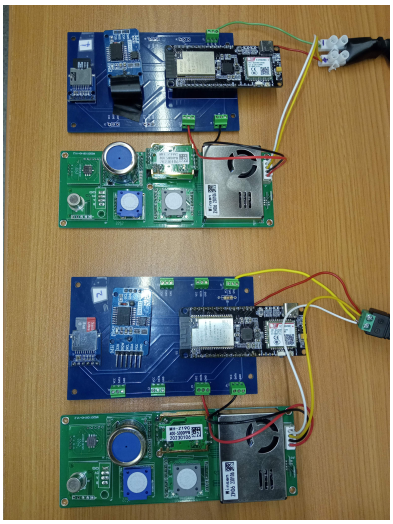


Figura 2. Nodos IoT con microntrolador ESP32 y sensores ZPHS01B.

Figura 3. Prototipo nodo IoT para monitorizar calidad del aire.

sólo ofrece valores orientativos de contaminación, es complementar la monitorización de la red de vigilancia oficial en aquellas zonas de pobre cobertura.

El nodo IoT está basado en el microcontrolador ESP32, por sus prestaciones y calidad/precio, ya que ofrece en cada modelo la posibilidad de tener diferentes antenas y/o externalizarla, así como implementar diferentes estándares de comunicación. Basado en este microcontrolador, hacemos una mención especial al módulo FiPy [14] de Pycom Ltd [1], que incluye tecnologías como Lora/Sigfox [2], WiFi, Bluetooth, y tecnologías celulares, como Long Term Evolution (LTE) for machines (LTE-M) y Narrow Band IoT (NB-IoT).

En la Figura 2, se muestra la conexión del microcontrolador ESP32 al módulo de sensores ZPHS01B [8]. Estos

[1]Pycom Ltd. entró en concurso de acreedores en septiembre de 2022, pero la recién creada Pycom BV se hizo cargo de esta empresa.

[2]En 2022, Sigfox entró en concurso de acreedores y la empresa Unabiz tomó su control en el momento de esta redacción.

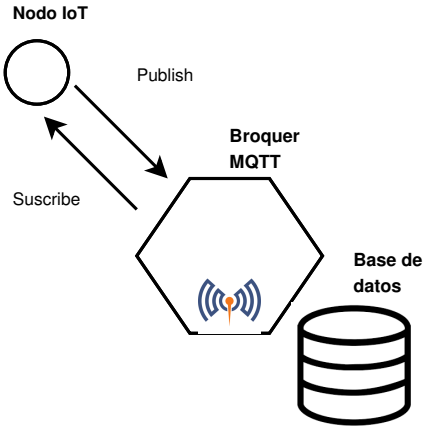


Figura 4. Conexión y protocolo de comunicaciones del nodo IoT.

nodos se instalan en exteriores como se muestra en la Figura 3, donde la entrada de aire se realiza por la parte inferior del tubo que succiona un pequeño ventilador.

El esquema de comunicaciones del nodo IoT con la infraestructura, se detalla en la Figura 4. Se basa en el protocolo IoT Message Queue Telemetry Transport (MQTT), que transmite información mediante mensajes entre los nodos y el broker MQTT. Destacar que MQTT permite 3 niveles de calidad de servicio (QoS) para verificar la entrega de los mensajes y también varios mecanismos de seguridad respecto a los datos transmitidos. Nosotros hemos elegido el nivel de calidad más alto, QoS-2, que garantiza la entrega de mensajes una sola vez, sin pérdidas ni duplicaciones. En cuanto a la seguridad, utilizamos conexión basada en usuario y contraseña, tanto en el broker como en los clientes, y el cifrado basado en certificación SSL para los datos transmitidos. Los datos recibidos se almacenan localmente en una base de datos.

IV. FUSIÓN DE DATOS Y PLANIFICADOR DE RUTAS

En base a los datos recabados tanto de las estaciones de medición como los nodos IoT asistentes, dado que el muestreo espacial es aún limitado, para poder realizar un mapeado preciso y en tiempo real de la distribución de los contaminantes en la ciudad, se utilizan técnicas estadística de interpolación espacial basadas en Kriging [15]. Kriging permite analizar información geolocalizada, basada en la autocorrelación espacial, a diferencia de otras técnicas como Inverse distance weighting (IDW) y Splines.

Una vez disponemos de toda esta información, para el desarrollo de la aplicación de usuario del cálculo de rutas saludables, se ha definido el diagrama de flujo que se muestra en la Figura 5. En este caso, inicialmente el usuario lanza una petición para el cálculo de ruta y se analiza su perfil. En base a dicho perfil, se procesan los contaminantes críticos, realizando una interpolación completa en el área de interés definida para la búsqueda. Sus concentraciones son superpuestas sobre el mapa geográfico y definen la métrica a minimizar en la búsqueda de las rutas. En la Figura 6 se muestra un ejemplo de

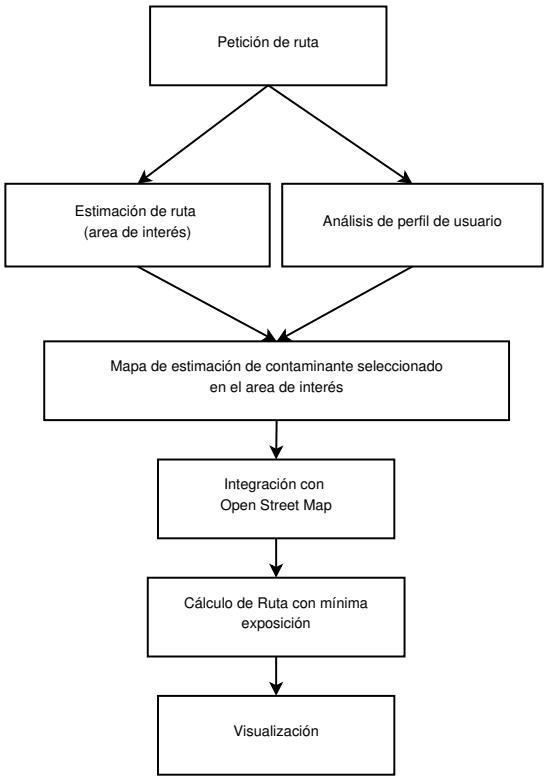


Figura 5. Diagrama de flujo de la aplicación de usuario *Eco4rupa*.

cálculo de ruta entre la ETSE-UV y el restaurante *"Tu y yo"*, ambos en Burjassot el 15/6/2023 16:00. En dicha figura, arriba se muestre el camino más corto y abajo el camino menos contaminado, donde el algoritmo ha planificado la ruta pasando cerca de un parque (La Granja), a costa de un recorrido un poco más largo. En particular, la métrica aplicada para minimizar la exposición a la contaminación que afecta al usuario, viene dada por una función de coste aplicada a los tramos del trayecto con los valores de los contaminantes extraídos de la interpolación espacial y con pesos fijados en 50 % Ozono (O3), 30 % NO2 y 20 % PM2,5. El valor de dichos contaminantes en media diezminutal en los puntos medidos más próximos al ejemplo mostrado, con su localización (latitud y longitud en grados decimales), se indican en la Tabla II en $\mu g/m^3$. Según la OMS, los máximos para estos contaminantes son 100, 25 y 15 respectivamente. En base a un análisis estadístico (no incluido en este artículo) con diferentes perfiles de usuario y trayectos realizados, se observa una reducción en media de la exposición a la contaminación en un 17.82 %, a cambio de un aumento de la distancia recorrida en un 9.8 %.

Para el cálculo de rutas y su planificación, hacemos uso de OpenStreetMap (OSM) [16], que es un proyecto colaborativo para la creación de mapas editables y libres, a través de la librería OSMnx [17] en *Python*, que nos permiten el análisis de estos mapas de forma coherente.

V. CONCLUSIONES

Los sensores de calidad del aire de bajo coste, junto con las estaciones oficiales de monitorización, nos permiten

Tabla II
CONTAMINANTES MEDIDOS [$\mu g/m^3$], 15/6/2023 16:00

Localización	O3	NO2	PM2.5
39.50961041,-0.41796381	114	8	9
39.55208592,-0.46170948	95	8	8
39.48135806,-0.44655002	107	14	9
39.47948825,-0.36955032	102	17	15
39.47071883,-0.37648469	116	39	12
39.46923859,-0.40603766	115	38	12

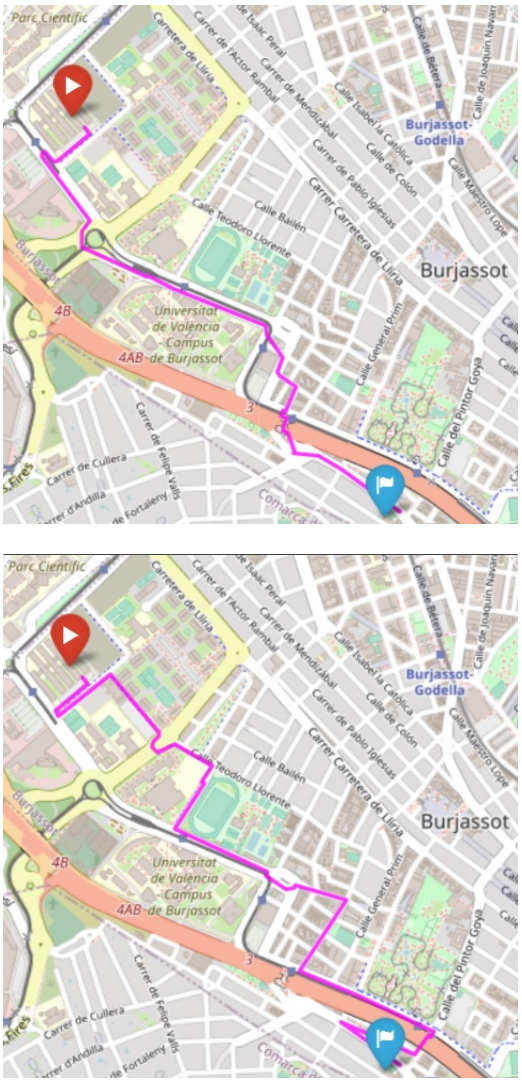


Figura 6. Arriba: ruta más corta, abajo: ruta menos contaminada

obtener información más precisa de la distribución de los contaminantes en las ciudades. Con ello, utilizando interpolación espacial con técnicas de Kriging, hemos mapeado las concentraciones de dichos contaminantes sobre el mapa de la ciudad, a fin de poder aplicar métricas basadas en estos contaminantes para el cálculo de rutas de peatones y ciclistas. Los resultados obtenidos muestran una reducción en la exposición, a costa de un aumento de la distancia recorrida.

AGRADECIMIENTOS

Este trabajo ha sido posible a la ayuda "PID2021-126823OB-I00"(*Eco4rupa*) financiada por MCIN/AEI/10.13039/501100011033 y por "European Union Next Generation EU/PRTR", junto con la Generalitat Valenciana con las subvenciones CIAICO/2022/179 (*Greenish*) y CIAEST/2022/64, y la Univèrsitat de València por la ayuda UV-INV-EPDI-2647726 y el Ministrio de Educación con la ayuda PRX22/00503.

REFERENCIAS

[1] H. Adair-Rohani, "Air pollution responsible for 6.7 million deaths every year," https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants, 2023, accessed: 27/02/2023.

[2] Eurostat: Statistics Explained, "Respiratory diseases statistics," https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Respiratory_diseases_statistics#Deaths_from_diseases_of_the_respiratory_system, September 2022, accessed: 22/05/2023.

[3] C. C. Molinari G, Colombo G, "Respiratory allergies: a general overview of remedies, delivery systems, and the need to progress." *International Scholarly Research Allergy. Hindawi*, vol. 1, pp. 1–16, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S157087051930215X

[4] 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, 2023, accessed: 27/05/2023.

[5] Nova Fitness Co., Ltd., "Air quality sensor SDS011," https://cdn-reichelt.de/documents/datenblatt/X200/SDS011-DATASHEET.pdf, 2023, accessed: 27/04/2023.

[6] DecentLab, Ltd., "Air quality sensor DL-LP8P," https://www.catsensors.com/media/Decentlab/Productos/Decentlab-DL-LP8P-datasheet.pdf, 2023, accessed: 27/04/2023.

[7] SGX, SensorTech, "Air quality sensor MiCS-6814," https://www.sgxsensortech.com/content/uploads/2015/02/1143_Datasheet-MiCS-6814-rev-8.pdf, 2023, accessed: 21/05/2023.

[8] Winsen, Ltd., "Air quality sensor zphs01b," https://www.winsen-sensor.com/d/files/zphs01b-english-version1_1-20200713.pdf, 2023, accessed: 20/03/2023.

[9] 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

[10] 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/

[11] K. T. S.L., "Calidad del Aire Urbano: Información ambiental y parámetros meteorológicos en entornos urbanos," https://www.kunak.es/, 2023, accessed: 28/2/2023.

[12] O. I. P. Ltd., "Accurate and Affordable Air Quality Monitoring Solutions," https://oizom.com, 2023, accessed: 28/2/2023.

[13] Jussi Kulonpalo, Project manager, "HOPE - Healthy Outdoor Premises for Everyone," https://uia-initiative.eu/en/uia-cities/helsinki, 2023, accessed: 21/05/2023.

[14] Pycom.io, "Fipy, five network development board for IoT," https://pycom.io/product/fipy/, 2022, accessed: 28/12/2022.

[15] I. E.H. and S. R.M., *An Introduction to Applied Geostatistics*. New York: Oxford University Press, 1989.

[16] OSM contributors, "Open Street Map," https://www.openstreetmap.org/, 2023, accessed: 27/05/2023.

[17] G. Boeing, "Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks," *Computers, Environment and Urban Systems*, vol. 65, pp. 126–139, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0198971516303970

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Converting a Weather Station into a LoRaWAN-enabled Device

Jorge Navarro-Ortiz, Natalia Chinchilla-Romero, Felix Delgado-Ferro,
Juan Jose Ramos-Munoz, Juan M. Lopez-Soler
Departamento de Teoría de la Señal, Telemática y Comunicaciones,
Universidad de Granada
C/ Periodista Daniel Saucedo Aranda, s/n
{jorgenavarro, nataliachr, felixdelgado, jjramos, juanma}@ugr.es

This paper presents a method for transforming an affordable device, such as a weather station, into a LoRaWAN-enabled device. For this purpose, a low-cost node with an ESP32 microcontroller and a SX1276 LoRa modem was employed which, in addition to provide LoRaWAN connectivity, is able to capture and decode the transmissions from the weather station in the 868 MHz band. A proof of concept was carried out to verify the success of this approach.

Keywords **– LoRaWAN, weather station, RTL-433, IoT**

I. Introduction

According to a report by Internet of Things (IoT) Analytics, the number of IoT connections is projected to rise from 7 billion in 2018 to 22 billion by 2025 [1]. In 2019, Low Power Wide Area Networks (LPWANs) accounted for 231 million of these IoT connections [2]. LPWANs are wireless networks designed specifically for IoT, enabling the connection of numerous low-cost devices. They offer long-range communication capabilities while transmitting small amounts of traffic and consuming minimal energy.

Currently, there are four dominant LPWAN technologies available in the market: LoRaWAN, Sigfox, NB-IoT, and LTE-M. Among these, LoRaWAN stands out as a highly promising solution that meets the requirements for long-range and low-power communication. It supports bit rates ranging from 250 bps to 11 kbps within the non-licensed ISM band. LoRaWAN follows an open standard, defining the medium access control (MAC) layer and network topology. Moreover, it has been deployed in 157 countries worldwide and is supported by the LoRa Alliance, a non-profit association consisting of over 500 members [3]. LoRaWAN is intended for applications that send few messages per day, e.g., logistics, smart farming, environmental monitoring, smart cities, smart grids, etc.

To get a device to generate environmental data and send it through LoRaWAN we have two possibilities. On the one hand, you could build a device, e.g., based on Arduino, that includes different types of sensors (e.g., temperature, humidity, wind, precipitation, light, etc.) and uses LoRaWAN natively. On the other hand, you could use a low-cost device that includes all these sensors and implement a gateway between the original device's wireless technology and LoRaWAN. This last approach is the one used in this paper, which can even be cheaper (e.g. the Bresser weather station [4] is available for under 100 euros, whereas assembling a weather station using Arduino [5] might incur costs exceeding 150 euros (excluding expenses tied to solar energy components)).

Thus, the objective of this article is to enable commercial products to function as LoRaWAN devices inexpensively, thereby achieving the inherent advantages of this technology (wide coverage with reduced energy cost and security in the LoRaWAN part). This will allow developers to generate LoRaWAN traffic with real data, which can be useful in research or for creating demonstrators.

The rest of the paper is organized as follows. Section II makes a brief description of LoRaWAN. The characteristics of the weather station are described in Section III, while Section IV explains how to turn it into a LoRaWAN device. Section V summarizes the proof of concept carried out, and Section VI concludes the paper.

II. LoRaWAN Overview

LoRa, the physical layer of a LoRaWAN network, is a modulation technique developed by Cycleo in 2009 and later acquired by Semtech in 2012. It operates using chirp spread spectrum modulation. One of the advantages of LoRa modulation is its receiver design, which has low complexity due to the timing and frequency offsets between the transmitter and receiver being equivalent. In this modulation scheme, the data signal is modulated onto a chirp signal that varies its frequency over time. The data

rate can vary based on the spectral bandwidth (BW, typically 125 KHz), the spreading factor (SF) employed (ranging from 7 to 12), and the coding rate (ranging from 1 to 4) following Eq. (1).

$$R_b = SF \times \left(\frac{BW}{2^{SF}}\right) \times \left(\frac{4}{4+CR}\right) \quad (1)$$

where the spreading factor (*SF*) is the first term in the equation, the symbol rate (*Rs*) is represented by the second term (symbols per second), and the third term is dependent on the coding rate (*CR*), which can vary between 1 and 4. Assuming a fixed bandwidth and coding rate, it is observed that as the spreading factor increases, the data rate decreases and the Time over Air (*ToA*) becomes higher.

The LoRaWAN open standard defines the network architecture and depicts the communication protocol that enables connectivity among all entities within the network. Fig. 1 illustrates a typical LoRaWAN network configuration.

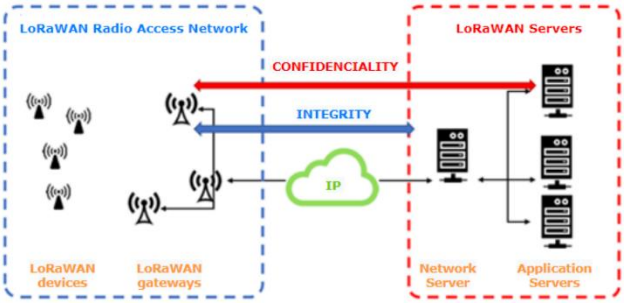


Fig. 1. LoRaWAN network architecture.

In this network configuration, a star-of-stars topology is employed. It consists of one or more gateways, also known as concentrators, which serve as intermediaries for relaying downlink and uplink traffic between a diverse range of end-nodes. These end-nodes can either be mobile or fixed at specific locations. On one hand, the gateways establish connections with the end-nodes using single-hop LoRa or FSK connections. On the other hand, the gateways utilize standard IP connections to communicate with the central network server. The network server assumes the responsibility of managing the traffic between each end-device and its associated application server. In addition, LoRaWAN provides integrity using a Message Integrity Code (*MIC*) between the end-device and the network server, using a Network Session Key (*NwkSKey*). The transmission between end-devices and the network server is encrypted using an Application Session Key (*AppSKey*). This guarantees the security of communications and allows the application owner to contract the usage of gateways and network server from an infrastructure operator, i.e., the application operator and the infrastructure operator can be different.

III. Weather Station Description

We own two low-cost weather stations, a Bresser professional 7-in-1 Wi-Fi Weather Station with Light Intensity and UV Measurement Function, located at the roof of the Higher Technical School of Computer and Telecommunication Engineering (ETSIIT) of the University of Granada (see Fig. 2), and a Bresser 7-in-1 ClimateConnect Tuya Smart Home Weather Station, located at the roof of the house of one member of the research team (see Fig. 3). Both weather stations include the same sensors: temperature, humidity, wind speed, wind direction, rainfall, UV and light intensity. The first weather station can be found for around 280 euros, and the second for around 100 euros (depending on promotions).

Both weather stations have a similar behavior. Both devices measure data, which are encoded with Pulse Code Modulation (PCM, with a pulse width of 124 μs) and then transmitted using the Frequency Shift Keying (FSK) modulation on the 868 MHz band. Data is transmitted every 12 seconds for UV, light intensity, wind speed and wind direction, and every 24 seconds for temperature, humidity and rain data.

Once the measurements have been sent to the display over 868 MHz, the display is able to post-process these data or forward them to different web weather services such as Weather Underground, Weathercloud or Awekas. This allows us to have beautiful dashboards with the current stats and to aggregate data or compute more complex calculations, e.g., for weather forecasting. These

Fig. 2. Weather station (and display) at the roof ot the ETSIIT.

Fig. 3. Weather station (and display) at one house.

features depend on the display and not on the weather station so, in our case, the web weather services are only available for the first weather station.

IV. Becoming a LoRaWAN-Enabled Device

Converting a Bresser 7-in-1 weather station to a LoRaWAN-enabled device is done in two steps. First, we have to capture and decode the data sent over the native wireless technology. Second, we have to forward this data through a LoRaWAN device.

Capturing and decoding is based on code from RTL-433 [6], adapted to the ESP32 microcontroller [7]. RTL-433 is software that allows capturing and decoding many devices using MC (Manchester Code), PCM, RZ (Return to Zero), PPM (Pulse Position Modulation), PWM (Pulse Width Modulation), DMC (Differential Manchester Code), etc., and transmission using FSK, OOK (On-Off Keying) and ASK (Amplitude Shift Keying) modulations. RTL-433 initially supported capture by RTL-SDR devices [8] and decoding by devices that used the 433 MHz band, hence its name, but now it also supports the usual bands (315, 345, and 915 MHz).

Since we want the final prototype to be low-cost, we have opted for a device based on ESP32 (which can be easily programmed using Arduino, LUA, MicroPython, etc.) that includes a LoRa chip. In particular, due to its availability, we have used a Heltec Wireless Stick mote [9], which costs less than 20 euros and has an ESP32 together with a SX1276 LoRa chip [10]. In addition to being a LoRa modem, SX1276 supports FSK, GFSK, MSK, GMSK and OOK modulations. This will allow us to capture and, with the appropriate software, decode 868 MHz transmissions from devices such as the weather stations we are using in this paper before sending the data over LoRaWAN with the same chip.

The code was developed using Arduino and allows the ESP32 to choose the period to read the data and send it via LoRaWAN. After each transmission, the device enters deep-sleep mode the rest of the time. This has two advantages. On the one hand, its energy consumption is reduced. On the other hand, every time the device wakes up it performs a reset, which improves its stability. In the latter sense, watch dog timers have also been used to eliminate the possibility of the microcontroller getting frozen. Fig. 4 shows the pseudocode for the ESP32 microcontroller.

In addition, since web weather services are only available for the weather stations with expensive displays, we have implemented an application that reads data from the LoRaWAN application server using its MQTT API and forwards them to Weather Underground, Weathercloud or Awekas using their corresponding APIs.

It should be noted that security is covered in the LoRaWAN part, but not around the weather station since it transmits unencrypted data over 868 MHz.

V. Proof-of-Concept

Our proof of concept uses a Pycom PyGate [11] as the LoRaWAN gateway and the ChirpStack framework [12] for the network and application servers. ChirpStack has different APIs to manage the different entities, including the use of MQTT (Message Queuing Telemetry Transport)

```
setup()
    lastTimeWDTReset = currentTime
    lastTransmission = -1
    dataAvailable = False
    sleepReq = True

    AddWatchdogTimer()

       if firstBoot == True:
          ResetStats()
       InitializeLoRaWAN()
       InitializeSensors()

loop()
    ResetWatchdogTimer()
    LoRaWAN.loop()
    Sensors.loop()
    checkGoingToDeepSleep()

LoRaWAN::loop()
    If dataAvailable
        LoRaWAN.sendWeatherData()
    If LoRaWAN.txComplete()
        dataAvailable = False
        sleepReq = True

Sensors::loop()
    If time > (lastTransmission + txPeriod)
        dataAvailable = ReadWeatherData()

checkGoingToDeepSleep()
    If sleepReq
        timeToSleep = computeTimeToNextTx()
        ESP.deepSleep(timeToSleep)
```

Fig 4. Pseudocode for the ESP32 microcontroller.

to query the received messages. We have also tested the usage of The Things Network [13] as LoRaWAN platform, including both the network and application servers (see the steps in [14]). The layout of the proof of concept can be observed in Fig. 5.

For the visualization of the received data, we selected the Grafana tool. For data storage, we used the InfluxDB database, which is specialized in time series. We employed Telegraf as an MQTT-InfluxDB gateway, which acts as an MQTT subscriber and stores the received data in InfluxDB. An example of this visualization can be found in Fig. 6.

The repositories with our implementation are publicly available [15][16][17].

VI. Conclusions

This paper presents a method for transforming an affordable device, such as a weather station, into a LoRaWAN-enabled device. To achieve this, we utilized the Heltec Wireless Stick, which costs less than 20 euros and incorporates an ESP32 microcontroller and a LoRa SX1276 modem. The LoRa chip's built-in FSK demodulation and other modulations eliminate the need for additional hardware to capture and decode the weather station's messages. This approach is valid for other devices transmitting on the 868 MHz band.

Furthermore, we conducted a proof of concept that involved a LoRaWAN test network comprising a PyGate gateway and the ChirpStack platform. Data visualization was achieved using Grafana, and we integrated the capability to send this data to web weather services like WeatherUnderGround, WeatherCloud, and AWEKAS.

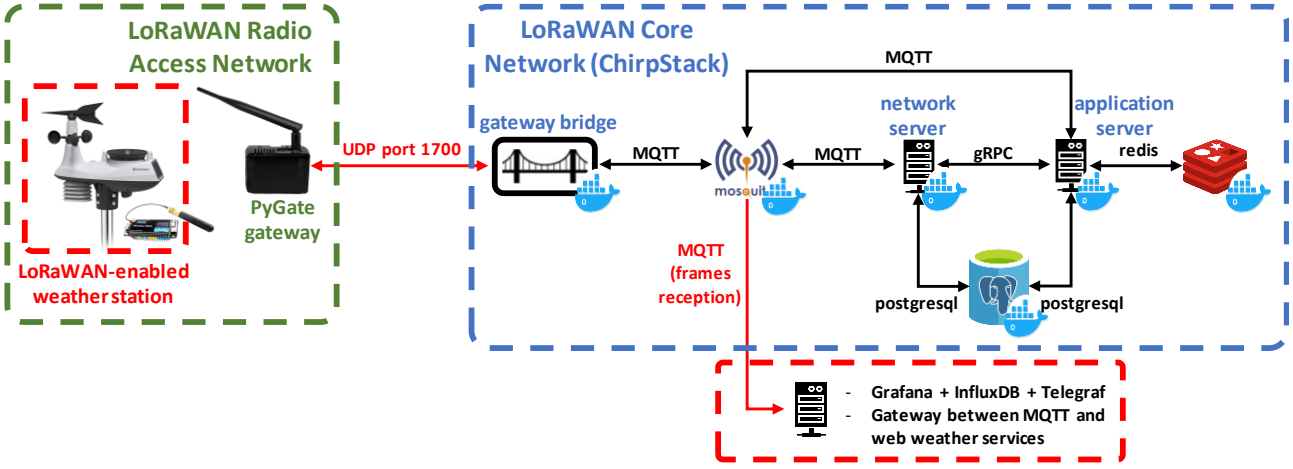


Fig. 5. Proof of concept.

ACKNOWLEDGMENTS

This research was partially funded by the Andalusian Knowledge Agency (project B-TIC-568-UGR20), the Spanish Ministry of Science and Innovation (projects PID2019-108713RB-C53 and PID2022-137329OB-C43), the Spanish Ministry of Economic Affairs and Digital Transformation (project TSI-063000-2021-28) and the Spanish Ministry of Universities (FPU Grant Number: 20/02621).

Fig. 6. Grafana dashboard with some stats from the weather station.

REFERENCES

[1] IoT Analytics. LPWAN Market Report 2018–2023. Available online: https://iot-analytics.com/lpwan-market-report-2018-2023-new-report/ (accessed on 30th May 2023).

[2] IoT Analytics. 5 Things to Know about the LPWAN Market in 2020. Available online: https://iot-analytics.com/5-things-to-know-about-the-lpwan-market-in-2020/ (accessed on 30th May 2023).

[3] LoRa Alliance. Available online: https://lora-alliance.org/ (accessed on 30th May 2023).

[4] BRESSER 7-in-1 ClimateConnect Tuya Smart Home Weather Station. Available online: https://www.bresser.de/en/Weather-Time/BRESSER-7-in-1-ClimateConnect-Tuya-Smart-Home-Weather-Station.html (accessed on 30th May 2023).

[5] Solar Powered WiFi Weather Station V3.0. Available online: https://www.instructables.com/Solar-Powered-WiFi-Weather-Station-V30/ (accessed on 30th May 2023).

[4] RTL-433 generic data receiver. Available online: https://github.com/merbanan/rtl_433 (accessed on 30th May 2023).

[5] ESP32, a feature-rich MCU with integrated Wi-Fi and Bluetooth connectivity for a wide range of applications. Available online: https://www.espressif.com/en/products/socs/esp32 (accessed on 30th May 2023).

[6] RTL-SDR and software defined radio news and projects. Available online: https://www.rtl-sdr.com/ (accessed on 30th May 2023).

[7] Heltec Wireless Stick. Available online: https://heltec.org/project/wireless-stick/ (accessed on 30th May 2023).

[8] Semtech's SX1276. Available online: https://www.semtech.com/products/wireless-rf/lora-connect/sx1276 (accessed on 30th May 2023).

[9] Pycom's PyGate. Available online: https://docs.pycom.io/tutorials/expansionboards/pygate/ (accessed on 30th May 2023).

[10] ChirpStack, open-source LoRaWAN Network Server. Available online: https://www.chirpstack.io/ (accessed on 30th May 2023).

[11] The Things Network. Available online: https://www.thethingsnetwork.org/ (accessed on 30th May 2023).

[12] LoRaWAN testbed repository. Available online: https://github.com/jorgenavarroortiz/testbed-LoRaWAN (accessed on 30th May 2023).

[13] BresserWeatherSensorReceiver with Bresser 7-in-1 decoder. Available online: https://github.com/jorgenavarroortiz/BresserWeatherStationLoRaWAN (accessed on 30th May 2023).

[14] Platform to show weather station data in Grafana. Available online: https://github.com/jorgenavarroortiz/weatherstation-mqtt-grafana (accessed on 30th May 2023).

[15] LoRaWAN network using Pycom gateway (PyGate) and ChirpStack platform. Available online: https://github.com/jorgenavarroortiz/testbed-LoRaWAN (accessed on 30th May 2023).

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Transmission of Images over LoRa

Jorge Navarro-Ortiz, Natalia Chinchilla-Romero, Felix Delgado-Ferro,
Juan Jose Ramos-Munoz, Juan M. Lopez-Soler, Fernando Tejero-Rodríguez
Departamento de Teoría de la Señal, Telemática y Comunicaciones,
Universidad de Granada
C/ Periodista Daniel Saucedo Aranda, s/n
{jorgenavarro, nataliachr, felixdelgado, jjramos, juanma}@ugr.es, fernandotr24@correo.ugr.es

In this research paper, a proposal is presented for a MAC layer designed specifically for image transmission over LoRa. The proposed layer utilizes polite spectrum access to overcome the limitations imposed by the duty cycle constraint. By implementing channel sensing, transmitting nodes can ensure efficient utilization of the available spectrum. Moreover, a contention mechanism has been incorporated to effectively reduce collisions. The feasibility of this solution has been demonstrated through its implementation in a testbed, serving as a proof of concept. During the initial performance evaluation, it was observed that the transmission of JPEG images via LoRa, with a resolution of 640x480 pixels and an average size of 11.3 KB, requires an average of 7.9 seconds.

Keywords **– LoRa, multimedia communications, polite spectrum access, listen-before-talk, contention**

I. Introduction

According to a report by Internet of Things (IoT) Analytics, the number of IoT connections is projected to rise from 7 billion in 2018 to 22 billion by 2025 [1]. In 2019, Low Power Wide Area Networks (LPWANs) accounted for 231 million of these IoT connections [2]. LPWANs are wireless networks designed specifically for IoT, enabling the connection of numerous low-cost devices. They offer long-range communication capabilities while transmitting small amounts of traffic and consuming minimal energy.

Currently, there are four dominant LPWAN technologies available in the market: LoRaWAN, Sigfox, NB-IoT, and LTE-M. Among these, LoRaWAN stands out as a highly promising solution that meets the requirements for long-range and low-power communication. It supports bit rates ranging from 250 bps to 11 kbps within the non-licensed ISM band. LoRaWAN follows an open standard, defining the medium access control (MAC) layer and network topology. Moreover, it has been deployed in 157 countries worldwide and is supported by the LoRa Alliance, a non-profit association consisting of over 500 members [3]. LoRaWAN is intended for applications that send few messages per day, e.g., logistics, smart farming, environmental monitoring, smart cities, smart grids, etc.

The purpose of the developed MAC layer is to be able to use low-cost LoRa devices to transmit images without the need to include new wireless technology hardware. This will be useful in a multitude of use cases, such as visual verification of fire alarms in rural areas, visual aid to locate lost UAVs without coverage, or for agricultural monitoring.

There are several works about image transmission over LoRa. Chen [4] introduced a communication protocol called MPLR for transmitting images via LoRa, which employs bit-vector acknowledgments to reduce the number of ACKs. He also utilized a channel reservation protocol to reduce the collision probability. Ching-Chung [5] multiplexed the transmission of an image using, thanks to their orthogonality, different spreading factors, thus reducing the image transmission time. In [6], Jebril proposed to use the LoRa physical layer with a new encryption method for image transmission. Wei [7] tested different image encoding methods to analyze their impact on the transmitted file size and image transmission efficiency. To the best of our knowledge, these works did not consider contemption mechanisms nor other MAC layer modifications to overcome the duty cycle limitations in LoRa.

The rest of the paper is organized as follows. Section II makes a brief description of LoRa. Section III describes the modifications to the MAC layer proposed, while Section IV summarizes the proof of concept carried out. Section V concludes the paper.

II. LoRa Overview

LoRa [8] is a modulation technique developed by Cycleo in 2009 and later acquired by Semtech in 2012. It operates using chirp spread spectrum modulation. One of the advantages of LoRa modulation is its receiver design, which has low complexity due to the timing and frequency

offsets between the transmitter and receiver being equivalent. In this modulation scheme, the data signal is modulated onto a chirp signal that varies its frequency over time. The spectral bandwidth (BW) of a LoRa signal is determined by the chirp rate, where a 125 kHz BW corresponds to a chip rate of 125,000 chips per second. The data rate, on the other hand, can vary based on the spreading factor (SF) employed, with SF ranging from 7 to 12. The spreading factor represents the number of raw bits carried per symbol. Consequently, the data rate (*Rb*) can be calculated using the following formula:

$$R_b = SF \times \left(\frac{BW}{2^{SF}}\right) \times \left(\frac{4}{4+CR}\right) \quad (1)$$

where the spreading factor (*SF*) is the first term in the equation, the symbol rate (*Rs*) is represented by the second term (symbols per second), and the third term is dependent on the coding rate (*CR*). LoRa modulation incorporates a FEC (Forward Error Correction) scheme to detect and correct erroneous bits by introducing redundancy in the code, thereby enhancing the robustness of the transmitted signal. The coding rate (*CR*) can vary between 1 and 4. Assuming a fixed bandwidth and coding rate, it is observed that as the spreading factor increases, the data rate decreases and thus, the Time over Air (*ToA*) becomes higher. Table 1 presents the data rate configuration for different bandwidths with a Coding Rate of 4/5 (*CR* = 1). The bit rate increases and the ToA decreases as the SF is decreased.

In the case of Europe, the band used for LoRa and LoRaWAN is ISM868. The use of this band forces us to use a duty cycle that depends on the specific subband, being the most restrictive of 0.1% among all the channels [9]. This means that, for example, if a transmitter takes 1 second to transmit, it would not be able to transmit again until 999 seconds later ($duty\ cycle = \frac{1}{999+1} = 0.001 = 0.1\%$). However, the regulations [10] also allow the use of *polite spectrum access*. In this case, the device senses the channel for at least the Clear Channel Assessment (CCA) interval to determine if it is free. The main parameters of CCA are included in Table II.

It shall be noted that the maximum allowed cumulative on time can be increased by changing the operating channel (*Adaptive Frequency Agility* or AFA).

The use of *polite spectrum access* will allow the transfer rate of LoRa and LoRaWAN devices to be significantly increased, since it will eliminate the need to comply with such a restrictive duty cycle. In the case of Europe, LoRa devices can generally use 8 channels of 125 kHz, i.e., a total of 1000 kHz, so it could transmit up to 500 seconds per hour (100 seconds/hour for each 200 kHz).

Table I. DATA RATE CONFIGURATION IN THE EU868 ISM BAND (24-BYTE FRAME, CODE RATE = 4/5, WITH CRC).

	Effective Data Rate (bps)			ToA (ms) for a 24B Packet		
BW→ SF ↓	125 kHz	250 kHz	500 kHz	125 kHz	250 kHz	500 kHz
12	293	586	1172	1319	659	330
11	537	1074	2148	741	371	185
10	977	1953	3906	371	185	93
9	1758	3516	7031	206	103	51
8	3125	6250	12500	113	57	28
7	5469	10937	21875	62	31	15
6	9375	18750	37500	34	17	9

Table II. CLEAR CHANNEL ASSESMENT PARAMETERS.

Parameter	Value	Notes
Min CCA interval	160 µs	Min CCA listening period
Min deferral period	CCA interval	Min value of the deferral interval
Dead Time	Declared by vendor, not exceeding 5 ms	Max time between the end of a listening interval and the start of a transmission
Max Tx duration (T_{on} *max*)	1 s	For a single transmission
Max Tx duration (T_{on} *max*)	4s	For a transmission dialogue or a polling sequence
Max T_{cum} *on* over 1 hour	100 s / 1 h per 200 kHz spectrum	Max allowed Cumulative On Time over a 200 kHz proportion of spectrum per hour
Min T_{off} *min* on the same operating frequency	100 ms	The min T-off time period where a specific transmitter shall remain off after a transmission on the same operating frequency
CCA radiated threshold limit (e.r.p < 100 mW)	-79 dBm	15 dB above Rx sensitivity level limit, value for devices with 200 kHz
CCA radiated threshold limit (e.r.p from 100 mW to 500 mW)	-84 dBm	11 dB above Rx sensitivity level limit, value for devices with 200 kHz

III. PROPOSED MAC LAYER

Based on the *polite spectrum access* regulations, we propose to employ a *listen-before-talk* approach to avoid the duty cycle limitations. This will allow us to transmit low resolution images via LoRa in a reasonable amount of time. Since multiple packets will be transmitted for one image, we included a negative block ACK mechanism to reduce signaling due to retransmissions.

The process for transmitting an image with our modified MAC layer is as follows:

- First, the transmitter sends a signaling message with the content "#IMAGE# <ImageSize>". This notifies the receiver that the transmitter is going to send an image of a certain size.
- Later, the transmitter splits the image into packets according to their maximum size, which will depend on the chosen spreading factor.
- The LoRa frames associated to an image have an application header that includes the source (4-bytes *src* field), the destination (4-bytes *dst* field), the transaction identifier (2-bytes *transID* field) and the sequence number within the transaction (4-bytes *seq* field). Each image is identified with a different transaction.
- Similar to IEEE 802.11, our MAC layer also employs a contention procedure to reduce the number of collisions.
- Once the packets are received, a BLOCKNACK message is sent that includes the chunks (consecutive frames) not received. If all packets have been received correctly, an empty BLOCKNACK is sent. This is done when the last packet of the image is received or when no packets have been received from this transaction during the *packetReceptionTimeout* parameter. The maximum number of retransmissions

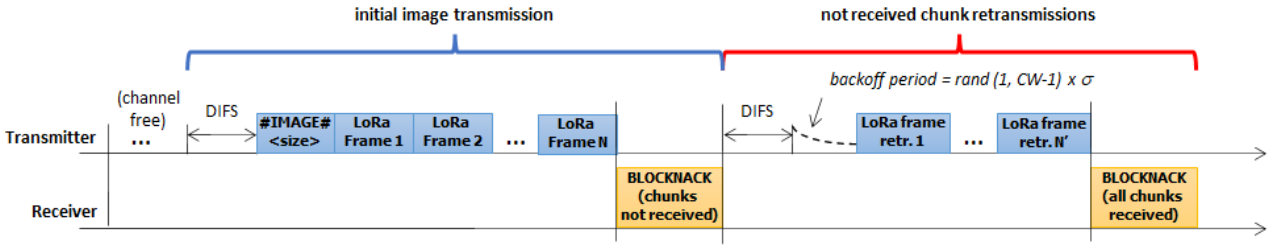


Fig. 1. Example of the transmission of one image using our modified MAC layer, including the contemption procedure.

of signaling messages is indicated by the maxNoRetx parameter.

- Once a BLOCKNACK is received, the packets corresponding to the indicated chunks are retransmitted.
- The process ends when all the image packets have been correctly received or when the maximum number of signaling message transmissions has been exceeded (assuming there are coverage problems).

The contention procedure employs the following parameters: DIFS (Data InterFrame Space), a slot time (σ), CW_{min} and CW_{max} (Contention Window minimum and maximum values). If the channel is free, the transmitter node senses the channel during a DIFS. For consecutive (re)transmissions, the transmitter will sense the channel for DIFS plus a *backoff* period, which is computed as a random number of slots between 0 and CW – 1. If the channel is busy, CW is doubled (up to its maximum value). If the channel is free, CW returns to its minimum value. Hence, in contrast to IEEE 802.11, the evolution of the contention window depends on whether the medium has been found to be busy or not, not on whether there have been collisions. This is easy to implement and avoids the need to receive ACKs and their corresponding delays.

To increase the cumulative transmission time by more than an hour, the transmitter uses the next channel for a new image to overcome the T_{cum} *on* limitation. In addition, the transmitter will have to wait T_{off} *min* if the transmission of an image takes longer than T_{on} *max*.

Fig. 1 shows an example of this procedure, considering that there has only been one retransmission.

IV. Proof-of-Concept

Our proof of concept uses Pycom Fipy modules as LoRa devices. One of the main advantages of the Fipy board is that it supports LoRaWAN class C, which will be useful to extend this work to LoRaWAN. As the maximum transmit power of these devices is 25 mW (14 dBm), the threshold to sense the channel would be -79 dBm according to Table I.

We connected a 2MP camera to the transmitter node. The camera includes an OV2640 sensor with a SPI (Serial Peripheral Interface) interface. The tests were carried out by taking photos with a resolution of 640x480 pixels, which generated JPG images.

To maximize the data rate in the tests, we chose SF7 with a bandwidth of 500 kHz, which allows a maximum LoRa frame of 222 bytes. Considering a 14-byte application header, this means that we can send 206 bytes of data per frame. Assuming a typical 10KB image, approx. 50 frames were sent per image.

The proposed MAC layer has been implemented using MicroPython in the FiPy motes, both for the transmitter and the receiver. The repository with our implementation will be made publicly available [11] once we have extended this work for a journal paper.

For our experimentation, the following values were used: DIFS = 5 ms, σ = 2 ms, CW_{min}=16, CW_{max}=1024. These values are similar to those of 802.11 but scaled by the difference in data rate.

Additionally, we developed a web site which receives the images (sent over Wi-Fi from the receiver node) and shows them in real-time. For that purpose, we utilized Ajax (to seamlessly update the browser) and PHP (to process and store the images).

Our testbed is shown in Fig. 2, where you can see the transmitter node with the camera and the receiver node. In addition, Fig. 3 shows an example of traces (including busy channel occurrences and BLOCKNACK sending) a nd the web server showing the photograph sent via LoRa.

As an initial performance evaluation of our solution, we conducted the transmission of 100 images. The transmitting node was moved during the experiment to ensure that the pictures were not always the same. The size of the images and their transmission time were measured, and the results are presented in Fig. 4. As observed, image size ranges from 9.3 KB (10th percentile) to 14.8 KB (90th percentile), with an average of 11.3 KB. This corresponds to transmission times varying from 5.7 s (10th percentile) to 9.8 s (90th percentile), with an average of 7.9 s.

V. Conclusions

This paper presents a proposal for a MAC layer for image transmission over LoRa. This layer is based on the use of polite spectrum access, which eliminates limitations due to the usage of a duty cycle. To achieve this, the transmitting nodes must sense the channel before transmitting. Additionally, a contention mechanism has been included to reduce the number of collisions. This solution has been implemented in a testbed as proof of concept. In an initial performance evaluation, it was

Fig. 2. Proof of concept.

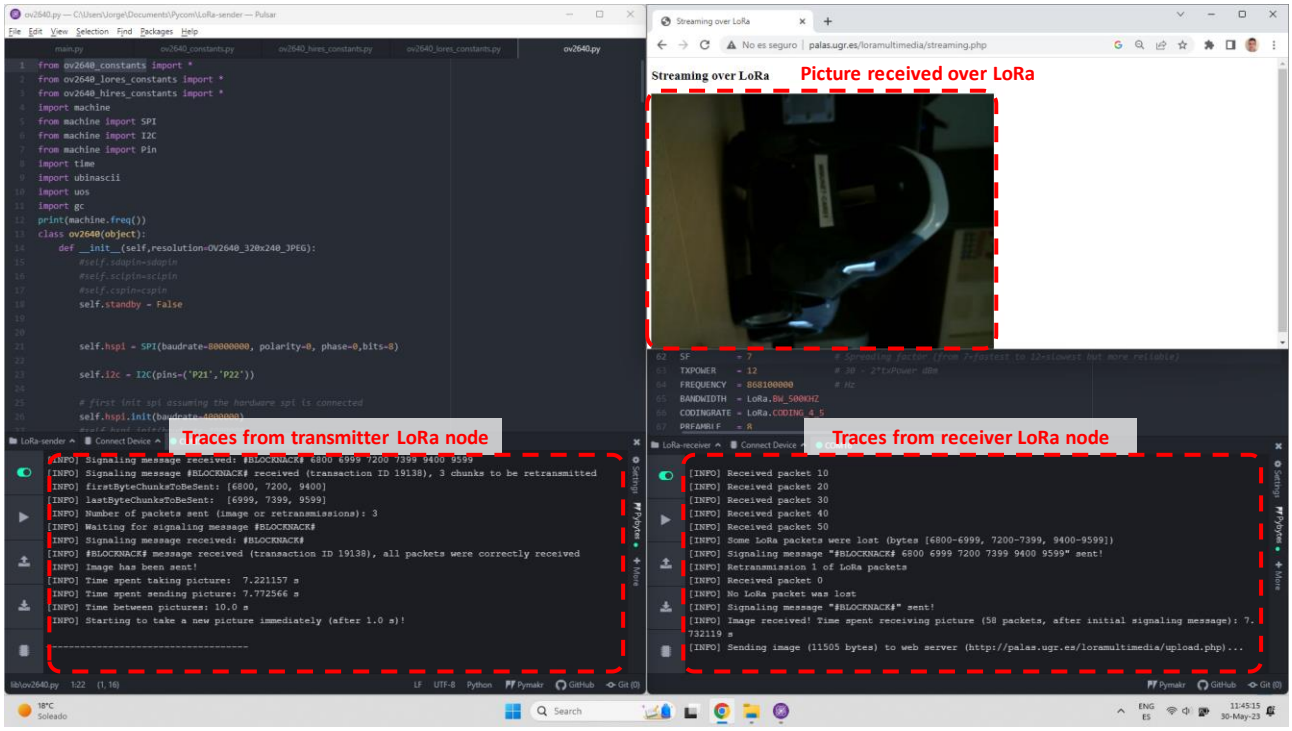


Fig. 3. Proof of concept: traces and image received.

observed that the LoRa transmission of JPEG images with a resolution of 640x480 pixels, with an average size of 11.3 KB, takes an average of 7.9 seconds.

As future work, the following aspects are included: 1) studying the impact of different parameter values (DIFS, slot time, CWmin, CWmax), 2) developing a mathematical model to estimate the transmission time of images and the collision probability of frames, 3) creating a simulator to verify the proper functioning of the mathematical model and 4) expanding the solution to transmit H.264 video frames (e.g., sent by UAVs).

ACKNOWLEDGMENTS

This research was partially funded by the Andalusian Knowledge Agency (project B-TIC-568-UGR20), the Spanish Ministry of Science and Innovation (projects PID2019-108713RB-C53 and PID2022-137329OB-C43), the Spanish Ministry of Economic Affairs and Digital Transformation (project TSI-063000-2021-28) and the Spanish Ministry of Universities (FPU Grant Number: 20/02621).

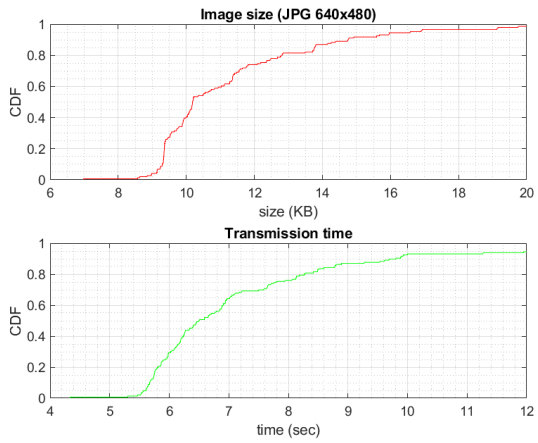


Fig. 4. CDFs of image size and transmission time.

REFERENCES

[1] IoT Analytics. LPWAN Market Report 2018–2023. Available online: https://iot-analytics.com/lpwan-market-report-2018-2023-new-report/ (accessed on 30th May 2023).

[2] IoT Analytics. 5 Things to Know about the LPWAN Market in 2020. Available online: https://iot-analytics.com/5-things-to-know-about-the-lpwan-market-in-2020/ (accessed on 30th May 2023).

[3] LoRa Alliance. Available online: https://lora-alliance.org/ (accessed on 30th May 2023).

[4] T. Chen, D. Eager and D. Makaroff, "Efficient Image Transmission Using LoRa Technology In Agricultural Monitoring IoT Systems," 2019 International Conference on Internet of Things (iThings), Atlanta, GA, USA, 2019, pp. 937-944, doi: 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00166.

[5] C. -C. Wei, S. -T. Chen and P. -Y. Su, "Image Transmission Using LoRa Technology with Various Spreading Factors," 2019 2nd World Symposium on Communication Engineering (WSCE), Nagoya, Japan, 2019, pp. 48-52, doi: 10.1109/WSCE49000.2019.9041044.

[6] A. Jebril, A. Sali, A. Ismail, and M. Rasid, "Overcoming Limitations of LoRa Physical Layer in Image Transmission," Sensors, vol. 18, no. 10, p. 3257, Sep. 2018, doi: 10.3390/s18103257.

[7] C.-C. Wei, P.-Y. Su, and S.-T. Chen, "Comparison of the LoRa Image Transmission Efficiency Based on Different Encoding Methods," IJIEE 2020 Vol.10(1): 1-4 ISSN: 2010-3719, Mar. 2020, doi: 10.18178/IJIEE.2020.10.1.712.

[8] Application Note AN1200.22, LoRa modulation basics, rev. 2, May 2015, Semtech.

[9] Technical Note TN1300.01, rev. 1.0, Semtech, Feb. 2018. Available online: https://lora-developers.semtech.com/?ACT=72&fid=30&aid=48_0znCpZpvImL3agza59hG&board_id=1 (accessed on 30th May 2023).

[10] ETSI EN 300 220-1 v3.1.1 (2017-02), Short Range Devices (SRD) operating in the frequency range 25 MHz to 1000 MHz; Part 1: Technical characteristics and methods of measurement. Available online: https://www.etsi.org/deliver/etsi_en/300200_300299/30022001/03.01.01_60/en_30022001v030101p.pdf (accessed on 30th May 2023).

[11] GitHub repository "LoRa multimedia". Available online: https://github.com/jorgenavarroortiz/LoRa-multimedia (accessed on 30th May 2023).

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Measuring the influence of spam on public opinions about COVID 19 news: a study of YouTube comments on the daily reports of the first spokesperson for the Spanish public health system

María Asunción Vicente[1], César Fernández[1], Mercedes Guilabert[2], Irene Carrillo[2] & José Joaquín Mira[2,3]
[1]Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche, Universidad Miguel Hernández (UMH)
[2]Departamento de Psicologia de la Salud, Universidad Miguel Hernández (UMH)
[3]Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO)

1,2 (UMH), Avenida de la Universidad s/n Elche (03202) Alicante, España
3 (FISABIO), Hospital General Universitario, Ed. Anexo II, 3ª planta, Camí de L´Almassera 11, Elche (03203), Alicante, España

suni@umh.es, c.fernandez@umh.es, mguilabert@umh.es, icarrillo@umh.es, jose.mira@umh.es

Social media, especially YouTube, is a crucial source of COVID-19 information for the public. Viewers often rely on the comments section to gather additional information and opinions. While social networks claim to combat fraud and misinformation, their automatic spam-filtering falls short. This study aimed to examine spam presence in YouTube's COVID-19-related comments, measure its influence on prevalent topics, and classify common spam messages. The study analyzed periodic speeches by the first official Spanish spokesperson, uploaded by a public Spanish TV channel during COVID-19 waves. NLP methods processed the extracted comments dataset. Findings revealed that 16% of comments consisted of copied and pasted spam. The most repeated message occurred 137 times in a single video. Manual classification showed that political, hate, and satirical messages accounted for 71.2% of total spam. Modified versions of messages often went undetected by automatic filters. Utilizing advanced NLP techniques could enhance spam detection, and removing repetitive messages would improve opinion sharing among readers.

Keywords **- misinformation, spam, fake news detection, NLP, COVID-19, social media**

I. Introduction

A. Sources of COVID-19 related information preferred by users.

There are multiple studies analyzing the preferred sources of information about COVID-19 data. In [1], a survey (completed by 5948 adults in central Pennsylvania during the first months of the pandemic) concluded that the most widely used source of information was government websites (Center for Disease Control [CDC], National Institutes of Health [NIH], and the World Health Organization [WHO]) (42.8%), followed by television news (27.2%). A different study from the same group (5911 valid survey answers), presented in [2] compared the sources of information preferred by healthcare workers (HCWs) and non-HCWs, concluding that television news channels were trusted more among non-HCWs and non-CDMS (non-Clinical Decision Makers). In [3], the study was focused on adolescents and young adults with cancer (663 participants), and the conclusion was that their preferred COVID-19 information sources were their cancer institutes and social media. In Greece [4], a study covering young to senior participants concluded that television, electronic press, and news websites were preferred, while social media had limited acceptance, and concerns about fake news were detected. In [5], an online survey during March 2020 showed that those respondents that used Facebook as an additional source of information were less likely to answer COVID-19 questions correctly than those who did not. A study in India with 1310 participants [6] ranked social media (including YouTube) in 5th place among the preferred information sources. Above social media in the rank there were 1) news websites; 2) government websites; 3) medical staff; and 4) family and friends. Below social media, there were 6)

paper or digital newspapers; 7) television; 8) emails/circulars; 9) posters/brochures; and 10) radio. Concerning which particular social media is preferred, a different study with 159 respondents, also in India [7] showed that YouTube was the second most used platform for seeking COVID-19 information, after WhatsApp and before Facebook, Telegram, Twitter, or Instagram.

B. *Quality of COVID-19 related information*

The quality of the COVID-19 related information presented in social media has also been analyzed. In [8], a review of 22 studies was presented. Depending on the study, the proportion of misinformation ranged from 0.2% to 28.8% of posts. Those studies that addressed the possible consequences of misinformation reported that it led to fear or panic. Focusing on YouTube, the quality and engagement capability of the 137 most viewed COVID-19 related videos was analyzed in [9]. The results showed that, although news channel videos were majority among those 137, they had the lowest quality. Highest quality videos were those uploaded by physicians, health organizations and education channels. A similar study was presented in [10], where among the 113 most-widely viewed videos about COVID-19, 69.9% were classified as useful, and 8.8% were classified as misleading. Besides, news agencies were more likely to post useful videos, whereas independent users were most likely to post misleading videos.

It is also important to consider not only the content posted (through YouTube or any other social media) but also the comments of the viewers. It is common to read these comments as an extra source of information and to know the opinions of other persons. Misinformation, fake news, or spam can be present also in the comments section and can create confusion in viewers. Focusing on YouTube, there have been multiple attempts to automatically filter out spam-like comments. The study presented in [11] evaluated different algorithms that can learn to filter spam comments. According to this study, the best performing algorithm was Adaptive Genetic Algorithm (AGA), capable of classifying correctly 99.1% of comments as spam or not-spam, over a dataset of 100 YouTube channels and 10,000 samples overall. There are many other studies proposing automatic spam-filtering techniques for YouTube, among them [12], where Artificial Network classifiers proved to perform better than standard classifiers; [13], which confirmed the prediction accuracy of Artificial Neural Networks over a database obtained from the UCI machine learning repository; or [14], where instead of a single detection model, a Cascaded Ensemble Machine Learning Model was proposed, and proved to outperform strategies based on using one single detection model. Even though these automatic filtering methods could perform correctly in most cases, they are usually not applied.

C. *Content filtering approaches*

According to the information given in the YouTube Help website [15], the platform YouTube detects spam based on the text of a comment, or by the behavior from a particular commenter. For instance, repeatedly posting comments can be detected as spam. Additionally, uploaders have more control over comments made on their videos. They can review any comment before it's displayed or remove it altogether. In the same website, YouTube considers spam those contents that create a negative experience by making it difficult to find more relevant and substantive material.

Concerning COVID-19 related information, YouTube (as well as Facebook, Twitter, and other companies) issued a joint statement stating that they were jointly combating fraud and misinformation about the virus [16]. The main mobile application platforms (Android and iOS) also established strict policies [17, 18] about COVID-19 related applications, which were not published unless they were backed up from a public organization.

In this sense, YouTube removes content which they consider contradictory with health authorities or the WHO, even if it has been posted by researchers, like, for example, the post by John P.A. Ioannidis [19, 20] (please note that discussing the ethics of these content filtering approaches is not one of the goals of this paper).

However, these YouTube content filtering strategies seem to focus mainly on the videos uploaded, not on the viewer comments.

D. *Focus of our study*

In our study, we measured the influence of spam in public opinions to daily COVID-19 reports from Spanish official spokesman. The goals were:

- To detect whether either YouTube or the uploader had successfully removed spam comments.
- If spam is present, to compare the main topics and trends before and after cleaning out spam comments, to measure the influence of spam.
- If spam is present, to detect the type of spam most found.

II. Methods

Video search was limited to presentations of the official spokesman for the Spanish public health system (Dr. Fernando Simón) uploaded to the YouTube channel of the Spanish public TV (RTVE or Radio Televisión Española). Each video represented one of the periodic speeches given by the spokesman from April 29, 2020 to September 9, 2021 (covering the worst episodes of the pandemic in Spain).

An ad-hoc created Python [21] script was used to extract all the comments (and answers to comments) from each of the videos, by means of the official YouTube API [22]. Data extraction was performed on December 10, 2021; so, the study includes all comments posted up to such date.

For the purpose of the paper, we only considered as spam comments those comments that were repeated literally (copied and pasted comments, even if they were comments to different videos). Fragments of comments (sentences) were also searched for repetitions, up to a minimum length of 6 words. Repeatedly copied and pasted sentences are considered spam as they try to impose its message above the others and mislead the readers. To

avoid this bias, only one instance of each copy was kept in the filtered dataset. Textacy NLP tool [23] was used, through a Python script, to detect repeated sentences through all videos (up to a minimum length of 6 words).

No other spam detection was carried out, in order not to add subjectivisms to the study. Besides, the comments section of a YouTube video is supposed to allow the viewers to express their opinions freely.

As a result of the filtering, two datasets were generated for further comparative analysis: original comments and spam-filtered comments.

Stacy NLP tool [24] was used, through a Python script, to remove stop words and lemmatize the text information of each dataset (original and spam-filtered). The goal was to remove non-meaningful words and to group variations of the same words (e.g., singular or plural) or verbs (e.g., different tenses). After lemmatization, Stacy was also used to perform word count analysis. Finally, word clouds were displayed with Matlab [25].

Words counts for each dataset were compared both numerically (via Chi-squared test) and graphically (through word clouds). The most deleted terms after filtering were also identified (as percentage of deleted items vs original items).

An extra dataset with all detected spam (repeated comments) was also generated for an additional quantitative analysis. First, the number of repetitions of each spam was registered. Second, the most prominent spam messages were selected for a manual classification. The criteria for selection was: messages appearing 3 or more times and having more than 10 words (shorter messages may not be meaningful enough for classifying them). Messages were classified in the following categories: (1) opinions, (2) fake news, (3) hate messages, (4) satiric messages, or (5) politic messages. It must be stated that messages were classified as spam due to their repetitions, but not due to the content of the message itself. That is the reason why a "opinions" category is present as spam: obviously, expressing our opinion is not spam, but copying and pasting such opinion several times is. These classes were selected as the most appropriate for classifying comments, because standards for classifying misinformation are usually focused on content (not comments). Further details are given in the discussion section.

Classification was carried out by two independent coders (discrepancies were accounted for by adding 0.5 to both classes). A Kappa test was used to measure the coherence of the classifications.

All statistics were computed using R software, version 3.6.2 for Mac OS X.

III. Results

A. Descriptive statistics

A total of 109 videos were retrieved. The number of comments per video (including answers to comments) ranged from 1 to 206 (mean value 35.94). The total length of the video comments, measured as the total number of words, ranged from 14 to 7025 (mean value 1100). These values, as well as video dates and URLs are available as multimedia appendix #1.

Figure 1 compares the evolution of the COVID-19 incidence in Spain (number of deaths per day) against the number of comments to the videos published (when several videos were published in the same day, the comment counts were added). An increase in the number of comments was clear up to 02/07/2020, where a peak number of 341 comments was reached (two videos were posted that day).

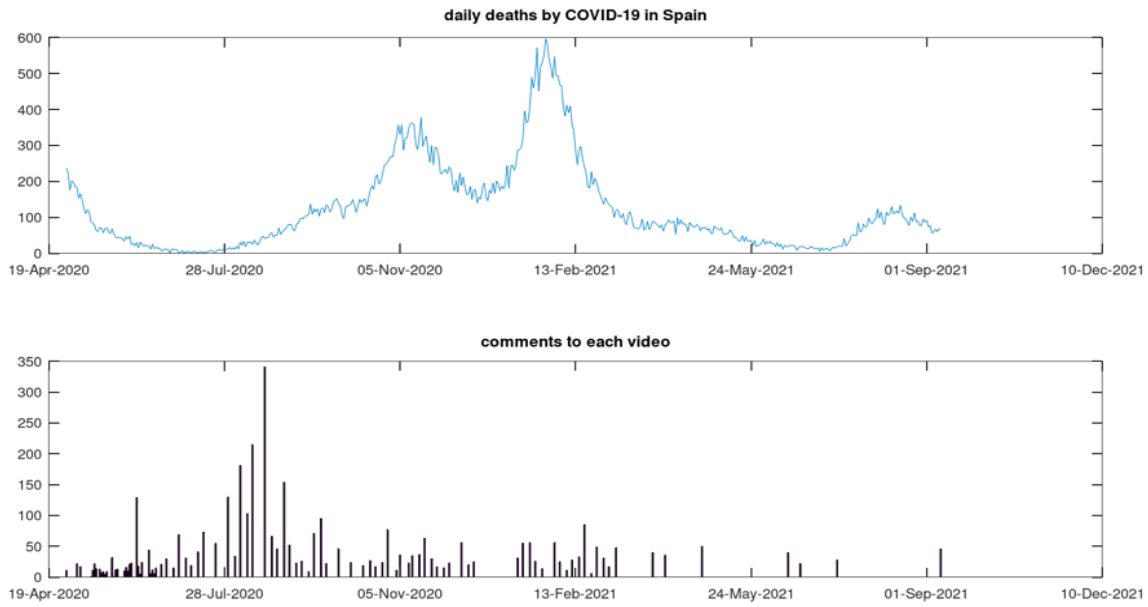


Fig. 1. Temporal evolution in the number of comments per video.

Afterwards, the number of comments decreased. Comparing both curves, the dates with a higher number of comments did not seem to be related to the worst episodes of the pandemic.

B. Detection of repeated messages

The detection and elimination of repeated messages, or spam messages, converted the initial dataset of 118807 words (all comments to all videos) to a filtered dataset of 99792 words (16% reduction in size or, in other words, 16% of spam content). The most repeated message was found 137 times in the comments (in particular, it was a message blaming electromagnetic contamination as the cause of the disease). Interestingly, the 137 repetitions of this message were comments to the same video (dated 30/11/2020), which clearly shows that YouTube spam-protection measures did not work correctly in this case.

Table I
NUMBER OF REPETITIONS PER SPAM MESSAGE (COMPLETE)

Number of repetitions	Different messages
137	1
62	1
40	1
27	1
25	2
24	1
21	1
20	2
18	4
17	1
16	1
14	1
10	1
9	2
8	4
7	3
6	5
5	1
4	15
3	11
2	145

Other highly repeated messages (62, 40 and 27 times respectively) were hate messages against the spokesperson (Dr. Fernando Simón) and the Spanish Ministry of Health (Salvador Illa). These messages were repeatedly found as comments to multiple videos. Table 1 shows all histogram values. Multimedia appendix #2 contains the list of the most repeated messages (messages repeated 3 or more times and containing 10 or more words), both in Spanish and translated to English.

C. Measures of the influence of spam

The goal of spam comments is to impose its message above the others and mislead the readers. Our first measure of influence was the direct comparison of the influence of all terms detected (after lemmatization) in both datasets (original and spam-filtered). The influence of each term was computed as the word count, or the number of times each term appeared in the comments. Table 2 shows the comparison for the most prominent terms. It becomes clear that word count for certain terms was reduced considerably after removing spam and that the most prominent terms were not the same in both datasets (both the original Spanish terms and their translation are included). A slightly different representation is shown in table 3, where the items have been reordered according to the ratio deleted vs kept.

Words counts before and after spam-filtering were compared using a Chi-squared test (only for terms with at least 50 repetitions in the original dataset). The result showed statistically significative differences between both datasets ($P<0.001$, Chi-squared $109 = 405.43$).

The previous results can also be represented graphically through word clouds. Figure 2 shows the word clouds for the original dataset and the spam-filtered dataset. Although the same terms appear before and after filtering, their relevance changes considerably. Particularly, certain hate comments (like “ssimon”, a game on words with “ss” and the surname of the spokesperson) reduce their influence considerably; while other non-spam terms, like “face mask” increase their presence.

Concerning the topics appearing more often in the repeated messages (also shown in table 3), they can also be represented graphically through the word cloud of figure 3 (please note that the relevance ordering in this word cloud is proportional to the ratio deleted vs kept; i.e., terms that appear more in the “removed after filtering” dataset than on the “kept after filtering” dataset have a higher relevance). The most spammed terms stand out clearly in the word cloud: “electromagnetic”, “pollution”, “pseudoctor”, et

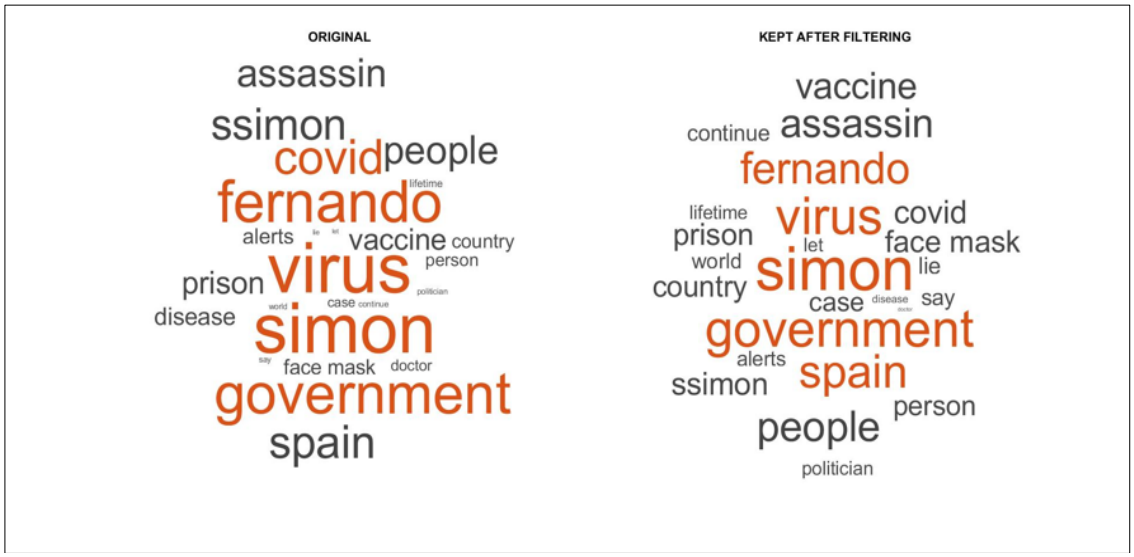


Fig 2. Word clouds: original and spam-filtered

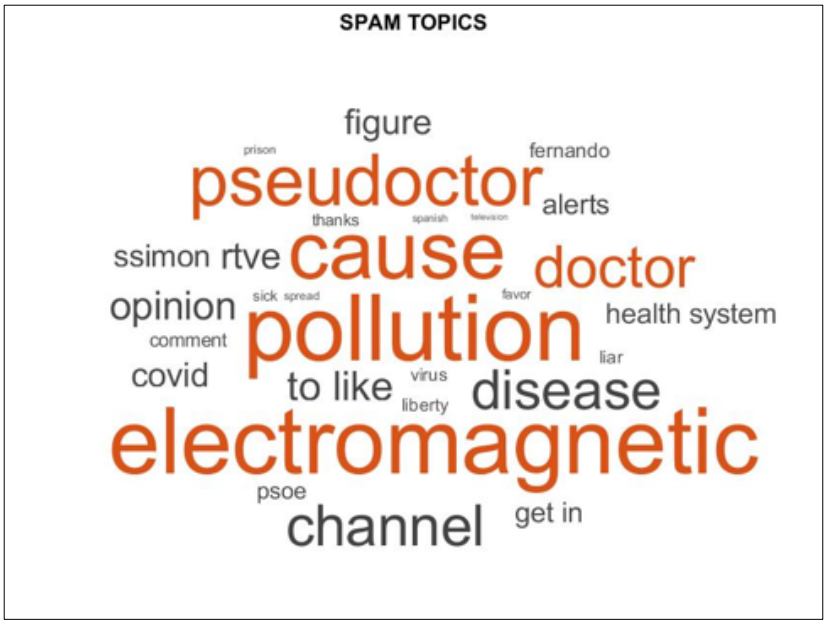


Fig 3. Word cloud of spam terms according to the ratio deleted vs kept

Table II
MOST PROMINENT TERMS:
RELEVANCE BEFORE AND AFTER SPAM-FILTERING

Topics			Original	Kept after filtering	Deleted
Id	**English**	**Spanish**	**n**	**n (%)**	**n**
1	virus	virus	***423***	296 (70.0%)	127 (30.0%)
2	simon	simon	***423***	367 (86.8%)	56 (13.2%)
3	fernando	fernando	***356***	242 (68.0%)	114 (32.0%)
4	govern-ment	gobierno	***351***	298 (84.9%)	53 (15.1%)
5	covid	covid	***303***	175 (57.8%)	128 (42.2%)
6	spain	españa	***289***	260 (90.0%)	29 (10.0%)
7	ssimon	ssimon	***282***	170 (60.3%)	112 (39.7%)
8	people	gente	***259***	233 (90.0%)	26 (10.0%)
9	assassin	asesino	***258***	218 (84.5%)	40 (15.5%)
10	prison	prision	***224***	165 (73.7%)	59 (26.3%)
11	vaccine	vacuna	***216***	199 (92.1%)	17 (7.9%)
12	disease	enfermedad	***191***	75 (39.3%)	116 (60.7%)
13	alerts	alertas	***183***	114 (62.3%)	69 (37.7%)
14	face mask	mascarilla	***178***	161 (90.5%)	17 (9.5%)
15	country	país	***175***	159 (90.9%)	16 (9.1%)
16	person	persona	***169***	147 (87.0%)	22 (13.0%)
17	doctor	doctor	***158***	58 (36.7%)	100 (63.3%)
18	case	caso	***152***	144 (94.7%)	8 (5.3%)
19	lifetime	vida	***140***	106 (75.7%)	34 (24.3%)
20	world	mundo	***133***	119 (89.5%)	14 (10.5%)
21	politician	político	***132***	107 (81.1%)	25 (18.9%)
22	continue	seguir	***132***	122 (92.4%)	10 (7.6%)
23	say	decir	***132***	123 (93.2%)	9 (6.8%)
24	lie	mentira	***130***	129 (99.2%)	1 (0.8%)
25	let	dejar	***127***	107 (84.3%)	20 (15.7%)

Table III
MOST DELETED TERMS IN FILTERING
(ORDERED BY PERCENTAGE OF DELETION)

Topics			Origi-nal	Kept after filtering	Deleted
Id	**English**	**Spanish**	**n**	**n(%)**	**n(%)**
39	pollution	contamina-ción	99	1(1,0%)	***98(99,0%)***
40	electromag-netic	electro-magnético	98	1(1,0%)	***97(99,0%)***
36	cause	causar	111	14(12,6%)	***97(87,4%)***
75	pseudoctor	pseudoctor	67	16(23,9%)	***51(76,1%)***
17	doctor	doctor	158	58(36,7%)	***100(63,3%)***
96	channel	canal	55	21(38,2%)	***34(61,8%)***
12	disease	enfermedad	191	75(39,3%)	***116(60,7%)***
77	to like	gustar	65	34(52,3%)	***31(47,7%)***
108	opinion	opinión	50	27(54,0%)	***23(46,0%)***
81	rtve	rtve	62	35(56,4%)	***27(43,6%)***
99	figure	cifra	54	31(57,4%)	***23(42,6%)***
5	covid	covid	303	175(57,8%)	***128(42,2%)***
7	ssimon	ssimon	282	170(60,3%)	***112(39,7%)***
13	alerts	alertas	183	114(62,3%)	***69(37,7%)***
107	get in	entrar	51	32(62,8%)	***19(37,2%)***
83	health sys-tem	sanidad	62	39(62,9%)	***23(37,1%)***
89	psoe	psoe	58	38(65,5%)	***20(34,5%)***
3	fernando	fernando	356	242(68,0%)	***114(32,0%)***
78	comment	comentario	64	44(68,8%)	***20(31,2%)***
30	liar	mentiroso	121	84(69,4%)	***37(30,6%)***
26	liberty	libertad	125	87(69,6%)	***38(30,4%)***
1	virus	virus	423	296(70,0%)	***127(30,0%)***
73	thanks	gracias	68	48(70,6%)	***20(29,4%)***

D. Analysis of spam messages

Among the detected spam messages, the most prominent ones were selected for a manual classification (as detailed in the methods section).

The results of the classification are show in table 4. Those messages that were classified differently by the coders were accounted for by adding 0.5 to both categories, that is the reason for the fractional numbers in "hate messages" and "satiric messages". As expected, opinions (or neutral messages) accounted only for a 18.6% of all repeated messages. Most repeated messages had political, hate or satiric connotations (71.2% altogether). Fake news represented only 8.5% of all repeated messages. The complete table, with the text of the 59 spam messages (in Spanish and English) and the classification of both coders, can be found in multimedia appendix #2.

The coherence of the classification between both coders was high, as proved by a further Kappa analysis (k=0.785, P<.001). Table 5 shows the confusion matrix obtained, where the most common discrepancies found where: fake instead of neutral (2 discrepancies); and satiric instead of political (2 discrepancies)

Table IV
CLASSIFICATION OF SPAM MESSAGES

Contents of the repeated message	n (%)
Political messages	15 (25.4%)
Hate messages	14.5 (24.6%)
Satiric messages	12.5 (21.2%)
Neutral (opinions)	11 (18.6%)
Fake news	5 (8.5%)
Uncategorized	1 (1.7%)
TOTAL	59 (100%)

Table V
CONFUSION MATRIX FOR THE SPAM CLASSIFICATION OF BOTH CODERS

	fake	hate	satiric	political	neu-tral	unclassi-fied
fake	3	1	0	0	2	0
hate	0	13	1	0	0	1
satiric	0	0	11	1	0	0
political	1	0	1	13	0	0
neutral	0	0	0	1	9	0
unclas-sified	0	0	0	0	1	0

IV. DISCUSSION

A. Principal results

The analysis performed shows that spam (or repeated messages) in YouTube comments can alter substantially the impression given to the reader, increasing artificially the influence of certain terms.

Although social platforms have increased their security measures to offer the most accurate (and spam free) information about COVID-19, different spam techniques are still possible. As stated in the introduction section, YouTube is supposed to be prepared to filter spam messages both manually (the viewers can mark comments as spam) and automatically (comments likely to be spam, according to the text of comments and the behavior of commenters). According to the definition given by YouTube, "spam is content or correspondences that create a negative experience by making it difficult to find more relevant and substantive material. It can sometimes be used to indiscriminately send unsolicited bulk messages to people on YouTube".

It seems that simple repetitions of exactly the same comment may be easily filtered out, but the automatic (or manual) publication of slightly modified versions of the message is still possible. The results obtained in this study show clearly that the comments section of YouTube videos is vulnerable to this kind of spam technique. Novel spam protection measures that successfully limit the repetitions of messages can improve the reader's experience in the comment section (still giving the viewers the freedom to share their opinions). The current state of the art of NLP techniques may serve as a more effective spam detection mechanism.

There are other studies where the incidence of spam and toxic comments is also measured. For example, in [26], Twitter, Facebook and Instagram comments about a public tobacco prevention campaign are analyzed, with 1.61% of posts being considered spam posts, and 12.88% being considered toxic comments. In [27], Facebook and Twitter bots, as well as bot-detection algorithms are analyzed in the context of the misinformation spread about COVID-19 pandemic.

Although this is not an objective of the present study, the review of the comments found revealed much more hate, sarcastic or political messages than opinions or scientific/medical information in these videos. This may be an important source of information about the validity of these media as informative of COVID-19 opinions in Spain, in contrast to studies carried out in different countries, like the one analyzing the daily COVID-19 briefings of the Canadian Prime Minister [28], where opinion messages were the most common comments found.

Standard misinformation classifications consider the following categories: false connection, false context, manipulated content, satire, misleading content, imposter content, and fabricated content [29]. Obviously, these categories are focused on contents, not on comments. For the purposes of this work, a different set of categories, not standard but more adapted to comments, was used: opinions, fake news, hate messages, satiric messages, and politic messages.

The comments section of YouTube videos contains a huge amount of valid information mixed with misinformation. Such misinformation may lead to health risks, particularly in the context of an epidemic like COVID-19.

Misinformation present in the comments section of YouTube may have different impact depending on the particular reader. Susceptibility to misinformation, as analyzed in [30] varies largely depending on numeracy and literacy levels, as well as demographic factors. Interventions to reduce this susceptibility should be encouraged.

V. CONCLUSIONS

Spam is present in YouTube comments and alters substantially the impression given to the reader, increasing artificially the influence of certain terms. Automatic (or manual) publication of slightly modified versions of the same message does not seem to be detected by automatic spam filters. The current state of the art of NLP techniques may serve as a more effective spam detection mechanism. The simple removal of repetitive messages proposed in this paper improves the quality of opinion sharing between readers.

ACKNOWLEDGEMENTS

This study was funded by APOTIP/2021/033 and FEDER-COVID-26 (European Regional Development Fund for the Valencian Community 2014-2020, within the framework of the REACT-EU program, the European Union's response to the COVID-19 pandemic.), both financed by the Conselleria de Educación, Generalitat Valenciana.

MULTIMEDIA APPENDICES

Appendix 1: List of all video URLs, dates and other details used for this study. Available from: https://lcsi.umh.es/spamcovidnews/appendix_1.xlsx

Appendix 2: List of the most repeated messages (messages repeated 3 or more times and containing 10 or more words), both in Spanish and translated to English. Available from: https://lcsi.umh.es/spamcovidnews/appendix_2.xlsx

REFERENCES

[1] Van Scoy LJ, Miller EL, Snyder B, Wasserman E, Chinchilli VM, Zgierska AE, Rabago D, Lennon CL, Lipnick D, Toyobo O, Ruffin MT. Knowledge, perceptions, and preferred information sources related to COVID-19 among central Pennsylvania adults early in the pandemic: a mixed methods cross-sectional survey. *The Annals of Family Medicine*. 2021 Jul 1;19(4):293-301. https://doi.org/10.1370/afm.2674

[2] Sathianathan S, Van Scoy LJ, Sakya SM, Miller E, Snyder B, Wasserman E, Chinchilli VM, Garman J, Lennon RP. Knowledge, perceptions, and preferred information sources related to COVID-19 among healthcare workers: results of a cross sectional survey. *American Journal of Health Promotion*. 2021 Jun;35(5):633-6. https://doi.org/10.1177/0890117120982416

[3] Yan A, Howden K, Mahar AL, Glidden C, Garland SN, Oberoi S. Gender differences in adherence to COVID-19 preventative measures and preferred sources of COVID-19 information among adolescents and young adults with cancer. Cancer Epide-miology. 2022 Jan 6:102098. https://doi.org/10.1016/j.canep.2022.102098

[4] Skarpa PE, Garoufallou E. Information seeking behavior and COVID-19 pandemic: A snapshot of young, middle aged and senior individuals in Greece. International journal of medical informatics. 2021 Jun 1;150:104465. https://doi.org/10.1016/j.ijmedinf.2021.104465

[5] Sakya SM, Scoy LJ, Garman JC, Miller EL, Snyder B, Wasserman E, Chinchilli VM, Lennon RP. The impact of COVID-19-related changes in media consumption on public knowledge: results of a cross-sectional survey of Pennsylvania adults. Current Medical Research and Opinion. 2021 Jun 3;37(6):911-5. https://doi.org/10.1080/03007995.2021.1901679

[6] Nafees N, Khan D. Health information seeking among general public in India during COVID 19 outbreak: Exploring healthcare practices, information needs, preferred information sources and problems. Library Philosophy and Practice. 2020:1-5.

[7] Chauhan P, Ansari MS, Sharma NK. Exploring Information Seeking Behavior of the People during COVID-19 Outbreak in India. Library Philosophy and Practice (ejournal). 2021 Jun 1;5588:1-3.

[8] Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bulletin of the World Health Organization. 2021 Jun 1;99(6):455. https://doi.org/10.2471/BLT.20.276782

[9] Szmuda T, Syed MT, Singh A, Ali S, Özdemir C, Słoniewski P. YouTube as a source of patient information for coronavirus disease (Covid-19): a content-quality and audience engagement analysis. Reviews in Medical Virology. 2020 Sep;30(5):e2132. https://doi.org/10.1002/rmv.2132

[10] D'Souza RS, D'Souza S, Strand N, Anderson A, Vogt MN, Olatoye O. YouTube as a source of medical information on the novel coronavirus 2019 disease (COVID-19) pandemic. Global public health. 2020 Jul 2;15(7):935-42. https://doi.org/10.1080/17441692.2020.1761426

[11] Abdullah AO, Ali MA, Karabatak M, Sengur A. A comparative analysis of common YouTube comment spam filtering tech niques. In2018 6th international symposium on digital forensic and security (ISDFS) 2018 Mar 22 (pp. 1-5). IEEE. https://doi.org/10.1109/ISDFS.2018.8355315

[12] Das RK, Dash SS, Das K, Panda M. Detection of spam in Youtube comments using different classifiers. InAdvanced Compu-ting and Intelligent Engineering 2020 (pp. 201-214). Springer, Singapore. https://doi.org/10.1007/978-981-15-1081-6_17

[13] Abd T, Altabrawee H, Ajmi SQ. YouTube spam comments detection using Artificial Neural Network. Journal of Engineering and Applied Sciences. 2018;13(22):9638-42. http://dx.doi.org/10.36478/jeasci.2018.9638.9642

[14] Oh H. A YouTube Spam Comments Detection Scheme Using Cascaded Ensemble Machine Learning Model. IEEE Access. 2021 Oct 19;9:144121-8. http://dx.doi.org/10.1109/ACCESS.2021.3121508

[15] YouTube Help Center: "Manage spam in comments"; 2022. URL: https://support.google.com/youtube/answer/9482362 [accessed 2022-09-10]

[16] Niemiec E. COVID-19 and misinformation: Is censorship of social media a remedy to the spread of medical misinfor-mation?. EMBO reports. 2020 Nov 5;21(11):e51420. https://doi.org/10.15252/embr.202051420

[17] Play Console Help: "Requirements for coronavirus disease 2019 (COVID-19) apps"; 2022. URL: https://support.google.com/googleplay/android-developer/answer/9889712 [accessed 2022-09-10]

[18] Apple Developer: " Ensuring the Credibility of Health & Safety Information"; 2022. URL: https://developer.apple.com/news/?id=03142020a [accessed 2022-09-10]

[19] Michael A. Alcorn in Medium: "How wrong was Ioannidis?"; 2020. URL: https://michaelaalcorn.medium.com/how-wrong-was-ioannidis-5940e49c9af6 [accessed 2022-09-10]

[20] John P-A- Ioannidis in Statnews.com: "A fiasco in the making? As the coronavirus pandemic takes hold, we are making deci-sions without reliable data"; 2020. URL: https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/ [accessed 2022-09-10]

[21] Python Official Web, 2022. URL: https://www.python.org/ [accessed 2022-09-10]

[22] YouTube Data API: "API Reference", 2022. URL: https://developers.google.com/youtube/v3/docs/ [accessed 2022-09-10]

[23] TextacyNLP, 2022. URL: https://pypi.org/project/textacy/ [accessed 2022-09-10]

[24] Spacy NLP, 2022.URL: https://spacy.io/ [accessed 2022-09-10]

[25] Matlab, 2022. URL: https://es.mathworks.com/products/matlab.html [accessed 2022-09-10]

[26] Majmundar A, Le N, Moran MB, Unger JB, Reuter K. Public response to a social media tobacco prevention campaign: Content analysis. JMIR public health and surveillance. 2020 Dec 7;6(4):e20649. https://doi.org/10.2196/20649

[27] Himelein-Wachowiak M, Giorgi S, Devoto A, Rahman M, Ungar L, Schwartz HA, Epstein DH, Leggio L, Curtis B. Bots and misinformation spread on social media: Implications for COVID-19. Journal of Medical Internet Research. 2021 May 20;23(5):e26933. https://doi.org/10.2196/26933

[28] Zheng C, Xue J, Sun Y, Zhu T. Public Opinions and Concerns Regarding the Canadian Prime Minister's Daily COVID-19 Briefing: Longitudinal Study of YouTube Comments Using Machine Learning Techniques. Journal of medical Internet re-search. 2021;23(2):e23957. https://doi.org/10.2196/23957

[29] Wardle C. 6 types of misinformation circulated this election season. Columbia Journalism Review. 2016 Nov 18;18. Retrieved from URL: https://www.cjr.org/tow_center/6_types_election_fake_news.php [accessed 2022-09-10]

[30] Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman AL, Recchia G, Van Der Bles AM, Van Der Linden S. Susceptibility to misinformation about COVID-19 around the world. Royal Society open science. 2020 Oct 14;7(10):201199

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Herramientas para el Desarrollo de Gemelos Digitales mediante Nube de Puntos 3D orientados a la Agricultura 5.0

Paula Catala-Roman, Jaume Segura-Garcia, Miguel Garcia-Pineda
Departamento de Informàtica, ETSE-UV
Universitat de València
Av. de la Universitat, s/n. Burjassot - Valencia.
paucaro7@uv.es, jaume.segura-garcia@uv.es, miguel.garcia-pineda@uv.es

RESUMEN: Los gemelos digitales son esenciales en la agricultura 5.0 al proporcionar una representación precisa y digital de objetos y procesos agrícolas, permitiendo la toma de decisiones basadas en datos, la simulación de escenarios futuros y la innovación para una agricultura más eficiente y sostenible. El objetivo principal de este artículo es hacer una revisión y comparativa de las principales herramientas para el desarrollo de gemelos digitales orientados a aplicaciones de agricultura 5.0 mediante modelos 3D de nube de puntos creados a partir de técnicas de fotogrametría. Para ello se han presentado las cuatro herramientas más utilizadas para el desarrollo de estos modelos 3D. Se ha realizado una comparativa cualitativa de las características principales de estas herramientas. A continuación, a partir de unas imágenes tomadas en un campo de naranjos se ha realizado un análisis de calidad de los modelos 3D de nube de puntos obtenidos por cada una de las herramientas analizadas. Finalmente, se ha indicado que herramienta puede ser la más interesante según el estudio llevado a cabo.

***Palabras Clave*- digital twin, point clouds, agricultura digital, herramientas**

I. Introducción

La agricultura 5.0 es una visión futurista y avanzada de la agricultura que busca integrar tecnologías de vanguardia y sistemas inteligentes para impulsar la eficiencia, la sostenibilidad y la productividad en el sector agrícola [1]. Se basa en los avances de la agricultura de precisión (agricultura 4.0) y lleva la digitalización y la automatización al siguiente nivel.

La Agricultura 5.0 utiliza tecnologías como la inteligencia artificial (IA), el aprendizaje automático, la robótica, los drones, el Internet de las cosas (IoT) y los sistemas de sensores para recopilar y analizar datos en tiempo real. Estos datos se utilizan para optimizar las operaciones agrícolas, mejorar la toma de decisiones y maximizar los rendimientos de los cultivos. Uno de los aspectos clave de la agricultura 5.0 es la conectividad y la interconexión de diferentes dispositivos y sistemas agrícolas. Los agricultores pueden monitorear y controlar remotamente sus cultivos, animales y equipos a través de plataformas digitales. Esto permite una gestión más eficiente de los recursos, como el agua y los fertilizantes, y ayuda a prevenir enfermedades y plagas al detectarlas de manera temprana.

Por otra parte, hoy en día los gemelos digitales (o *digital twin*) están siendo muy utilizados en diversos campos de nuestras vidas, como industria, sanidad, transporte, logística, etc [2]. Un gemelo digital es una representación o réplica digital precisa de un objeto físico, proceso o servicio. En la agricultura 5.0, el gemelo digital es importante porque permite optimizar la planificación, monitorear y controlar la granja en tiempo real, tomar decisiones basadas en datos, simular escenarios futuros y fomentar la innovación [3]. Proporciona a los agricultores información precisa, les ayuda a anticiparse a problemas, maximizar rendimientos y desarrollar soluciones más eficientes y sostenibles. En resumen, el gemelo digital es clave para impulsar la eficiencia y la productividad en la agricultura 5.0.

El objetivo principal de este artículo es hacer una revisión y comparativa de las principales herramientas para el desarrollo de gemelos digitales de aplicaciones orientadas a agricultura 5.0 mediante nubes de puntos 3D creados a partir de técnicas de fotogrametría.

Para ello, una vez introducidos los conceptos de agricultura 5.0 y gemelos digitales, y propuesto el objetivo principal del artículo, el resto del documento se ha organizado como se va a describir a continuación. La sección 2 presenta un estado del arte sobre gemelos digitales aplicados en agricultura, y a continuación la sección 3, donde se enumeran los posibles beneficios que puede ofrecer este tipo de tecnología para la agricultura

5.0. La sección 4 describe varias herramientas actuales que puede resultar útiles para el desarrollo de este digital twin y lo que nos ofrecen, sobre modelado 3D. En la siguiente sección se han probado y comparado estas herramientas para posteriormente seleccionar la mejor. Por último, en la sección 6 se discutirán las conclusiones a las que se han llegado mediante este trabajo y cuales puedes ser los posibles trabajos futuros.

II. Estado del Arte

La sección del estado del arte que se presenta a continuación está divida en dos subsecciones. La primera de ellas donde se puede ver la importancia de los gemelos digitales en el sector de la agricultura y la segunda, donde se hace un análisis de algunos trabajos sobre el uso de técnicas de fotogrametría para realizar modelos 3D en la agricultura.

A. Gemelos Digitales en la Agricultura 5.0

Para el estado del arte de gemelos digitales en agricultura, se ha realizado una búsqueda literaria en Google Scholar de "digital twins agriculture". Esto nos ha devueltos varios artículos de resultado que nos pueden ser interesantes sobre el tema

Los trabajos [3] y [4] se basan en una búsqueda literaria, donde el primer artículo que encuentran sobre gemelos digitales en el ámbito de la agricultura es de 2017. En ellos vemos varios ejemplos como el de un prototipo de gemelo digital de un invernadero en el cual utilizan información recogida por un robot, por ejemplo sobre los nutrientes del suelo y humedad, y simulaciones del jardín para saber qué debe hacer el robot para garantizar que el cultivo se encuentre en las condiciones ideales para crecer [5], u otro de un gemelo digital de un cultivo de tomates [6] cuyo equipo se encuentra desarrollando un modelo 3D del cultivo al que añadirle información en tiempo real a partir de unos sensores, con el objetivo también de tomar decisiones correctas para que el cultivo real crezca en las mejores condiciones. También vemos un prototipo de un gemelo digital en una granja inteligente en [7], donde llegan a hacer una prueba de su sistema con una planta de ejemplo, como trabajos futuros se indica que dicho sistema quieren aplicarlo a proyectos más grandes, utilizando también IA (Inteligencia Artificial), para permitir un desarrollo sostenible y también mejorar la seguridad alimentaria. Muchos de los trabajos que hablan son conceptos, como el desarrollado en [8], done se muestra la idea de un servicio/producto para agricultura, donde utilizan inteligencia artificial y gemelos digitales para la toma de decisiones para la correcta gestión de un invernadero. El gemelo digital tiene que la función de controlar las condiciones ambientales del invernadero, como por ejemplo la humedad y el nivel de CO_2, así como la posibilidad de controlar las ventanas y ventiladores del propio invernadero.

B. Fotogrametría para desarrollo de modelos 3D en Agricultura 5.0

En la actualidad, los drones se han convertido en una fuente invaluable de datos para actividades como inspección, vigilancia, cartografía y modelado 3D [9]. Mediante un proceso fotogramétrico convencional, es posible obtener resultados tridimensionales, como modelos de superficie (DSM/DTM), contornos de elevación, modelos texturizados en 3D y datos vectoriales, entre otros, de manera considerablemente automatizada.

El uso de la fotogrametría permite el desarrollo de mapas de nube de puntos 3D, a través de esta técnica en [10] se presenta un método no supervisado de detección de viñedos y extracción de características a partir de mapas 3D de nubes de puntos. El método propuesto permite generar automáticamente mapas de las regiones de terreno cubiertas por viñedos y, además, proporciona información sobre la orientación local de las hileras de viñas y del espaciado entre hileras, organizada espacialmente en mapas. Otro ejemplo es [11], donde se presenta un enfoque alternativo, para estimar el momento correcto de la cosecha del maíz; el método propuesto se centra en la relación entre los datos de madurez obtenidos mediante fotogrametría y los parámetros producidos por el análisis químico del maíz. Otro trabajo donde se desarrolla un enfoque fotogramétrico es [12]. En él se desarrolla un sistema automático para medir la rugosidad de la superficie del suelo a partir de imágenes tomadas sobre el terreno con una simple cámara digital, sin restricciones geométricas. La precisión del sistema se determinó en modelos artificiales construidos con poliestireno, cuyas precisiones de posición y elevación fueron de aproximadamente 1,5 mm, mientras que el error en la estimación de la superficie fue inferior al 0,76% de la superficie del sitio. Estos resultados muestran que dos índices de rugosidad, el índice de tortuosidad de la superficie y el valor medio de la altura, son los más eficaces para discriminar los niveles de laboreo del suelo agrícola.

La mayoría de los trabajos encontrados son conceptuales, aunque hay algún prototipo que se ha llegado a probar. Durante toda la búsqueda se ha podido observar cómo ninguno de los artículos habla detalladamente sobre las herramientas/tecnologías que pretenden utilizar a la hora de desarrollar el gemelo digital a partir de técnicas fotogramétricas. A partir de la escasez de este tipo de estudios, surge la idea de desarrollar un artículo en el que ofrecer una comparación de varias de estas herramientas útiles para la creación de gemelos digitales mediante nube de puntos 3D obtenidos a partir de técnicas fotogramétricas y que sean orientados a la agricultura digital.

III. La Fotogrametría para generar Gemelos Digitales en Agricultura 5.0

La fotogrametría es una técnica que se utiliza para medir y crear modelos tridimensionales de objetos y entornos mediante el análisis de imágenes. Consiste en el proceso de capturar y analizar fotografías desde diferentes ángulos y utilizar la información visual y geométrica contenida en esas imágenes para reconstruir la forma, la posición y la escala de los objetos en un espacio tridimensional. La fotogrametría se basa en el principio de la triangulación, que implica encontrar puntos comunes en diferentes imágenes y calcular sus posiciones en el espacio 3D utilizando la geometría y la paralaje de las imágenes capturadas desde diferentes perspectivas. Estos puntos

comunes se denominan puntos de control y pueden ser puntos visuales identificables en las imágenes o marcadores artificiales colocados en los objetos. El proceso de fotogrametría generalmente implica varias etapas, que pueden incluir:

1. Adquisición de imágenes: Se capturan fotografías del objeto o área de interés desde diferentes ángulos y posiciones. Es importante tener una buena cobertura de imágenes para obtener una reconstrucción precisa.
2. Orientación de imágenes: Se determina la posición y la orientación relativa de cada imagen en relación con las demás. Esto se logra identificando puntos de control comunes en las imágenes y utilizando técnicas de emparejamiento y ajuste para estimar los parámetros de orientación.
3. Extracción de puntos característicos: Se identifican puntos clave y características distintivas en las imágenes. Estos puntos se utilizan para realizar un seguimiento y establecer correspondencias entre las diferentes imágenes.
4. Emparejamiento y triangulación: Se establecen correspondencias entre los puntos característicos en las diferentes imágenes y se realiza la triangulación para calcular las posiciones 3D de los puntos en el espacio.
5. Generación del modelo 3D: Se crea un modelo tridimensional a partir de los puntos 3D calculados. Dependiendo de la precisión requerida y el nivel de detalle deseado, se pueden generar modelos de malla poligonal, nubes de puntos densas o modelos superficiales.

En el contexto de la agricultura, la fotogrametría se utiliza para generar gemelos digitales, que son representaciones virtuales precisas de cultivos, terrenos y estructuras agrícolas. Estos gemelos digitales proporcionan información detallada y en tiempo real sobre el estado de los cultivos y el entorno agrícola, lo que ayuda a los agricultores en la toma de decisiones y en la optimización de sus prácticas agrícolas. El uso de la fotogrametría y los gemelos digitales en la agricultura brinda una serie de beneficios, tales como:

- Monitoreo de cultivos: La generación de gemelos digitales a partir de imágenes permite a los agricultores obtener información actualizada sobre el crecimiento y el estado de los cultivos. Pueden detectar áreas con estrés hídrico, enfermedades o plagas, y tomar medidas preventivas o correctivas de manera oportuna.
- Planificación y gestión de terrenos: Los gemelos digitales facilitan la planificación de siembras, el diseño de sistemas de riego y el análisis de las características del terreno. Los agricultores pueden evaluar la topografía, la pendiente, la exposición solar y otros factores para tomar decisiones informadas sobre la preparación del terreno y la distribución de cultivos.
- Optimización de insumos agrícolas: Mediante la generación de gemelos digitales, los agricultores pueden realizar un monitoreo preciso de la salud de los cultivos y aplicar insumos agrícolas (como fertilizantes o pesticidas) de manera localizada y específica. Esto reduce el desperdicio de insumos y minimiza el impacto ambiental.
- Análisis de rendimiento y pronóstico de cosechas: Los gemelos digitales permiten llevar un seguimiento del crecimiento y el rendimiento de los cultivos a lo largo del tiempo. Los agricultores pueden evaluar la salud de las plantas, estimar el rendimiento esperado y planificar la cosecha de manera más eficiente.

Finalmente, el gemelo digital es una herramienta que impulsa la innovación en la agricultura. Permite a los agricultores probar nuevas tecnologías, prácticas y modelos de negocio en el entorno virtual antes de implementarlos en la realidad. Esto fomenta la experimentación y el desarrollo de soluciones más eficientes, sostenibles y rentables.

IV. Herramientas para el Desarrollo de Gemelos Digitales en Agricultura 5.0

Para esta sección se ha realizado una búsqueda de las herramientas más comunes para el desarrollo de modelos 3D a partir de fotogrametría, que posteriormente pueden aplicados a gemelos digitales para la agricultura.

A. Agisoft Metashape

Agisoft Metashape[1], permite generar modelos 3D de alta calidad a partir de fotografías, utilizando algoritmos avanzados de estructura desde el movimiento y coincidencia de puntos densos (ver figura 1). Con esta herramienta, se pueden alinear imágenes, generar nubes de puntos densas, crear mallas poligonales y aplicar texturas realistas. Agisoft Metashape es utilizado en diversas industrias, como la arquitectura, la arqueología y la agricultura, para obtener modelos precisos y visualmente atractivos.

B. Pix4d Mapper

Pix4d Mapper[2] nos ofrece un paquete software especializado en fotogrametría (ver figura 2). Esta herramienta nos permite generar modelos 3D y mapas a partir de imágenes hechas por cámaras de drones, y entre sus principales aplicaciones se encuentran la agricultura de precisión y la topografía, ya que permite generar mapas de índices de vegetación para asistir en la gestión agrícola, y modelos 3D detallados de terrenos. Con Pix4d se pueden nivel y suavizar superficies digitales, así como clasificar automáticamente nubes de puntos.

[1] Agisoft Metashape. Disponible en: https://www.agisoft.com/
[2] Pix3dMapper. Disponible en: https://www.pix4d.com/es/producto/pix4dmapper-fotogrametria-software/

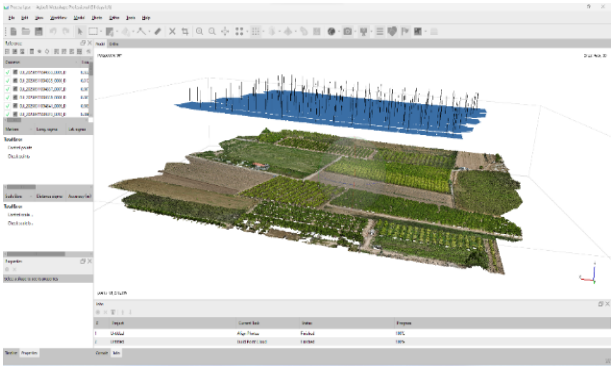


Fig. 1. Modelo 3D con Agisoft Metashape.

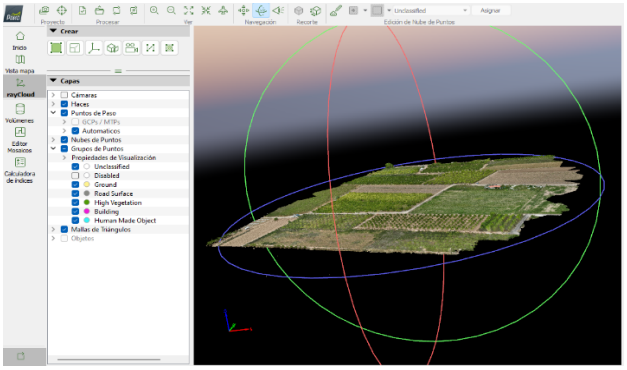


Fig. 2. Modelo 3D con Pix4d Mapper.

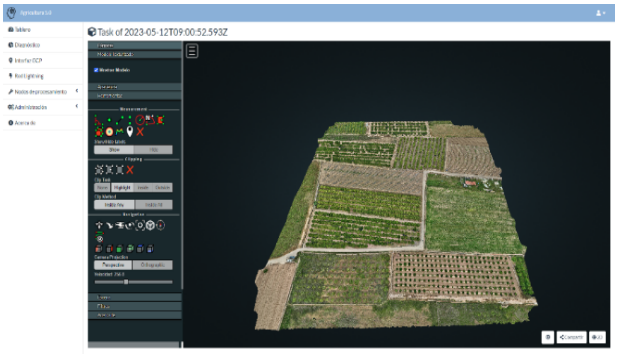


Fig. 3. Modelo 3D con OpenDroneMap.

Fig. 4. Modelo 3D con DJITerra.

C. *OpenDroneMap*

OpenDroneMap[3] es un software de código abierto para el procesamiento de imágenes captadas por dron y generación de modelos 3D, mapas y ortofotos (ver figura 3). Entre las diferentes aplicaciones de ODM (OpenDroneMap), se encuentra la agricultura de precisión, ya que, además de la creación de nubes de puntos densas y de alta resolución de cultivos, también ofrece un paquete en R que permite obtener información sobre cómo están estos cultivos (índice de vegetación, detección de mal estado de las plantas, etc).

D. *DJI Terra*

DJI Terra[4] es una plataforma desarrollada por DJI, permite la construcción de modelos 3D a partir de fotogrametría e imágenes de drones, permitiendo transformar lugares físicos en digitales (ver figura 4). A parte también ofrece una planificación detalla de misiones para vuelos automáticos, enlazándolo con la función de modelado 3D. Entre sus aplicaciones se encuentra el mapeo y la topografía, la agricultura de precisión, y gestión de desastres/respuesta a emergencias.

V. Comparativa de Herramientas

En esta sección vamos a realizar una comparativa de las herramientas presentadas en la sección anterior. Para ello, primero mostraremos un análisis cualitativo de las cuatro herramientas y posteriormente analizaremos la calidad de los modelos 3D obtenidos a partir de las técnicas de fotogrametría implementadas en cada software.

A. *Análisis cualitativo*

En la tabla I, se muestra el análisis cualitativo llevado a cabo. El primer elemento analizado son los requisitos hardware. En este ítem el software DJI Terra es él que requiere más memoria RAM, mientras de ODM y Pix4d Mapper son los que menos. Este aspecto es interesante tenerlo en cuenta, por ejemplo, si alguno de estos servicios quisiera implantarse en la nube. De las herramientas analizadas, el único que posee una licencia libre es ODM. El resto de software posee licencias de pago con un coste superior a los 3000€.

Por lo que se refiere al procesado de imágenes multiespectrales, georreferenciación y medidas de distancias, todas las herramientas analizadas tienen la capacidad de realizar dichas tareas. En el caso de presentar modelos de elevación tanto ODM como Agisoft Metashape pueden crear dichos modelos a partir de las imágenes tomadas por un dron, en cambio DJI Terra y Pix4d Mapper no tiene disponible dicha información de manera clara. Otro aspecto interesante es el ofrecimiento de una API para el desarrollo de nuevas herramientas que hagan uso de estas aplicaciones. En este caso todas las herramientas, excepto DJI Terra, ofrecen esta posibilidad. Finalmente, la última característica analizada es la posibilidad de cómputo distribuido, en este caso tanto ODM como Agisoft Metashape permiten esta opción.

[3] OpenDroneMap. Disponible en: https://opendronemap.org/
[4] DJI Terra. Disponible en: https://enterprise.dji.com/es/dji-terra

Tabla I. Análisis cualitativo de las herramientas analizadas. N/D (No disponible).

	Agisoft Metashape	Pix4d Mapper	OpenDroneMap	DJI Terra
Requisitos hardware	Mínimos: -entre 16 y 32 GB RAM -CPU: Procesador Intel o AMD de 4-8 núcleos -GPU: NVIDIA o AMD con más de 700 núcleos	Mínimos: -Entre 4 y 16 GB RAM Recomendados: -Entre 16 y 32 GB RAM -Disco duro SS, entre 15 y 120 GB espacio libre	Mínimos: -CPU 64 bit -20GB espacio disco -4GB RAM Recomendados: -CPU última generación -100GB espacio disco -16GB RAM	Mínimos: -tarjeta gráfica NVIDIA -32GB RAM - SO Windows 7 64-bit o más nuevos Recomendados: -Tarjeta gráfica NVIDIA GeForce GTX 2070 o por arriba
Software libre	✗	✗	✓	✗
API	✓	✓	✓	✗
Procesado de imagen multiespectral	✓	✓	✓	✓
Georreferenciación	✓	✓	✓	✓
Modelos de elevación	✓	N/D	✓	N/D
Medidas de distancias	✓	✓	✓	✓
Procesamiento distribuido	✓	✗	✓	✗

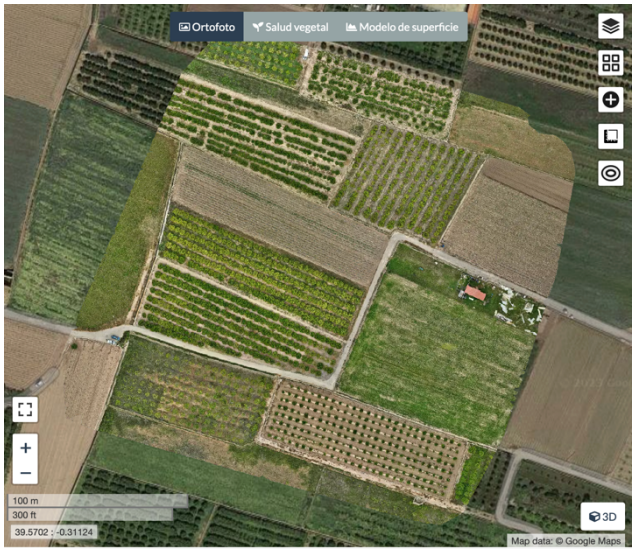


Fig. 5. Ortofoto generada por ODM del vuelo de prueba llevado a cabo en mayo de 2013 en El Puig, València.

B. Análisis cuantitativo

Para realizar el análisis cuantitativo de las herramientas comentadas a lo largo del artículo hemos creado un modelo 3D con nube de puntos a partir de las imágenes captadas por el dron DJI Mavic 3M, en unos campos de El Puig, València. Para generar esta nube de puntos hemos partido de un vuelo que posee 142 imágenes, en la figura 5 puede verse la ortofoto del vuelo realizado.

El primer parámetro que se ha analizado es el tiempo de proceso de la fotogrametría para crear un modelo 3D de nube de puntos a partir de las 142 imágenes. Estos tiempos dependerán de las características de la máquina donde estén instaladas las herramientas.

En nuestro caso, todas ellas han sido instaladas sobre la misma máquina. Las características de esta son: CPU 12th Gen Intel(R) Core (TM) i7-12700F @2.10 GHz, una memoria de caché de 25 MB, 2x16 GB DIMM DDR4 3200 RAM, un disco WD Blue SN570 SSD 1TB M.2 NVMe, Sistema operativo Windows 11 x64 bits.

En cuanto al tiempo que utiliza cada una para generar la nube de puntos, las más rápidas han sido DJI Terra y ODM, con 14 y 25 minutos respectivamente. Agisoft Metashape y Pix4d Mapper han tardado unos 40 minutos.

Para realizar el estudio de calidad, cada modelo generado por cada herramienta analizada se ha guardado en formato ".ply". Estos modelos 3D de nube puntos han sido evaluados mediante la herramienta NR-3DQA [13] disponible en Github[5].

De las diversas métricas que extrae la herramienta NR-3DQA, hemos seleccionado varias características geométricas que pueden ser útiles para evaluar la calidad de las nubes de puntos:

- Curvatura: La curvatura es la cantidad en la que una curva se desvía de ser una línea recta, lo que suele utilizarse para describir la rugosidad o la suavidad. Para un mismo modelo, valores altos mejor calidad.
- Anisotropía: La anisotropía se utiliza para mostrar variaciones en las propiedades geométricas en diferentes direcciones. Para un mismo modelo, valores bajos mejor calidad.
- Linealidad: La linealidad es la propiedad para estimar la similitud con una línea recta. Para un mismo modelo, valores altos mayor calidad.
- Planitud: La planitud se utiliza para medir la similitud con una superficie plana. Para un mismo modelo, valores bajos mayor calidad.
- Esfericidad: La esfericidad es la medida de cuánto se parece la forma de un objeto a la de una esfera perfecta. Para un mismo modelo, valores altos mayor calidad.

En la figura 6, se presentan los valores de curvatura, anisotropía, linealidad, planitud y esfericidad para cada modelo 3D obtenido de las herramientas analizadas en este trabajo.

[5] NR-3DQA. Disponible en: https://github.com/zzc-1998/NR-3DQA

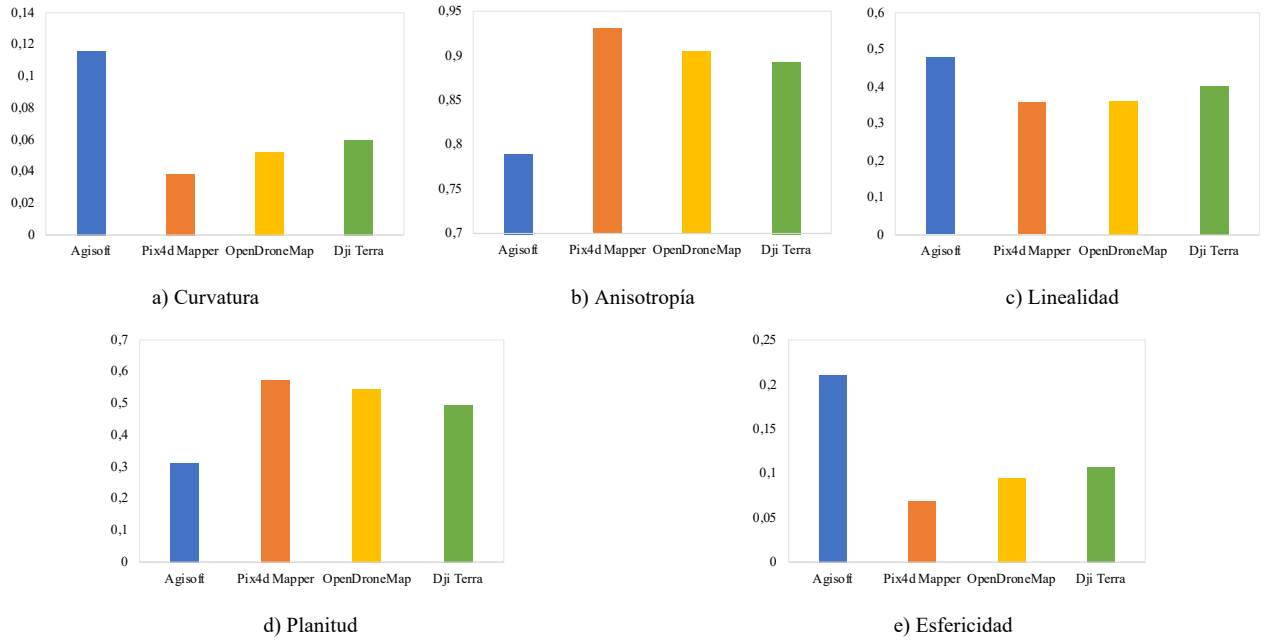


Fig. 6. Valores de las características geométricas obtenidas del análisis de calidad.

Según lo indicado en la explicación de las características podemos ver que la herramienta Agisoft Metashape es el que más calidad aporta, seguidos de DJI Terra y ODM (muy igualados), y finalmente Pix4d Mapper que es el que peor resultado ha obtenido de esta prueba.

VI. Conclusiones

Como hemos visto existen muchos beneficios al incorporar gemelos digitales para agricultura. En este artículo se ha realizado un estudio acerca de los beneficios del uso de la fotogrametría en la creación de gemelos digitales para el ámbito de la agricultura. Posteriormente se han realizado varios análisis cualitativo y cuantitativo de las cuatro herramientas más utilizadas para implementar estas técnicas fotogramétricas y crear modelos 3D.

Haciendo una valoración global del estudio llevado a cabo podemos indicar que la herramienta ODM es la mejor relación calidad/características. Cabe destacar que presenta una calidad buena y se trata de una herramienta "open source" que mejora día a día. Además, sus modelos requieren menos tiempo de computación y las características que ofrece la herramienta son muy adecuadas para adaptarla a diversos tipos de proyectos.

Agradecimientos

La publicación es parte del proyecto Agriculture6.0 con referencia TED2021-131040B-C33, financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea "NextGenerationEU"/PRTR.

Referencias

[1] L. Ahmad y F. Nabi, Agriculture 5.0: Artificial Intelligence, IoT and Machine Learning, Boca Raton: CRC Press, 2021.

[2] B. R. Barricelli, E. Casiraghi y D. Fogli, «A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications,» *IEEE Access,* vol. 7, pp. 167653-167671, 2019.

[3] C. Pylianidis, S. Osinga y I. N. Athanasiadis, «Introducing digital twins to agriculture,» *Computers and Electronics in Agriculture,* vol. 184, p. 105942, 2021.

[4] W. Purcell y T. Neubauer, «Digital Twins in Agriculture: A State-of-the-art review,» *Smart Agricultural Technology, vol. 3, p. 100094,* 2023.

[5] A. Barnard, «In the digital indoor garden,» 2019. [En línea]. Available: https://www.siemens.com/global/en/company/stories/research-technologies/digitaltwin/digital-indoor-garden.html.

[6] W. U. &. Research, 2020. [En línea]. Available: https://www.wur.nl/en/newsarticle/wur-is-working-on-digital-twins-for-tomatoes-food-and-farming.htm.

[7] R. Gomes Alves, G. Souza, R. Filev Maia, A. Lan Ho Tran, C. Kamienski, J.-P. Soininen, P. T. Aquino-Jr y F. Lima, «A digital twin for smart farming,» de *IEEE Global Humanitarian Technology Conference*, 2019.

[8] «Digital Twin Solutions for Smart Farming,» 2019. [En línea]. Available: https://www.rdworldonline.com/rd-100-2019-winner/digital-twin-solutions-for-smart-farming/.

[9] G. J. Grenzdörffer, A. Engel y B. Teichert, «The photogrammetric potential of low-cost UAVs in forestry and agriculture,» *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,,* vol. 31, nº B3, pp. 1207-1214, 2008.

[10] L. Comba, A. Biglia, D. R. Aimonino y P. Gay, «Unsupervised detection of vineyards by 3D point-cloud UAV photogrammetry for precision agriculture,» *Computers and Electronics in Agriculture,* vol. 155, pp. 84-95, 2018.

[11] J. Janoušek, V. Jambor, P. Marcoň, P. Dohnal, H. Synková y P. Fiala, «Using UAV-Based Photogrammetry to Obtain Correlation between the Vegetation Indices and Chemical Analysis of Agricultural Crops,» *Remote Sensing,* vol. 13, nº 10, p. 1878, 2021.

[12] J. Gilliot, E. Vaudour y J. Michelin, «Soil surface roughness measurement: A new fully automatic photogrammetric approach applied to agricultural bare fields,» *Computers and Electronics in Agriculture,* vol. 134, pp. 63-78, 2017.

[13] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu y G. Zhai, «o-Reference Quality Assessment for 3D Colored Point Cloud and Mesh Models,» *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 32, nº 11, pp. 7618-7631, 2022.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Estrategia local vs. remota para el desarrollo de pacientes virtuales conversacionales

César Fernández[1], M. Asunción Vicente[1], Susana Lorenzo[2], Adriana López[3], M. Concepción Carratala[3], José Joaquín Mira[4,5].

[1]Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche, Universidad Miguel Hernández (UMH)
[2]Hospital Universitario Fundación Alcorcón (HUFA)
[3]Departamento de Medicina Clínica, Universidad Miguel Hernández (UMH)
[4]Departamento de Psicologia de la Salud, Universidad Miguel Hernández (UMH)
[5]Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO)

1,3,4 (UMH), Avenida de la Universidad s/n Elche (03202) Alicante, España
2 (HUFA), Hospital Universitario Fundación Alcorcón, Calle Budapest, 1 Alcorcón (28922) Madrid, España
5 (FISABIO), Hospital General Universitario, Ed. Anexo II, 3ª planta, Camí de L´Almassera 11, Elche (03203), Alicante, España

c.fernandez@umh.es, suni@umh.es, slorenzom@salud.madrid.org, adriana.lopezp@umh.es, maria.carratala@umh.es, jose.mira@umh.es

Los pacientes virtuales conversacionales son la mejor herramienta para la práctica de la consulta clínica en titulaciones de medicina. Las plataformas de aprendizaje basadas en pacientes virtuales pueden crearse con herramientas específicas o genéricas (modelado 3D, chatbots, etc.). En todos los casos, su funcionamiento en la parte conversacional se basa en el procesamiento remoto. Esto supone retrasos que afectan a la naturalidad de la conversación entre estudiante y paciente virtual. Proponemos una plataforma con procesamiento 100% local y basada en software libre (*CMU-Sphinx* para reconocer la voz del estudiante; librerías *MLKit* de Google para el procesamiento del lenguaje natural; *Blender* más *MPFB2* para la generación de personajes; y *StyleTalk* para la gesticulación y el movimiento de los labios de los avatares mientras hablan). Como trabajo futuro, establecemos un protocolo de pruebas que permita comparar esta plataforma con los mejores sistemas comerciales actuales.

***Palabras Clave*- paciente virtual, reconocimiento de voz, PLN, avatar 3D, procesamiento local, procesamiento remoto.**

I. Introducción

En todos los campos de las ciencias de la salud, la formación de futuros profesionales incluye la realización de prácticas en un entorno lo más cercano posible a la realidad. En los últimos años, se han desarrollado herramientas conocidas como *pacientes virtuales* que simulan a pacientes reales, facilitando en gran medida la realización de estas prácticas. La evolución de los pacientes virtuales ha estado ligada al aumento del realismo en la simulación de pacientes reales, desde los primeros desarrollos en este campo [1][2]; hasta los trabajos más recientes, como los presentados en [3] o [4].

En todos los casos, un paciente virtual es un programa de ordenador que simula a un paciente con síntomas concretos. En algunos casos, el objetivo para los estudiantes es aplicar tratamientos al paciente y comprobar su evolución [5]. En otros casos, el objetivo es entrevistar al paciente para practicar técnicas de consulta clínica [6]. También existen pacientes virtuales para cirugía, en los que, a través de entornos de realidad virtual con realimentación háptica de contactos, los estudiantes pueden practicar técnicas quirúrgicas [7]. En todos los casos, el realismo es fundamental para un aprendizaje correcto.

Un paciente virtual se puede desarrollar desde cero, utilizando herramientas genéricas, como *Unity* [8] para las simulaciones tridimensionales o *ChatScript* [9] para los módulos conversacionales; pero existe multitud de software específico que facilita en gran medida la creación de pacientes virtuales. Entre este software, cabe mencionar el entorno *Web-SP*, cuya primera versión se presentó en 2003 con el nombre *WASP* [1], o el entorno *VIC* (*Virtual Interactive Case System*) de la Universidad de Toronto [10]. También existen productos comerciales, como los ofrecidos por la compañía *Laerdal* [11], el software *Full Code Medical Simulation* [12], o el software *Body Interact*

[13]. En la Fig. 1 se muestra un ejemplo de las capacidades de uno de estos productos.

Fig. 1: Ejemplo de software de pacientes virtuales, con avatar y pruebas médicas que el estudiante puede solicitar. Imagen obtenida de [13].

En este artículo, nuestro trabajo se centra en la parte conversacional de los pacientes virtuales, que simula una consulta clínica. Básicamente, el paciente virtual está a la espera de las preguntas del estudiante, y responde de acuerdo con un programa predefinido en función de sus síntomas y de la dolencia que padece. Se trata de un entrenamiento práctico que es difícil implementar sin pacientes virtuales. La única alternativa es la utilización de actores, que deben estar perfectamente entrenados para responder ante cualquier pregunta de forma coherente con su dolencia y sus síntomas. Son los llamados pacientes simulados (SP o *simulated patients* [6]). Se trata de una alternativa muy costosa que habitualmente sólo se utiliza en los exámenes finales de capacitación, denominados ECOE (Examen Clínico Objetivo Estructurado) [14].

Los pacientes virtuales conversacionales más comunes no permiten una interacción en lenguaje natural, sino que muestran un listado de posibles preguntas para que el estudiante elija entre ellas [15]. Posteriormente, se puede analizar si el estudiante ha realizado las preguntas adecuadas, en el orden correcto, o por el contrario si ha olvidado preguntas necesarias o ha realizado preguntas irrelevantes. Un ejemplo se muestra en la Fig. 2.

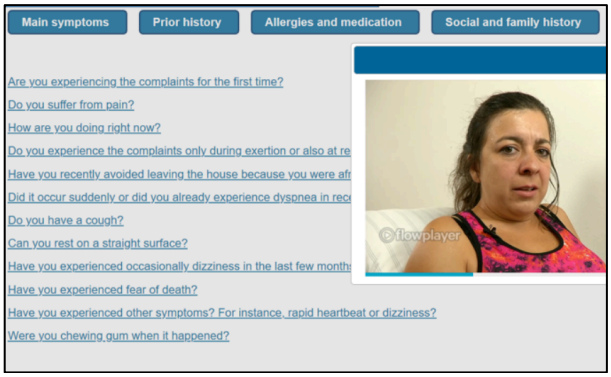


Fig. 2: Ejemplo de módulo conversacional basado en listados de posibles preguntas al paciente. Imagen obtenida de [15].

Los pacientes virtuales más recientes, para aumentar el realismo en la interacción, sí permiten una interacción en lenguaje natural, gracias a los avances en las técnicas de procesamiento del lenguaje natural (PLN). Además, la comunicación puede realizarse indistintamente mediante teclado o mediante voz. Esto supone un mayor realismo (no hay listas de preguntas posibles) y por tanto un mejor entrenamiento para el ejercicio real de la profesión.

La reciente aparición de *ChatGPT* [16] y otros chatbots avanzados permitiría, en teoría, crear pacientes virtuales de un modo 100% automático. Bastaría con dar instrucciones del tipo "Eres una persona de xxx años con la enfermedad yyy; ¿qué contestarías si te preguntan zzz?". El problema es la falta de garantías de que las respuestas ofrecidas sean exactamente las que debería ofrecer. Un ejemplo y un análisis se pueden encontrar en el trabajo presentado en [17]. Esta falta de garantías hace que esta posibilidad no sea adecuada por ahora, y sea más razonable una programación determinista y supervisada de las respuestas a ofrecer ante cualquier pregunta.

Teniendo en cuenta los antecedentes anteriores, en este trabajo proponemos la evaluación de dos alternativas para la creación de módulos conversacionales en lenguaje natural con pacientes virtuales: en primer lugar, la estrategia habitual en todas las plataformas actuales: procesamiento remoto (tanto para el reconocimiento de voz, como para las técnicas de PLN, la obtención de respuestas y la simulación de avatares 3D reproduciendo esas respuestas). En segundo lugar, una nueva estrategia con procesamiento 100% local, que no dependa de las posibles fluctuaciones de la conexión a internet y que se pueda implementar completamente mediante herramientas gratuitas o de código abierto (*open source*). Para esta segunda estrategia, analizaremos diferentes posibilidades para la comunicación por voz y para la generación de los avatares 3D realistas.

II. Alternativas para la comunicación por voz sin procesamiento remoto

El primer paso para la comunicación por voz es el reconocimiento de voz o *speech to text*, que generalmente se realiza online en las principales plataformas. Entre las pocas alternativas que se pueden implementar localmente, la más desarrollada es *CMU-Sphinx* [18], integrable tanto en aplicaciones para dispositivos Android e iOS, como en aplicaciones para ordenador (Windows, Mac o Linux). Una vez reconocida la voz, el procesamiento del texto resultante mediante PLN para generar la respuesta se puede realizar con múltiples herramientas [19][20]. El paso final, convertir la respuesta de nuevo en voz (*text to speech*) está implementado en todas las plataformas en modo local y no plantea dificultades.

En concreto, para nuestras pruebas desarrollaremos una app específica para el sistema operativo Android basada en la librería *CMU-Sphinx* [18] y el módulo PLN de las librerías *ML-Kit* de Google [19]. Aunque las pruebas se realicen en entorno Android, las herramientas utilizadas también están disponibles para iOS.

III. Alternativas para la representación de rostros animados

El objetivo es generar un video animado de un personaje reproduciendo un texto. Si deseamos realizar este proceso localmente, caben dos opciones: en primer lugar, es posible generar el vídeo sobre la marcha, adaptando los movimientos de los labios del avatar al texto que se reproduce. La mayor parte de las plataformas online utilizan esta técnica [21][22]. La segunda opción, menos costosa computacionalmente, es disponer de vídeos pregrabados con cada una de las frases, más un vídeo animado (con ligeros movimientos y cambios de gesto) para los periodos de pausa. Dada la aplicación deseada (paciente virtual), esta segunda opción es razonable, dado que el número de frases diferentes que deberá pronunciar el avatar será bastante reducido en todos los casos. Un inconveniente es la posible aparición de discontinuidades al enlazar varios vídeos: tras el vídeo de una respuesta del paciente, debe reproducirse un video de espera mientras el estudiante realiza la nueva pregunta y posteriormente el video de la siguiente respuesta, y así sucesivamente. En cada uno de los cambios pueden aparecer saltos, que se intentarán minimizar estandarizando los gestos de comienzo y finalización de cada respuesta. Las herramientas disponibles para la generación de estos vídeos son múltiples. Normalmente, el primer paso es la creación de un modelo 3D, para lo cual las alternativas actuales incluyen *Character Creator* y *HeadShot* de la empresa *Reallusion* [23] o también alternativas de código libre como *Blender* [24] y su add-on MPFB2 [25] especializado en el modelado de rostros humanos. El segundo paso, una vez disponible el modelo 3D, es la animación de éste para la generación de vídeos (ya sea una secuencia de reposo o una secuencia de pronunciación de fases). Para este segundo paso, se pueden usar diferentes softwares, como *Synthesia* [26] o *HuggingFace* [27], en los cuales los avatares están predeterminados y no se requiere generación previa; o *D-ID* [28], que sí permite importar avatares personalizados; o también los plugin *NaturalFront* de *Unity* [29] o *MetaHuman* de *Unreal Engine* [30]. Entre las alternativas de código libre, cabe destacar *StyleTalk*, recientemente presentado en [31], o *GeneFace*, descrito en [32].

Para nuestros experimentos, optaremos por la grabación previa de vídeos (que reduce los recursos necesarios en el dispositivo del usuario) y utilizaremos la plataforma *StyleTalk* [31] para la generación de vídeos y el software *Blender* [24] con el plugin *MPFB2* [25] para la creación de los personajes.

IV. Procedimiento de pruebas

Las pruebas buscarán comparar el funcionamiento de nuestra alternativa, basada en herramientas de código libre y sin procesamiento remoto, frente a una alternativa estándar de máxima calidad en cuanto a realismo de video y audio.

De acuerdo con lo comentado en los apartados previos, nuestra alternativa seguirá la estructura indicada en la Fig. 3. Desde el punto de vista del estudiante, la reacción del paciente virtual aparecerá como un vídeo continuo, que en los momentos de las respuestas gesticulará y moverá los labios en sincronismo con el audio; y en los momentos de pausa (por ejemplo, mientras el estudiante realiza sus preguntas), continuará gesticulando y moviéndose ligeramente. El objetivo es conseguir que los cambios de un vídeo a otro pasen inadvertidos para el estudiante, de modo que la conversación con el paciente parezca natural.

En cuanto a la alternativa de máxima calidad frente a la que realizaremos las pruebas, hemos seleccionado el software comercial *SoulMachines* [21]. Es una de las herramientas de más realistas para la interacción por voz con avatares 3D, y su programación es sencilla. Una imagen de esta aplicación se muestra en la Fig. 4.

El procedimiento de comparación consistirá en la generación de un caso clínico simple, con un conjunto muy limitado de preguntas. Las preguntas se programarán de la misma manera tanto en el entorno comercial *SoulMachines* como en el entorno propuesto en este artículo; garantizando que los comportamientos en términos de comprensión de las preguntas sean idénticos en ambas plataformas. Nuestro objetivo principal es comparar la naturalidad y el realismo de las conversaciones.

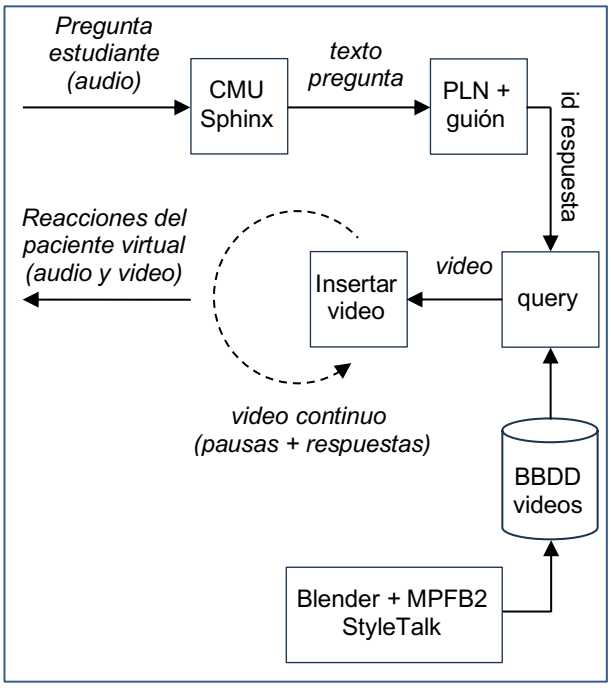


Fig. 3. Estructura software de nuestra propuesta para los pacientes virtuales, con procesamiento 100% local.

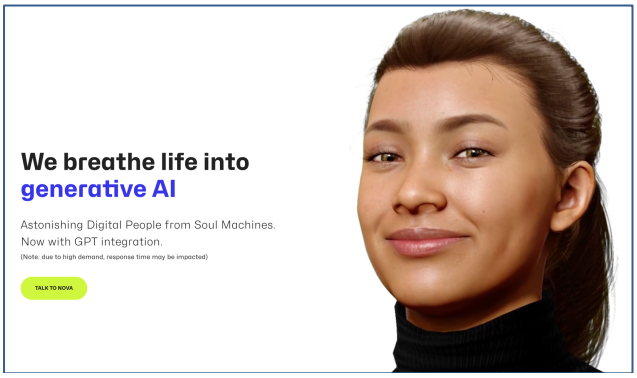


Fig. 4. Herramienta *SoulMachines* a utilizar para la comparación. Imagen obtenida de [21].

Ambas plataformas serán evaluadas mediante un procedimiento de doble ciego por un grupo de profesionales y estudiantes de medicina. Los evaluadores valorarán las plataformas sin conocer qué plataforma específica están utilizando en cada caso. Se realizarán pruebas adicionales en las que se simularán fallos puntuales de la conexión a internet (en la plataforma *SoulMachines*) para evaluar la pérdida de naturalidad a causa de estos problemas, y estimar su probabilidad de aparición en función de la calidad de la conexión a internet.

V. Conclusiones

La práctica de la consulta clínica es fundamental en los estudios en ciencias de la salud, y los pacientes virtuales conversacionales representan la herramienta óptima para realizar tales prácticas.

La mayoría de las plataformas conversacionales se basan en procesamiento remoto, realizando todo el procesamiento en servidores externos, lo que puede llevar a fallos de funcionamiento en caso de interrupciones de la conexión. Además, las plataformas de mayor realismo suelen tener costes muy elevados.

Proponemos una plataforma que realice todo el procesamiento en modo local (sin depender de la conexión a internet) y basada en herramientas de código abierto. Como trabajo futuro, planteamos un protocolo de pruebas comparativas frente a una de las plataformas actuales de mayor realismo, para determinar las diferencias en cuanto a la naturalidad de la conversación, tanto en ausencia como en presencia de fallos de conexión.

Referencias

[1] Zary, Nabil, and Uno GH Fors. "WASP-a generic web-based, interactive, patient simulation system." *The New Navigators: from Professionals to Patients*. IOS Press, 2003. 756-761.

[2] Freeman, Karen M., et al. "A virtual reality patient simulation system for teaching emergency response skills to US Navy medical providers." *Prehospital and Disaster medicine*, 2001, vol. 16, no 1, p. 3-8.

[3] Suárez, Ana, et al. "Using a Virtual Patient via an Artificial Intelligence Chatbot to Develop Dental Students' Diagnostic Skills." *International Journal of Environmental Research and Public Health*, 2022, vol. 19, no 14, p. 8735.

[4] Kiesewetter, Jan, et al. "Implementing Remote Collaboration in a Virtual Patient Platform: Usability Study." *JMIR Medical Education*, 2022, vol. 8, no 3, p. e24306.

[5] Harris, Margaret; O'Connor, Alison; Chang, Lih-Fan. "Effectiveness of a Virtual Patient Simulation at Improving Diagnosis and Treatment of Cardiac Amyloidosis." *Journal of Cardiac Failure*, 2022, vol. 28, no 5, p. S125.

[6] Walkiewicz, Maciej; Zalewski, Bartosz; Guziak, Mateusz. "Affect and Cognitive Closure in Students—A Step to Personalised Education of Clinical Assessment in Psychology with the Use of Simulated and Virtual Patients." En *Healthcare*. MDPI, 2022. p. 1076.

[7] Seifert, Lukas B., et al. "Virtual patients versus small-group teaching in the training of oral and maxillofacial surgery: a randomized controlled trial." *BMC Medical Education*, 2019, vol. 19, no 1, p. 1-10.

[8] https://unity.com/es

[9] https://github.com/ChatScript/ChatScript

[10] http://pie.med.utoronto.ca/VIC/

[11] https://laerdal.com/es/products/simulation-training/

[12] https://fullcodemedical.com/

[13] https://bodyinteract.com/

[14] CHAN, See Chai Carol, et al. "Implementation of virtual OSCE in health professions education: A systematic review." *Medical Education*, 2023.

[15] FINK, Maximilian C., et al. Assessment of diagnostic competences with standardized patients versus virtual patients: experimental study in the context of history taking. Journal of Medical Internet Research, 2021, vol. 23, no 3, p. e21196.

[16] https://openai.com/blog/chatgpt

[17] Eysenbach, Gunther, et al. "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, 2023, vol. 9, no 1, p. e46885.

[18] https://cmusphinx.github.io/

[19] https://developers.google.com/ml-kit

[20] https://www.nltk.org/

[21] https://www.soulmachines.com/

[22] https://www.digitalhumans.com/

[23] https://www.reallusion.com/

[24] https://www.blender.org/

[25] https://github.com/makehumancommunity/mpfb2

[26] https://www.synthesia.io/

[27] https://huggingface.co/spaces/CVPR/ml-talking-face

[28] https://www.d-id.com/

[29] https://assetstore.unity.com/packages/tools/animation/naturalfront-3d-face-animation-plugin-free-48380

[30] https://www.unrealengine.com/es-ES/metahuman

[31] MA, Yifeng, et al. StyleTalk: One-shot Talking Head Generation with Controllable Speaking Styles. arXiv preprint arXiv:2301.01081, 2023.

[32] YE, Zhenhui, et al. Geneface: Generalized and high-fidelity audio-driven 3d talking face synthesis. *arXiv preprint arXiv:2301.13430*, 2023.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Superando barreras digitales: Activación y alfabetización en Telemedicina para personas mayores en zonas rurales

José Joaquín Mira[1,2], María Asunción Vicente[3], César Fernández[3], Mercedes Guilabert[2] & Irene Carrillo[2]
[1]Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO)
[2]Departamento de Psicologia de la Salud, Universidad Miguel Hernández (UMH)
[3]Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche, Universidad Miguel Hernández (UMH)

1 (FISABIO), Hospital General Universitario, Ed. Anexo II, 3ª planta, Camí de L´Almassera 11, Elche (03203), Alicante, España
2,3 (UMH), Avenida de la Universidad s/n Elche (03202) Alicante, España

jose.mira@umh.es, suni@umh.es, c.fernandez@umh.es, mguilabert@umh.es, icarrillo@umh.es

Este estudio se centra en los retos que plantea el creciente envejecimiento de la población y el incremento de enfermedades crónicas, especialmente en zonas rurales. Destaca el concepto de "activación del paciente" como un factor clave para mejorar los resultados de salud y optimizar el consumo de recursos sanitarios. Se presenta la telemedicina como una alternativa viable, promoviendo la asistencia domiciliaria y permitiendo a los pacientes permanecer en su entorno natural. No obstante, se resalta la necesidad de superar la brecha digital y mejorar la alfabetización en salud para maximizar los beneficios de la telemedicina. Este documento propone el desarrollo de métodos analíticos orientados a detectar dificultades en la activación de pacientes y monitorizar el uso adecuado de dispositivos de telemedicina, enfocándose en personas mayores en áreas rurales y remotas.

***Palabras Clave*- telemedicina, brecha digital, activación del paciente, enfermedades crónicas y envejecimiento de la población**

I. Introducción

El envejecimiento de la población, particularmente en zonas más rurales, se acompaña de un incremento de las enfermedades crónicas y del fenómeno de la soledad, lo que ocasiona una limitación significativa en la calidad de vida y en el estado funcional de las personas que las padecen. En términos de "morbimortalidad", esta realidad supone un desafío para la capacidad organizativa y sostenibilidad de los sistemas sanitarios y de bienestar social en todo el mundo, especialmente en aquellos lugares, como España, que experimentan un notable incremento anual de este índice de envejecimiento.

En esta situación se dan coincidentemente una serie de factores: la disponibilidad cada vez más limitada de recursos, dificultades de acceso a las consultas, la creciente complejidad del régimen terapéutico y factores inherentes al propio paciente (p. ej., creencias negativas hacia ciertos fármacos, mayor fragilidad, etc.) que exhortan a buscar alternativas para garantizar una atención adecuada a estas personas.

Además, en los enfermos crónicos de mayor edad es común la polimedicación (consumo de 5 o más medicamentos al día), lo que conlleva un mayor incumplimiento terapéutico y un aumento en los errores de medicación en el hogar. Sin embargo, el compromiso del individuo (y de sus cuidadores, cuando los haya) en su autocuidado ha demostrado ser beneficioso para afrontar esta situación y obtener mejores resultados, evitando ingresos hospitalarios o la institucionalización en residencias de la persona, cuando esta prefiere vivir en su hogar. De este modo, ha surgido el concepto de "activación del paciente", que se refiere a la integración del conocimiento, la habilidad y el deseo de las personas para comprometerse a alcanzar mejores resultados en salud. Esto implica aprender a controlar la evolución de la enfermedad, saber gestionar su impacto físico, emocional y social, y decidir cómo y dónde obtener la asistencia sanitaria necesaria en cada momento. Se ha demostrado que este concepto está directamente relacionado con mejoras en el estado de salud (incluyendo una menor mortalidad), que

motiva a los profesionales y que contribuye a un consumo racional de recursos sanitarios.

En este sentido, la telemedicina emerge como una alternativa válida que aporta resultados positivos. Los programas de telemedicina se iniciaron en los años ochenta y actualmente se aplican a una gran variedad de enfermedades crónicas, principalmente a través de soluciones de mHealth (aplicaciones de salud, plataformas web, dispositivos wearables), cuyo número ha crecido a medida que más personas disponen de dispositivos inteligentes (smartphones, tablets, etc.) y a medida que avanzan las tecnologías. La mayoría de los estudios sugieren que la telemedicina es más costo-efectiva, activa a los pacientes y les permite permanecer en su entorno natural el mayor tiempo posible en comparación con la atención tradicional [1-5]. Los mejores resultados se obtienen con pacientes de zonas rurales y con acceso más dificultoso a la asistencia tradicional [6,7]. Nuestro grupo ha estado trabajando en esta dirección y, recientemente [8], hemos llevado a cabo un estudio en el que hemos proporcionado evidencias de que esta activación es posible, aunque existe una brecha digital y lagunas en la alfabetización en salud que deben abordarse para lograr mejores resultados en las experiencias en telemedicina. En concreto, es esencial incrementar el nivel de comprensión de la experiencia de uso de los dispositivos de telemedicina y se requiere más datos sobre cómo lograr el compromiso del paciente en un entorno digital [9,10].

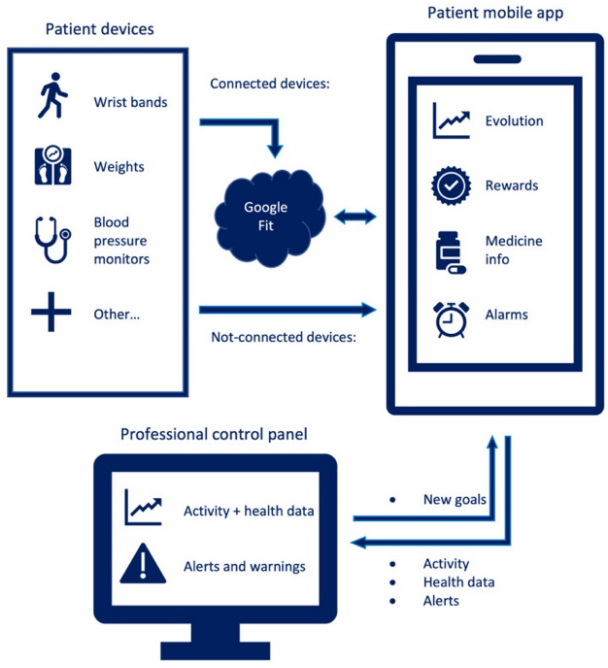


Fig 1. Gráfico de la arquitectura de un sistema de Telemedicina basado en el uso combinado de una app y varios wearables de salud [8].

Esta propuesta se enfoca en este aspecto y, a partir de la experiencia que acumulamos en los estudios sobre aplicaciones de la telemedicina y el desarrollo de medidas basadas en soluciones digitales [8,11], proponemos trabajar en el diseño de métodos de análisis de esta experiencia de uso de dispositivos para la telemedicina con el fin de lograr la activación del paciente mayor que vive en zonas más distantes de los grandes núcleos de población.

La Figura 1 presenta de manera gráfica la arquitectura de un sistema de Telemedicina que se basa en el uso combinado de una aplicación móvil y diversos wearables de salud. La figura 1 han sido extraída de la referencia [8], donde se detalla en profundidad la implementación y funcionamiento de este sistema de Telemedicina basado en el uso combinado de una app y varios wearables de salud.

El objetivo es desarrollar métodos de análisis para detectar dificultades en la activación de estas personas para la telemedicina y para la monitorización del uso apropiado, evitando abandonos y errores de forma proactiva, adaptándonos a las condiciones de las personas mayores que viven en zonas más alejadas de grandes núcleos urbanos (municipios de menos de 3.000 habitantes) con el fin de lograr una mayor comprensión de la realidad de uso y eliminar barreras y dificultades.

II. Metodología

La metodología de nuestra propuesta es la siguiente:

A. Fase 1: Sesiones introductorias

En esta fase inicial, se llevarán a cabo sesiones introductorias destinadas a los responsables de proyectos de Telemedicina, tutores de residentes MIR y demás personal encargado de implementar programas de Telemedicina. El propósito principal de estas sesiones es destacar la necesidad de medir y evaluar la experiencia de uso de las herramientas digitales en el ámbito de la telemedicina.

Para la realización de estas sesiones, se utilizarán herramientas de videoconferencia y plataformas colaborativas en línea, lo que permitirá la participación de múltiples actores de manera eficiente y escalable. Durante estas sesiones, se enfatizará la importancia de la medición precisa de la interacción del usuario con los dispositivos y las aplicaciones de telemedicina.

B. Fase 2: Estudio de las competencias necesarias

En esta fase, se llevarán a cabo sesiones grupales en las que se identificarán y analizarán las competencias necesarias para el uso efectivo de dispositivos digitales en el contexto de la telemedicina. Estas sesiones incluirán la discusión de temas técnicos y prácticos relacionados con la implementación y el uso de tecnologías de telemedicina.

Se emplearán herramientas de colaboración en línea y foros técnicos para facilitar la comunicación y el intercambio de conocimientos entre los participantes. Se

abordarán cuestiones específicas como el perfil del usuario adherente, los errores comunes en la operación de dispositivos, la adaptación a los objetivos de salud pactados, diferencias en el uso según el género, factores predictores de abandono y la utilización del feedback proporcionado por los dispositivos para fomentar cambios en los hábitos de los pacientes.

C. Fase 3: Captura, clasificación y análisis datos de experiencia de uso en estudio de campo de una muestra aleatoria de pacientes

En esta fase crítica del estudio, se procederá a la recopilación, clasificación y análisis de datos relativos a la experiencia de uso, en un estudio de campo con una muestra aleatoria de pacientes a quienes se les suministrarán dispositivos digitales.

Disponemos de una variedad de dispositivos digitales, como glucómetros, básculas, tensiómetros y oxímetros, y estamos trabajando activamente para fomentar proyectos de telemedicina. Contamos con acceso a centros y consultorios en la provincia de Alicante, gracias a nuestra posición en el sistema público de salud, lo que nos permite llevar a cabo el estudio. Poseemos una plataforma online propia que registra toda la información, incluyendo datos que nos permiten analizar la dinámica de uso de los dispositivos. Esta plataforma cumple con los criterios establecidos en la normativa española. Con los datos recopilados de uso, nos proponemos desarrollar rutinas de análisis e indicadores para mejorar nuestro entendimiento sobre cómo los usuarios de telemedicina utilizan los dispositivos digitales, identificar patrones de uso, proponer acciones para mejorar la adherencia y promover programas de soporte, entre otros. El estudio de campo tendría una duración de 9 meses.

Utilizaremos la plataforma PowerBI de Microsoft (o alternativamente R) para clasificar y analizar los datos y desarrollaremos indicadores que posteriormente podrán ser empleados en los programas que se están desarrollando actualmente en el ámbito sanitario para proporcionar alternativas de atención a este perfil de pacientes.

Nuestro equipo de trabajo está compuesto por estadísticos, ingenieros, psicólogos, farmacéuticos y médicos, trabajando en una dinámica multidisciplinar que nos permite atender las diferentes perspectivas que este tipo de proyecto implica, desde cuestiones de índole social hasta aspectos de diseño de procesos, automatización de análisis de datos, etc. Esta fase contará con la aprobación del Comité de Ética del Hospital de Sant Joan D'Alacant.

D. Fase 4: Difusión del estudio

La fase final de este proyecto implica la difusión de los resultados y hallazgos. Se organizará un seminario a nivel nacional que contará con la participación de miembros de equipos de investigación en Telemedicina, responsables de programas a nivel del sistema sanitario y de bienestar social, y expertos nacionales e internacionales en telemedicina y tecnología de la salud.

Este seminario se llevará a cabo en formato remoto, utilizando tecnologías de videoconferencia y plataformas en línea para garantizar la participación de una audiencia amplia y diversa. Durante el seminario, se presentarán los resultados del estudio y se discutirán las implicaciones técnicas y prácticas para el futuro desarrollo de programas de telemedicina.

III. Conclusión

En resumen, este estudio representa un esfuerzo significativo para abordar los retos intrínsecos al envejecimiento demográfico y la prevalencia de enfermedades crónicas en regiones rurales, a través de la aplicación de la telemedicina como herramienta fundamental. Nuestra principal misión es la de profundizar en la comprensión de las complejas barreras y desafíos que afrontan los individuos de edad avanzada en contextos rurales, con el propósito último de eliminar los obstáculos que dificultan la obtención de atención médica adecuada.

Aunque este proyecto se encuentra aún en su fase inicial desarrollo y de análisis, anticipamos que los resultados obtenidos poseen un potencial transformador. Esperamos que estos resultados contribuyan de manera sustancial a la superación de la brecha digital, facilitando un mayor acceso a servicios de salud de calidad para las poblaciones rurales, mientras que se promueve un enfoque altamente centrado en el paciente en el ámbito de la telemedicina. La información recopilada y los indicadores desarrollados tienen el potencial de informar de manera crucial las políticas de atención médica, impulsando innovaciones tecnológicas y cambios en la prestación de servicios, y, en última instancia, mejorando la calidad de vida de los residentes rurales de avanzada edad. Estamos ansiosos por compartir nuestros resultados con la comunidad médica, investigadora y de políticas de salud, con la esperanza de generar un impacto duradero en la atención médica de las poblaciones más vulnerables y alejadas.

Referencias

[1] Eze ND, Mateus C, Cravo Oliveira Hashiguchi T. Telemedicine in the OECD: An umbrella review of clinical and cost-effectiveness, patient experience and implementation. PLoS One. 2020;15(8):e0237585.

[2] Brouwers RWM, van der Poort EKJ, Kemps HMC, van den Akker-van Marle ME, Kraal JJ. Cost-effectiveness of Cardiac Telerehabilitation With Relapse Prevention for the Treatment of Patients With Coronary Artery Disease in the Netherlands. JAMA Netw Open. 2021;4(12):e2136652.

[3] Mileski M, Kruse CS, Catalani J, Haderer T. Adopting Telemedicine for the Self-Management of Hypertension: Systematic Review. JMIR Med Inform. 2017;5(4):e41

[4] Eberle C, Stichling S. Clinical Improvements by Telemedicine Interventions Managing Type 1 and Type 2 Diabetes: Systematic Meta-review. J Med Internet Res. 2021;23(2):e23244.

[5] de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health. 2015;21:81-5.

[6] Zhai YK, Zhu WJ, Cai YL, Sun DX, Zhao J. Clinical- and cost-effectiveness of telemedicine in type 2 diabetes mellitus: a systematic review and meta-analysis. Medicine (Baltimore). 2014;93(28):e312.

[7] Batsis JA, DiMilia PR, Seo LM, Fortuna KL, Kennedy MA, Blunt HB, Bagley PJ, Brooks J, Brooks E, Kim SY, Masutani RK, Bruce ML, Bartels SJ. Effectiveness of Ambulatory Telemedicine Care in Older Adults: A Systematic Review. J Am Geriatr Soc. 2019;67:1737-49.

[8] Vicente MA, Fernández C, Guilabert M, Carrillo I, Martín-Delgado J, Mira JJ. Patient Engagement Using Telemedicine in Primary Care during COVID-19 Pandemic: A Trial Study. Int. J. Environ. Res. Public Health 2022, 19,14682

[9] Khanijahani A, Akinci N, Quitiquit E. A Systematic Review of the Role of Telemedicine in Blood Pressure Control: Focus on Patient Engagement. Curr Hypertens Rep. 2022:1–12. doi: 10.1007/s11906-022-01186-5.

[10] Barello S, Triberti S, Graffigna G, et al. eHealth for Patient Engagement: A Systematic Review. Frontiers In Psychology. 2016;6:2013.

[11] Vicente, MA, Fernández, C. "Herramientas de telemedicina para autocuidado de pacientes y para ayuda a cuidadores. Prototipo para Android." JITEL 2021 Libro de Actas: XV Jornadas de Ingeniería Telemática, A Coruña 2021. Universidad de La Coruña, 2021.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

VQMTK: Una Herramienta Open Source para Evaluar la Calidad de los Vídeos

Wilmer Moina-Rivera, Juan Gutiérrez-Aguado, Miguel Garcia-Pineda.
Departamento de Informática, ETSE-UV
Universitat de València
Av. de la Universitat, s/n. Burjassot. València
wilmoiri@alumni.uv.es, miguel.garcia-pineda@uv.es, juan.gutierrez@uv.es

Resumen

Debido al auge del consumo de contenido multimedia en Internet, las plataformas de *streaming* deben garantizar un cierto nivel de calidad a la hora de preparar sus contenidos. Con este fin, se ha desarrollado una aplicación *open source* que agrupa varias métricas para evaluar la calidad del vídeo de manera sencilla. Este trabajo integra 14 métricas y el SI-TI en un contenedor, para así disponer de una herramienta multiplataforma, la cual hemos denominado Video Quality Metric ToolKit (VQMTK). La herramienta desarrollada ofrece una interfaz web a través de Jupyter notebooks y un script Bash que combina todas las métricas en una única herramienta. Esta herramienta podrá ser utilizada tanto en entornos científicos como educativos. La herramienta está disponible para su uso en: https://github.com/cloudmedialab-uv/vqmtk.

***Palabras Clave*—video, calidad, aplicación, métricas, vqmtk**

I. INTRODUCCIÓN

La evaluación de la calidad del vídeo es esencial para garantizar la calidad del servicio del vídeo proporcionado y mejorar la calidad de la experiencia del usuario final [1]. Los sistemas de evaluación de la calidad permiten a los proveedores de servicios de vídeo identificar posibles problemas de calidad, en la etapa inicial de preparación del contenido, que puedan afectar la experiencia del usuario.

En este trabajo se presenta la herramienta VQMTK, que integra 14 métricas objetivas (PSNR, APSNR, PSNR-HVS, SSIM, MSSSIM, VMAF, VIF, VQM, CAMBI, BRISQUE, NIQE, VIIDEO, STRRED, CIEDE2000) y el SI-TI en una imagen de contenedor para ponerlas a disposición de la comunidad científica como una herramienta de código abierto para la evaluación de la calidad de vídeos sin restricciones. El proyecto puede ser descargado desde nuestra cuenta de Github[1], donde se proporciona información detallada sobre su instalación y uso.

La solución contempla dos componentes principales que hacen uso de estos artefactos; por una parte esta el componente `Jupyter`[2], que proporciona una interfaz gráfica de usuario (GUI) vía navegador Web para ejecutar notebooks que ilustran cada una de las métricas indicadas previamente; por otra parte, el componente `vqmcli`, que utiliza todos los artefactos en una sola herramienta (el usuario utiliza la interfaz de línea de comandos para calcular una o varias métricas a la vez); a su vez, este componente puede ser ejecutado desde `Jupyter`. En [2] se puede obtener una descripción detallada de la herramienta VQMTK, así como una comparativa frente a otras herramientas similares que existen y un estudio del rendimiento de la propuesta en la extracción de métricas con vídeos 4K.

II. CONCLUSIONES

VQMTK permite evaluar fácilmente la calidad objetiva de distintas configuraciones de codificación de vídeo utilizando 14 métricas. Finalmente, esta herramienta también podría utilizarse en entornos educativos, donde se traten temas de codificación y calidad de vídeos, ya que podría ser muy útil para el alumnado por su facilidad de despliegue y utilización.

AGRADECIMIENTOS

Proyecto PID2021-126209OB-I00 financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa.

REFERENCIAS

[1] L. J. Karam, T. Ebrahimi, S. S. Hemami, T. N. Pappas, R. J. Safranek, Z. Wang, and A. B. Watson, "Introduction to the issue on visual media quality assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 189–192, 2009.

[2] W. Moina-Rivera, J. Gutiérrez-Aguado, and M. Garcia-Pineda, "Video quality metrics toolkit: An open source software to assess video quality," *SoftwareX*, vol. 23, p. 101427, 2023. [Online]. Available: https://doi.org/10.1016/j.softx.2023.101427

[1] https://github.com/cloudmedialab-uv/vqmtk

[2] https://jupyter.org

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Network Traffic Analysis for eXtended Reality Applications

Tianhua Chen[1], Elans Grabs[1], Igor Tasic[2], Maria-Dolores Cano[2]
[1]Telecommunications Institute, Riga Technical University, Azenes street 12, LV-1048 Riga, Latvia
[2]Department of Information Technologies and Communication, Universidad Politécnica de Cartagena, 30202 Cartagena, Spain

tianhua.chen@rtu.lv, mdolores.Cano@upct.es

Currently, new multimedia applications are booming, especially streaming services represented by eXtended Reality (XR) and the next generation B5G/6G networks. Network traffic generated by XR and traditional mobile and PC clients is becoming more and more intertwined and complex, especially the homogeneity of Quality of Services (QoS) parameters makes the analysis of traffic critical. In this paper, we review the research done so far in traffic classification for XR services and propose the use of graph neural networks and Quality of Experience (QoE) scores to detect and classify highly similar streaming traffic as well as to predict them in further works. The goal will be not only to secure network resources and devices but also to achieve dynamic traffic classification and resource management for heterogeneous networks with a mixture of multiple network standards to meet the requirements of Self Organizing Networks (SON).

***Keywords*- VR, AR, XR, spatial computing, metaverse, traffic classification, QoS, QoE, 5G.**

I. INTRODUCTION

Nowadays, conventional multimedia applications and services are evolved into the metaverse, including Augmented Reality (AR), Virtual Reality (VR), Mixed Reality (MR), and, in general, eXtended Reality (XR) or spatial computing, compared to unitary text, audio, images, videos, and animation that are more underlining interactive resources integrated.

Head Mounted Devices (HMDs) are one of the representative tools to realize their demands and support three degrees of freedom (3DoF), including pitch, yaw, and roll, or six degrees of freedom (6DoF) subdivided into forward, backward, left, right, up, and down. High degrees of freedom support combined with oversized Field of View (FoV) remarkably enhance users' immersive experience when using HMDs to watch graphics or videos, particularly for 360 degrees videos [1] [2].

While the content and form of multimedia resources continue to evolve, the transmission requirements for next-generation networks are becoming increasingly demanding, as reflected by parameters such as end-to-end latency, jitter, throughput, burst traffic, packet loss ratio, Round Trip Times (RTT), or network delay, among others. Those parameters embedded into Quality of Service (QoS) and Quality of user Experience (QoE) models will significantly impact users' perceptions of XR applications in 5G and future B5G/6G networks [3] [4].

Self Organizing Networks (SON) [5] [6] is one of the most vital development goals of 5G, which mainly refers to Self-Configuration, Self-Optimization, and Self-Healing. Usually, Self-Configuration is linked to Self-Learning. It lets the network understand topology, realize parameter adjustments, and update nodes. Self Healing refers to maintenance assistance, including load balancing, capacity expansion, or power outage. Self Optimization connects to traffic monitoring and performance optimization, dealing with QoS and QoE parameters.

On the other hand, traffic monitoring is a transversal action that plays a top-down role in the Internet. For example, classifying and analyzing network traffic allows for more efficient load balancing of network resources. In the face of new XR applications, how to identify new types of network traffic, i.e., traffic classification, might play a key role in allocating and optimizing network resources.

In this work-in-progress paper, we review the traffic characteristics of XR applications in heterogeneous networks and reflect on the role of combining QoS/QoE levels to detect and classify different XR application applications. The rest of the paper is organized as follows. Section 2 describes the state-of-the-art research, including QoS/QoE optimization for XR traffic and QoS-based traffic classification. The proposal for traffic classification is introduced in Section 3, and finally, the paper

summarizes XR applications' traffic classification importance and future works.

II. STATE OF THE ART

The multimedia represented by XR applications have come to HMDs after the development of PC clients and mobile phone clients. Over The Top (OTT) publishers have produced the corresponding streaming content, live video streaming, and game streaming to a larger extent.

Compared with PCs or smartphones, streaming is characterized by the larger volume size of the video source and the use of ultra-high resolution (4K or 8K). Resolution is mainly determined by Pixels Per Degree (PPD), color depth, and Frames Per Second (FPS); as the most recommended resolution for HMDs monocular resolution is 4K, the left and right eyes also support frame sequential multiplexing when playing streaming, which gives the user an excellent immersion experience.

Considering the bandwidth and delay in network transmission, it also requires powerful video compression coding protocols, including H.264/H.265, VP9 to guarantee lossless content, 50FPS, 60FPS, or even 90FPS to make users feel the video playback smoother. This is also crucial for game streaming. Streaming communication protocols are also key for XR applications, such as Real-time Transport Protocol (RTP/UDP), Real-time Streaming Protocol (RTSP), Dynamic Adaptive Streaming over HTTP (DASH), HTTP Adaptive Streaming (HAS), and Apple's HTTP Live Streaming (HLS) with matching buffer configuration, bitrate adaption, and so forth parameters that need to re-evaluate their transmission performance [7].

Next, we review the research works from the related literature that address traffic classification of XR services and applications. To the authors' knowledge, there are no works related to XR traffic classification based on QoS/QoE.

A. QoS/QoE optimization for XR traffic

For novel XR applications, many studies have started to model optimized QoE for their traffic characteristics in next-generation networks. Mattia *et al.* [8] captured the traffic of three different types of VR applications on the streaming media rendering servers, found the most suitable statistical distribution by calculating the Cumulative Distribution Function (CDF), and generated a model that conforms to the XR traffic type for simulating the scheduling of XR traffic strategy.

Jing *et al.* [9] used the Dynamic time division duplex (D-TDD) technique to improve QoE for MR users. They used a Deep Q-Network (DQN) algorithm to mitigate Inter-Cell Interference (ICI) and optimized Multidimensional Resource Allocation (MRA). DQN, as one of the reinforcement learning algorithms, has achieved a good advantage in QoE-based resource allocation. Similarly, Ismail *et al.* [10] proposed the DeepEdge framework to efficiently allocate edge resources for Edge-IoT applications, which include VR/AR-related applications. They employed Deep Neural Networks (DNNs) with Edge-IoT state mapping to a joint resource allocation operation consisting of resource allocation and QoS categories to maximize QoE scores.

Following the same trend, Cristina *et al.* [11] utilized Deep Recurrent Neural Network (DRNN) based on Gated Recurrent Units (GRUs) to assist in maximizing transmitted video block quality. To reduce VR frame latency, they combined predictive FoV correlation with viewer position viewing 360 degrees HD VR video to enable proactive FoV-centric millimeter Wave (mmWave) Small cell Base Stations (SBSs) physical layer multicast transmission.

From the perspective of Network Slicing (NS), Federico *et al.* [2] collected traffic in real VR applications and investigated their temporal correlation, focusing on the Constant Bit Rate (CBR) encoding mode, which produces more predictable traffic.

The above-mentioned research tasks is mainly focused on finding a balance between traffic and QoE for multimedia XR applications, especially in 5G scenarios. Researchers used advanced algorithms such as machine learning and reinforcement learning to dynamically predict the network resource allocation logic to optimize resource utilization and QoE scores. QoS and QoE parameters are usually set from international standards such as ITU-T G.1035 [12] standard, ITU-T P.1203 [13] standard, ITU-T P.1204 [14] standard, 3GPP TR 26.929 [15] standard. They are further subdivided into Key Performance Indexes (KPI) and Key Quality Indexes (KQI). KPI is more related to QoS parameters, including RTT, burst pulse, etc. KQI for traditional mobile phone calling applications comprises service availability, call drop rate, etc. KQI for XR application includes stalling, Motion To Photon (MTP), FoV, FPS, etc [16].

These highly segmented parameters design requires intelligent network learning to adjust the service capability dynamically. From the traffic monitoring point of view, XR applications traffic in 5G, 6G, WiFi, and other network standards in its QoS and QoE parameters have strong properties relative to other applications, which could have a better basis for this traffic detection classification and guarantee.

B. QoS-based traffic classification

The following works basically use QoS parameters as criteria for the classification task. Zheng *et al.* [17] used sliding window technology to capture slices of multimedia application flows and obtain QoS-related features. However, its multimedia applications are still limited to conventional client-based Conference Video (CV), Video on Demand (VOD), Voice on Demand (Voice), Live Video (LV), HTTP Download Video (HDV), and Peer-to-Peer video (P2P).

A different approach was followed by Huang *et al*. [18], which acted according to different QoS standardization methods, comprising Integrated Services (IntServ) and Differentiated Services (DiffServ). Specifically, DiffServ contains Differentiated Services Code Point (DSCP) tags with Assured Forwarding (AF) [19] categories used to make classifications for mixed video streaming applications. They also obtained high classification metrics.

Gabilondo *et al*. [20] employed the dynamic classification of traffic types for different applications and heterogeneous flows and transmitted them through specific slices with associated 5G QoS Identifiers (5QI). In addition, the authors proposed specific radio resource-sharing configurations to determine the most appropriate traffic priority through scheduling. The slicing advantage improves the efficiency and reliability of the most critical data regarding QoS-related packet loss or jitter.

Our previous work [21] focused on extracting the traffic intensity of different video quality traffic as features from PC clients and achieved a high accuracy of 97% using CNN for supervised classification. All labels were manually marked. However, we did not test XR applications streaming video quality classification performance, and currently, QoE score metrics are mainly used for traffic prediction and resource allocation adjustment. To the authors' knowledge, there is no related work as criteria for classification. The disadvantage of QoS parameters such as DiffServ is the inability to isolate and classify the same type of streaming traffic. Network slicing technology can provide a complete end-to-end virtual network for highly similar traffic to provide different QoS/QoE guarantees, which also could offer a reliable basis for the traffic classification task.

III. Proposal for Traffic Classification

To achieve the goal of traffic classification for XR applications based on QoE scores metrics, we propose the flow chart shown in Fig. 1. This approach could be used for traffic classification and forecasting for streaming applications in heterogeneous networks.

Assuming different types of streaming application traffic in heterogeneous networks, we will detect their corresponding QoE levels, such as mean opinion score (MOS), video quality model (VQM), peak signal-to-noise ratio (PSNR), etc., which are based on the above metrics and classify the traffic according to different calculated level labels. If XR applications are below the expected level, we will optimize the metrics enhancement and reclassify the existing traffic in combination with QoS and 5G network slicing levels parameters, etc. By doing so, the system would combine iterative optimization and dynamic traffic classification to analyze the traffic characteristics of XR applications better.

We plan to simulate and experiment with a simple, smart home scenario. Although many devices could be connected to the smart home, we only focus on devices that can run streaming applications, including PCs, mobile phones, tablets, webcam monitors, HMDs, etc. Streaming and network video-related applications include live video streaming, online games, online conferences, XR applications built into HMDs, and streaming between multiple devices such as STEAM VR clients and HMDs, among others. In a heterogeneous network environment consisting of WIFI, 4G, and 5G, different applications and network standards are used alternatively, where these applications generate highly similar network traffic. The traffic generated by streaming needs to be further modeled, correlation analyzed, and useful training features extracted, and machine learning algorithms are used to achieve traffic prediction and classification.

We propose to use Graphical Neural Networks (GNN) [22] for traffic classification and prediction in this scenario. GNN can be used not only for traffic classification but also for traffic prediction regression, which has an excellent theoretical basis for analyzing end-to-end traffic relationships. We could also add Network Intrusion Detection Prevention Systems (NIDPS) to protect users' cybersecurity. In general, the final purpose of classification is to help OTT and Internet service providers better understand the proportion of traffic by various applications and to achieve the goal of SON by combining traffic volume prediction for better allocation and load balancing of physical resources of the network system.

IV. Conclusions And Future Work

It is well known that one of the main challenges for streaming media applications in Internet-generated traffic is bandwidth consumption. This fact is enlarged with the introduction and adoption of all kinds of new XR applications; how to identify, secure, classify, and forecast this traffic are some of the questions suggested by the researchers. Besides, the challenge of having limited physical resources while maintaining QoE levels in heterogeneous networks is another important constraint. Our future work will focus on dynamic traffic classification and prediction under this XR paradigm to improve network resource utilization.

Acknowledgments

This work has been supported by the European Social Fund within the Project No 8.2.2.0/20/I/008 «Strengthening of PhD students and academic personnel of Riga Technical University and BA School of Business and Finance in the strategic fields of specialization» of the Specific Objective 8.2.2 «To Strengthen Academic Staff of Higher Education Institutions in Strategic Specialization Areas» of the Operational Programme «Growth and Employment» and also funded by grant PID2020-116329GB-C22 funded by MCIN/AEI/10.13039/501100011033.

References

[1] Y. Wang *et al.*, "A Survey on Metaverse: Fundamentals, Security, and Privacy," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 1, pp. 319–352, 2023, doi: 10.1109/COMST.2022.3202047.

[2] F. Chiariotti, M. Drago, P. Testolina, M. Lecci, A. Zanella, and M. Zorzi, "Temporal Characterization and Prediction of VR Traffic: A Network Slicing Use Case," *IEEE Trans. Mob. Comput.*, pp. 1–18, 2023, doi: 10.1109/TMC.2023.3282689.

[3] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research

Problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May 2020, doi: 10.1109/MNET.001.1900287.

[4] R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, "Revisiting VoIP QoE assessment methods: are they suitable for VoLTE?," *Netw. Protoc. Algorithms*, vol. 8, no. 2, p. 39, Jul. 2016, doi: 10.5296/npa.v8i2.9123.

[5] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 3, pp. 1617–1655, 2016, doi: 10.1109/COMST.2016.2532458.

[6] S. Latif, F. Pervez, M. Usama, and J. Qadir, "Artificial Intelligence as an Enabler for Cognitive Self-Organizing Future Networks," 2017, doi: 10.48550/ARXIV.1702.02823.

[7] A. Bentaleb, B. Taani, A. C. Begen, C. Timmerer, and R. Zimmermann, "A Survey on Bitrate Adaptation Schemes for Streaming Media Over HTTP," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 1, pp. 562–585, 2019, doi: 10.1109/COMST.2018.2862938.

[8] M. Lecci, M. Drago, A. Zanella, and M. Zorzi, "An Open Framework for Analyzing and Modeling XR Network Traffic," *IEEE Access*, vol. 9, pp. 129782–129795, 2021, doi: 10.1109/ACCESS.2021.3113162.

[9] J. Song, Q. Song, Y. Kang, L. Guo, and A. Jamalipour, "QoE-Driven Distributed Resource Optimization for Mixed Reality in Dynamic TDD Systems," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7294–7306, Nov. 2022, doi: 10.1109/TCOMM.2022.3208113.

[10] I. Alqerm and J. Pan, "DeepEdge: A New QoE-Based Resource Allocation Framework Using Deep Reinforcement Learning for Future Heterogeneous Edge-IoT Applications," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 4, pp. 3942–3954, Dec. 2021, doi: 10.1109/TNSM.2021.3123959.

[11] C. Perfecto, M. S. Elbamby, J. D. Ser, and M. Bennis, "Taming the Latency in Multi-User VR 360°: A QoE-Aware Deep Learning-Aided Multicast Framework," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2491–2508, Apr. 2020, doi: 10.1109/TCOMM.2020.2965527.

[12] "ITU Telecommunication Standardization Sector ITU-T Rec G 1035 Multimedia Quality of Service and performance – Generic and user-related aspects." [Online]. Available: https://www.itu.int/rec/T-REC-G.1035-202111-I

[13] "ITU Telecommunication Standardization Sector ITU-T Rec P 1203 Models and Tools for Quality Assessment of Streamed Media." [Online]. Available: https://www.itu.int/rec/T-REC-P.1203-201710-I

[14] "ITU Telecommunication Standardization Sector ITU-T Rec P 1204 Models and tools for quality assessment of streamed media." [Online]. Available: https://www.itu.int/rec/T-REC-P.1204-202001-I

[15] "3rd Generation Partnership Project Technical report 26.929 QoE parameters and metrics relevant to the Virtual Reality (VR) user experience." [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3327

[16] H. Dong and J. S. A. Lee, "The Metaverse From a Multimedia Communications Perspective," *IEEE Multimed.*, vol. 29, no. 4, pp. 123–127, Oct. 2022, doi: 10.1109/MMUL.2022.3217627.

[17] Z. Wu, Y. Dong, X. Qiu, and J. Jin, "Online multimedia traffic classification from the QoS perspective using deep learning," *Comput. Netw.*, vol. 204, p. 108716, Feb. 2022, doi: 10.1016/j.comnet.2021.108716.

[18] Y.-F. Huang, C.-B. Lin, C.-M. Chung, and C.-M. Chen, "Research on QoS Classification of Network Encrypted Traffic Behavior Based on Machine Learning," *Electronics*, vol. 10, no. 12, p. 1376, Jun. 2021, doi: 10.3390/electronics10121376.

[19] M.-D. Cano, F. Cerdán, J. García-Haro, and J. Malgosa-Sanahuja, "A New Proposal for Assuring Services in Internet," in *Proceedings of the International Conference on Internet Computing, IC 2002, Las Vegas, Nevada, USA, June 24-27, 2002*, H. R. Arabnia and Y. Mun, Eds., CSREA Press, 2002, pp. 379–384.

[20] Á. Gabilondo *et al.*, "Traffic Classification for Network Slicing in Mobile Networks," *Electronics*, vol. 11, no. 7, p. 1097, Mar. 2022, doi: 10.3390/electronics11071097.

[21] T. Chen *et al.*, "Multiclass Live Streaming Video Quality Classification Based on Convolutional Neural Networks," *Autom. Control Comput. Sci.*, vol. 56, no. 5, pp. 455–466, Oct. 2022, doi: 10.3103/S0146411622050029.

[22] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021, doi: 10.1109/TNNLS.2020.2978386.

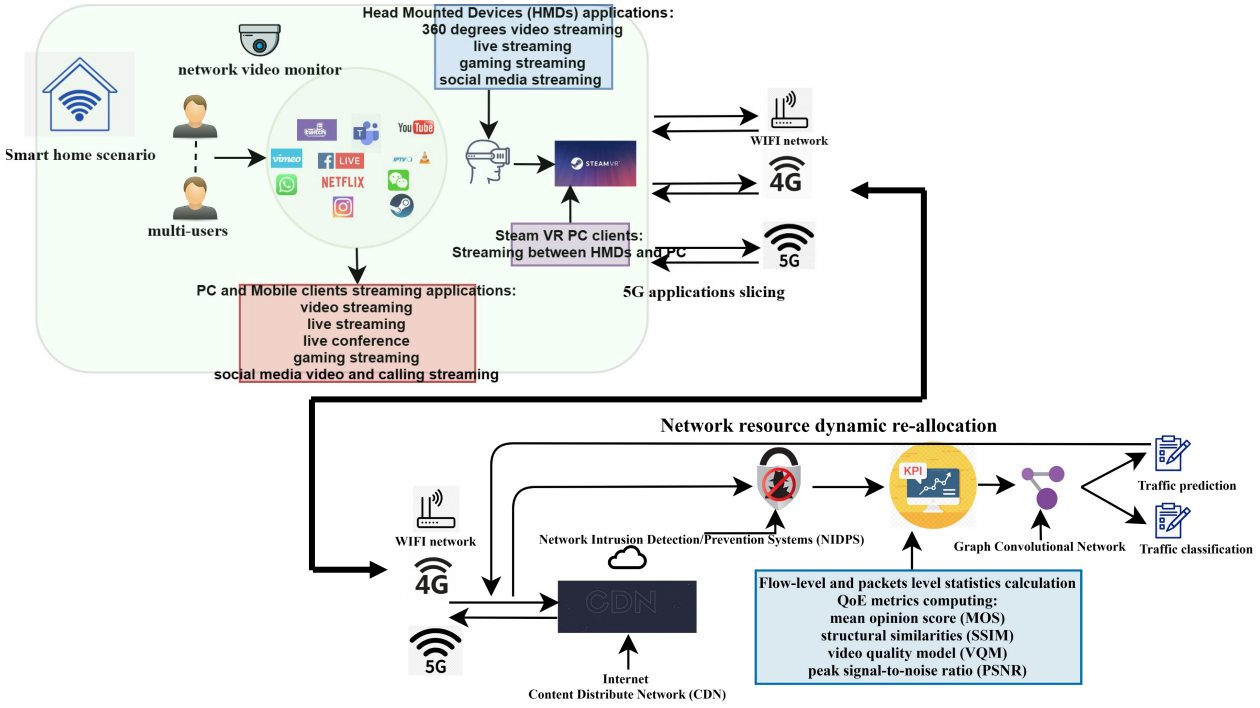


Fig. 1. Flow chart of traffic classification and prediction for streaming applications in heterogeneous network

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Innovando en el cuidado a través de la tecnología: la línea de investigación "REALITY CARE"

J. J. Mira[2,3], P. Ballester[3], M. Guilabert[2], I. Carrillo[2], E. Gil[3], M. A. Vicente[1], C. Fernández[1], A. Arroyo[4], Á. Cobos[5], S. Lorenzo[6], Á. Márquez[7], P. Navas[8] & P. Pérez-Pérez[9]

[1]Área de Ingeniería Telemática, Escuela Politécnica Superior de Elche, Universidad Miguel Hernández (UMH); [2]Departamento de Psicologia de la Salud, (UMH); [3]Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO); [4]Centro Universitario de Enfermería "San Juan de Dios", Universidad de Sevilla; [5]Hospital Universitario San Cecilio, Granada; [6]Hospital Universitario Fundación Alcorcón, Madrid; [7]Hospital San Juan Grande, Jérez; [8]Hospital Ramón y Cajal, Madrid; [9]Orden Hospitalaria de San Juan de Dios, Sevilla.

jose.mira@umh.es; maria.ballester04@umh.es; mguilabert@umh.es; icarrillo@umh.es; eva.gilh@umh.es; suni@umh.es; c.fernandez@umh.es; almudena.arroyo@sjd.edu.es; angel.cobos.sspa@juntadeandalucia.es; slorenzom@salud.madrid.org; alvaro.marquez@sjd.es; paloma.navas@salud.madrid.org; pastora.perez@sjd.es

La línea de investigación "REALITY CARE" aborda el desafío del envejecimiento poblacional y el incremento de enfermedades no transmisibles, centrando su esfuerzo en mejorar la formación de cuidadores de pacientes dependientes. Mediante el uso de tecnologías como la Realidad Virtual, Realidad Aumentada e Inteligencia Artificial, busca potenciar la eficacia y eficiencia de la formación de cuidadores, explorar la personalización de esta formación y desarrollar métricas apropiadas para evaluar su rendimiento.

***Palabras Clave*- salud, cuidadores, formación, realidad virtual y aumentada, inteligencia artificial (IA).**

I. Introducción

Nuestra sociedad se enfrenta a desafíos importantes como el envejecimiento poblacional y el incremento de enfermedades no transmisibles. Frente a esta realidad, nuestra línea de investigación, REALITY CARE [1], busca soluciones que permitan a las personas envejecer con dignidad y preservar su autonomía. Nos enfocamos en mejorar la formación de cuidadores, tanto formales como informales, mediante el uso de tecnologías emergentes como la Realidad Virtual (RV), la Realidad Aumentada (RA) y la Inteligencia Artificial (IA), las cuales pueden empoderar a los cuidadores y contribuir al desarrollo de nuevas oportunidades en la economía del cuidado, reduciendo la brecha de género en este ámbito.

Nuestro objetivo principal es investigar si la formación de cuidadores en el hogar a través de RV/RA es más eficiente que la formación tradicional. Además, pretendemos desarrollar rutinas de IA para crear materiales inmersivos personalizados que mejoren la efectividad de la formación. Para ello, actualmente estamos realizando un estudio multicéntrico con métodos mixtos en Comunidad Valenciana, Madrid, Andalucía y Región de Murcia. que incluye el desarrollo de vídeos 3D inmersivos y un nuevo procedimiento para analizar los datos del entrenamiento con RV/RA. Buscamos generar nuevas evidencias sobre la eficacia y eficiencia de la formación de cuidadores con RV/RA, explorar opciones para personalizar la formación, y desarrollar métricas para evaluar el rendimiento de los cuidadores en entornos virtuales. Desde España, buscamos liderar la investigación en este campo en Europa, aportando soluciones innovadoras para afrontar los retos de la economía de los cuidados.

II. Proyectos de investigación

El equipo de investigación REALITY CARE está compuesto por profesionales de la salud e investigadores de diferentes ámbitos de la ingeniería y las ciencias de la salud. Sus miembros han participado en varios proyectos nacionales y europeos, con una larga experiencia en el campo de la salud digital y la seguridad del paciente.

También indicamos nuestros proyectos más recientes en esta línea: *Uso seguro de la medicación en el hogar por cuidadores/as* (PI21/00646; UGP-21-062); *Capacitación de cuidadores/as mediante Realidad Virtual y Aumentada. Estudio Experimental* (TED2021-130383B-I00; UGP-22-003); *Soluciones educativas innovadoras enfocadas a aspectos de la labor de personas que son cuidadores/as de adultos/as mayores* (PI22/00868; UGP-22-016).

Referencias

[1] Grupo de Investigación "RealityCare", https://realitycare.es

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Sistema de apoyo a la gamificación en plataformas MOOC

Alejandro Ortega Arranz[1,2], Juan I. Asensio Pérez[1], Alejandra Martínez Monés[2], Miguel L. Bote Lorenzo[1], Héctor Ortega Arranz[2], Marco Kalz[3]
[1]Departamento de Teoría de la Señal, Comunicaciones e Ingeniería Telemática, Universidad de Valladolid
[2]Departamento de Informática, Universidad de Valladolid
[3]Institut für Kunst, Musik und Medien, Pädagogische Hochschule Heidelberg

alex@gsic.uva.es, juaase@tel.uva.es, amartine@infor.uva.es, migbot@tel.uva.es, hector@infor.uva.es, kalz@ph-heidelberg.de

I. Resumen de trabajo ya publicado

Las estrategias de gamificación basadas en recompensas se proponen como una técnica prometedora para aumentar la implicación de los estudiantes en los MOOC (*Massive Open Online* Courses, Cursos en Línea Masivos y Abiertos), tras su éxito en otros entornos educativos a pequeña escala. En una revisión sistemática de la literatura llevada a cabo sobre gamificación en MOOC se encontraron 8 propuestas de sistemas que podrían ser utilizados por los instructores para gamificar los MOOC. Sin embargo, la realización de un análisis sistemático de las características de dichos sistemas desveló que todos ellos cuentan con dos o más de las siguientes 6 limitaciones [1]:

1. Falta de interfaces gráficas y apoyo automático a la autoría y monitorización de la gamificación.
2. Expresividad limitada de los diseños de gamificación.
3. Diseño *ad hoc* para plataformas MOOC específicas incompatible con otras plataformas.
4. Imposibilidad de integrar en la gamificación las herramientas propias de la plataforma MOOC.
5. Imposibilidad de integrar en la gamificación herramientas de terceros habitualmente usadas en MOOC.
6. Ausencia de opción de participación en la gamificación a discreción de los estudiantes.

En [1] se presenta el trabajo de investigación realizado con el objetivo de proponer un nuevo sistema de apoyo a la gamificación en plataformas MOOC que supere estas limitaciones. Dicho trabajo ha sido llevado a cabo siguiendo la Metodología de Investigación para el Desarrollo de Sistemas (*Systems Development Science Research Methodology*).

El sistema propuesto, llamado GamiTool, cuenta con una arquitectura basada en adaptadores diseñada para facilitar su integración con un amplio rango de plataformas y herramientas. GamiTool proporciona a los instructores el apoyo necesario para evitar que la gamificación dé lugar a un incremento en la carga de trabajo que suponga una barrera para su adopción. Su modelo de datos ha sido concebido para dotar a los diseños de gamificación con una amplia expresividad. Además, Gamitool permite a los estudiantes participar en la gamificación de manera opcional.

En la actualidad existe un prototipo de GamiTool* que ya ha sido utilizado para la gamificación de un MOOC impartido en la plataforma Canvas Network. El prototipo también ha sido sometido a una evaluación basada en usuarios con 19 participantes de 10 instituciones diferentes pertenecientes a 6 países distintos. Todos ellos contaban con experiencia previa como instructores MOOC y/o como diseñadores de gamificación. Esta evaluación permitió comprobar que el sistema propuesto genera una baja carga de trabajo, es usable, tiene altas probabilidades de adopción y amplia expresividad. Por lo tanto, GamiTool puede ser utilizado por los instructores para mejorar la implicación de los estudiantes.

Agradecimientos

Este trabajo de investigación ha sido financiado por el Fondo Europeo de Desarrollo Regional y la Agencia Nacional de Investigación del Ministerio de Ciencia e Innovación bajo los proyectos TIN2017-85179-C3-2-R y PID2020-112584RB-C32.

Referencias

[1] A. Ortega-Arranz, J.I. Asensio-Pérez, A. Martínez-Monés, M.L. Bote-Lorenzo, H. Ortega-Arranz, M. Kalz, "GamiTool: supporting instructors in the gamification of MOOCs", *IEEE Access*, vol. 10, pp. 13165-131979, 2022, doi: 10.1109/access.2022.3228762

* https://www.gsic.uva.es/gamitool/index.html, última visita: septiembre de 2023.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Sistemas ADAS para la mejora de la seguridad en vehículos industriales off-road

Roberto García[(1)], Xabiel G. Pañeda[(1)], Dan Garcia-Carrillo[(1)], David Melendi[(1)], Victor Corcoba[(1)], Filipa Mourao[(2)], Sara Paiva[(2)].
Departamento de Informática, Universidad de Oviedo, Asturias[(1)]

ADiT-Lab, Instituto Politécnico de Viana do Castelo, Portugal [(2)]

{garciaroberto,xabiel,garciadan,melendi,corcobavictor}@uniovi.es[(1)], {fmourao,sara.paiva}@estg.ipvc.pt[(2)]

En este artículo se presenta la línea de investigación en sistemas de asistencia a la conducción para la mejora de la seguridad en vehículos industriales *off-road* en el contexto de la industria 4.0. En esta línea de investigación colaboran investigadores de la Universidad de Oviedo y del Instituto Politécnico de Viana do Castelo, en Portugal.

***Palabras Clave*- vehículo industrial, ADAS, seguridad, industria 4.0, sistemas de transporte inteligentes**

I. Descripción y Resultados

La conducción de vehículos industriales en las fábricas es una tarea compleja que debería requerir de sistemas de asistencia. Por lo general, los vehículos son de gran tamaño y potencia lo que, en caso de accidente, puede provocar daños importantes en la estructura de la fábrica o en otros vehículos. Si el impacto es sobre un operador, las consecuencias pueden ser muy graves. Por ejemplo, en el año 2022 se registraron en los Estados Unidos de Norteamérica 108 accidentes en los que había implicada una carretilla elevadora [1]. Estas condiciones de trabajo hacen que la conducción del vehículo requiera de una gran habilidad y esfuerzo cognitivo por parte de los profesionales. Deben cumplir con la tarea asignada a tiempo, velando también por la seguridad de las personas e infraestructuras con las que comparten el espacio.

Los objetivos de esta investigación, financiada a través del proyecto SEGURTRUCK [2], son: diseñar un sistema de asistencia a la conducción (ADAS) versátil, completo y conectado; crear una interfaz de comunicación persona-computador configurable y adaptable; establecer un ecosistema de comunicación de diferentes elementos de detección y actuación para alimentar al ADAS; y evaluar la usabilidad y utilidad del asistente. El sistema recibirá información de otros vehículos, de los operarios y de detectores fijos, para disponer de información sobre el estado del área de trabajo y poder anticiparse a la situación de riesgo. Asimismo, dispondrá de sistemas de aviso al conductor, lo que le permitirá ser configurado para cualquier contexto de actividad. El sistema debe ser de bajo coste y fácilmente adaptable a cualquier tipo de máquina industrial y contexto.

Uno de los trabajos realizados dentro de esta línea de investigación ha sido la construcción de un entorno ciberfísico que permite evaluar el diseño y configuración de estos ADAS [3]. El entorno de pruebas posibilita la monitorización de tiempos de respuesta y la estimación de la carga de trabajo, física y mental, que provoca el sistema de asistencia. Incluye hardware específico, un simulador basado en el motor UNITY adaptado al contexto industrial, un sistema de registro de eventos, un sistema de comunicación para interactuar con el ADAS a evaluar y un sistema de evaluación. Se muestran en la Fig. 1, a modo de ejemplo, los tiempos de respuesta a los avisos de obstáculo oculto y la respuesta cardiaca (intervalo R-R) de un conductor durante una prueba de conducción.

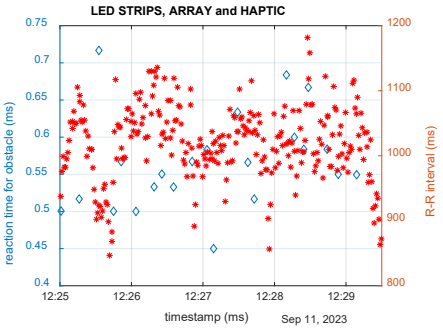


Fig. 1. Monitorización de tiempos de respuesta e intervalo R-R cardiaco

Referencias

[1] Occupational Safety and Health Administration (OSHA), "OSHA Fatality and Catastrophe Investigation Summaries", Accedido abril 2023. https://www.osha.gov/pls/imis/accidentsearch.html.

[2] SEGURTRUCK. Sistema de asistencia a la conducción para la mejora de la seguridad en vehículos industriales off-road en el contexto de una industria 4.0, TED2021-130919B-I00. Ministerio de Ciencia e Innovación, 2022-2024.

[3] X. G. Pañeda, R. García, D. Melendi, D. G. Carrillo, D. Martínez and V. Corcoba, "Testbed for industrial advanced driver assistance systems," in IEEE Latin America Transactions, vol. 21, no. 5, pp. 613-620, May 2023, doi: 10.1109/TLA.2023.10130832.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

A Self-Sustainable Opportunistic Solution for Emergency Detection in Ageing People Living in Rural Areas

Manuel Jesús-Azabal, Javier Berrocal, Vasco N. G. J. Soares, José García-Alonso and Jaime Galán-Jiménez
Departamento de Ingeniería de Sistemas Informáticos y Telemáticos,
Universidad de Extremadura
Avda. de la Universidad, S/N, 10003 Cáceres
manuel@unex.es, jberolm@unex.es, vasco.g.soares@ipcb.pt, jgaralo@unex.es, jaime@unex.es

There are contexts where the use of TCP/IP protocol is not possible due to the lack of infrastructure. In these cases, where latency is unpredictable, alternatives such as Opportunistic Networks (OPPNets) prove valid. Such challenges are common in rural areas, posing an obstacle to eHealth technologies for older adults. Thus, the present work introduces Interest-based System for Communication in Isolated Areas (ISCA), an OPPNet architecture for remote monitoring and emergency detection in ageing people who live alone. For this, the energy requirements are considered, providing sustainable operation and applying a routing algorithm based on interests. ISCA is evaluated over a realistic scenario and compared with state-of-the-art solutions. The experiment shows that ISCA improves the delivery probability in comparison to alternatives and provides a suitable average latency, overhead and number of hops.

***Palabras Clave*—Opportunistic networks, sustainability, routing algorithm, rural areas**

I. SUMMARY

Many rural areas cannot support the TCP/IP protocol due to its high investment requirements[1]. In these scenarios, services such as eHealth solutions for older adults become hard to deploy. Hence, the Interest-based System for Communication in Isolated Areas (ISCA) proposes a disruptive communication architecture which applies Opportunistic Networks (OPPNet) and a self-sustainable infrastructure to detect and transmit potential emergency events in ageing people who live alone in rural areas. For this, 1) Internet of Things (IoT) sensors installed in homes gather the presence data of the elders. Then, 2) gathered data is transmitted opportunistically with Bluetooth to intermediate devices such as surrounding smartphones. As a result, 3) the information is forwarded towards a medical center, exploiting the physical movements and interactions of the devices. Thus, it is possible to use only presence information to analyze and identify anomalous behaviors. This communication process has been designed to be deployed in rural contexts, enabling its use for other purposes. Hence, local businesses such as bee hives or livestock can potentially benefit from the proposed architecture. As a result, owners and workers can be notified about events and production statistics.

Communications are performed using a routing algorithm based on interest. This protocol is focused on selecting dynamic nodes paths considering the individual preferences and the carried information. Thus, nodes perform store-carry-forward tasks with the information that they are interested in. This solution has been assessed in a realistic simulated scenario to compare its outcomes with the alternatives and to study its energy consumption.

The performance is compared with state-of-the-art solutions such as LifeRouter, GeoSpray, WaveRouter, Simbet, and BubbleRap. The delivery probability reaches 97.5%, improving the second-best result by 52.25%, maintaining suitable average latency, overhead and hops. Furthermore, energy consumption displays good outcomes, providing viable battery requirements in those devices fed by solar power. As a result, the assessment manifests self-sustainable feasible power requirements with positive routing metrics.

[1]This work has been published in Wireless Networks, 2023. Impact factor: 2.701. DOI: https://doi.org/10.1007/s11276-023-03294-9. This work has been partially funded 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), by the by the Cap4IE project (0786_CAP4ie_4_P) funded by the Interreg V-A España-Portugal (POCTEP) 2014-2020 program, by the Regional Ministry of Economy, Science and Digital Agenda of the Regional Government of Extremadura (GR21133) and the European Regional Development Fund. V.N.G.J.S. acknowledges that this work is funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Estimating ideology and polarization in European countries using Facebook data

Francisco Caravaca, José González-Cabañas, Ángel Cuevas, Rubén Cuevas
Department of Telematic Engineering, Universidad Carlos III de Madrid.
Av. de la Universidad, 30, 28911 Leganés, Madrid
fcaravac@pa.uc3m.es, jose.gonzalez.cabanas@uc3m.es, acrumin@it.uc3m.es, rcuevas@it.uc3m.es

Researchers have studied political ideology and polarization in many different contexts since their effects are usually related to aspects and actions of individuals and societies. Hence, being able to estimate and measure the changes in political ideology and polarization is crucial. In this paper, we model the ideology and polarization of 28 countries using Facebook public posts from political parties' Facebook pages. We collected a three-year dataset from 2019 to 2021 with information from 234 political parties' Facebook pages and took advantage of the EU parliament elections of May 2019 to create our models. Our methodology efficiently spans 28 countries, offering cost-effective, high-resolution measurements of ideology and polarization. Validation against 19 countries' elections confirms accuracy. Additionally, we've created a website with detailed party information and temporal evolution of our estimations for accessibility.

***Palabras Clave*—Europe, Politics, Barometer, Ideology, Polarization, Facebook**

I. Methodology and results

Our work [1] discusses a methodology for predicting the ideology and polarization of European Union (EU) + UK countries using Facebook data. Ideology is computed as the weighted average ideology of political parties based on the percentage of votes they receive. Polarization measures the spread of ideologies in a country's political system using Dalton's formula [2]. We built an automated tool to collect posts from Facebook pages. We obtained 500k posts from a list of 246 parties, from 2019 to 2021.

To create the models, we consider various variables from Facebook data, including popularity, activity, and engagement metrics of political parties. We compute ideology and polarization values for each country using these variables. Election results were used as ground truth data to train the models. Those elections were the May 2019 EU elections, which were held simultaneously in all 28 countries. The linear regression models were built using different combinations of variables and different time windows to evaluate their performance based on the ground truth data. Finally, the models were primarily validated using results from 15 national parliament elections between 2019 and 2020.

Results show that the best-performing models for estimating ideology and polarization use a combination of post-activity and engagement. Those models have a R^2 of 0.752 (ideology) and 0.808 (polarization).

II. EU Political Barometer

To complete our research work we have built an interactive website: the *EU Political Barometer*, available at: https://eupoliticalbarometer.uc3m.es, providing easy access to our research findings. Our site displays information about the activity and engagement per political party and country, as well as the evolution of the political ideology and polarization based on our models. The site includes more information and insights from the Facebook dataset and it is regularly updated.

III. Conclusions

In conclusion, we introduced a method to estimate ideology and polarization in EU + UK countries using Facebook data. Our models, based on just two variables, proved highly accurate. This study enhances political insight and offers a tool for cross-country analysis. The full research contribution can be found in https://doi.org/10.1140/epjds/s13688-022-00367-1.

Acknowledgments

Authors acknowledge the funding from project ENTRUDIT (Grant TED2021-130118B-I00) funded by the MCIN/AEI/10.13039/501100011033 and the EU FEDER funds; and from the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with UC3M ("Fostering Young Doctors Research", UEMEASURE-CM-UC3M).

Referencias

[1] Caravaca, F., González-Cabañas, J., Cuevas, Á. et al. Estimating ideology and polarization in European countries using Facebook data. EPJ Data Sci. 11, 56 (2022). https://doi.org/10.1140/epjds/s13688-022-00367-1.

[2] Dalton RJ (2008) The quantity and the quality of party systems: party system polarization, its measurement, and its consequences. Comp Polit Stud 41(7):899–920

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Un sistema de alerta temprana de ventilación (SATV) para espacios de trabajo diáfanos considerando escenarios de COVID-19 y futuras pandemias

Gonçal Costa[1], Oriol Arroyo[2], Pablo Rueda[3], Alan Briones[4].

[1] Human Environment Research (HER), La Salle, Ramon Llull University, 08022, Barcelona, España.
[2] Noumena, 08019, Barcelona, España.
[3] CT Solutions Group, 08018, Barcelona, España.
[4] Research Group on Smart Society, La Salle, Ramon Llull University, 08022, Barcelona, España.

goncal.costa@salle.url.edu, oriol@noumena.io, prueda@ctgrupo.com, alan.briones@salle.url.edu.

La pandemia de COVID-19 ha generado nuevas necesidades debido a los riesgos sanitarios asociados y, más concretamente, a su rápida tasa de contagios. Las medidas de prevención para evitar contagios en espacios interiores, especialmente en oficinas y edificios públicos (por ejemplo, hospitales, administración pública, centros educativos, etc.), han llevado a la necesidad de una ventilación adecuada para diluir la posible concentración del virus. Este artículo presenta nuestra contribución a este nuevo reto a través de un sistema de alerta temprana de ventilación (SATV) que tiene como objetivo adaptar el funcionamiento de los actuales sistemas de climatización a las necesidades de ventilación de los espacios diáfanos de trabajo, basándose en un enfoque Smart Campus Digital Twin (SCDT), manteniendo la sostenibilidad. Para ello se han combinado diferentes tecnologías como el Internet de las cosas (IoT), el modelado de información de construcción (BIM), y algoritmos de inteligencia artificial (AI) para recopilar e integrar datos de monitoreo (registros históricos, información en tiempo real y patrones relacionados con la ubicación) para la realización de simulaciones de previsiones en este gemelo digital.

***Palabras Clave*- Smart Building, Gemelo Digital, COVID-19, IoT, Simulación, BIM, Gestión de instalaciones**

I. INTRODUCCIÓN

La eficiencia operativa de los edificios ha progresado sustancialmente en los últimos años con la integración de las tecnologías de Internet de las cosas (IoT) y Building Information Modeling (BIM). Su combinación proporciona un poderoso paradigma de aplicación para mejorar el funcionamiento de los edificios [1]. Aquí, el uso de datos de sensores en tiempo real se presenta como una alternativa más confiable al uso de simulaciones, que continúan mostrando discrepancias significativas entre el comportamiento esperado de un edificio y su comportamiento real [2]. La introducción del concepto de gemelo digital en el dominio de la gestión de instalaciones puede servir para obtener lo mejor de estos dos mundos como un medio para vincular modelos y simulaciones digitales con datos del mundo real [3].

Las soluciones basadas en un enfoque de gemelo digital se presentan como una forma para que los administradores de instalaciones obtengan una mejor comprensión del comportamiento del edificio en tiempo real y, por lo tanto, puedan, por ejemplo, hacer un uso más eficiente de los sistemas de climatización, mejorando su rendimiento operativo. En línea con este propósito, este artículo presenta un sistema de alerta temprana de ventilación (SATV) basado en un enfoque de gemelo digital para espacios de trabajo diáfanos que contempla la prevención de la propagación de virus que puedan representar un serio riesgo para la salud humana, como lo ha sido en su momento el COVID-19, manteniendo la sostenibilidad y el ahorro energético en los sistemas de climatización. El sistema complementa el envío de avisos con la posibilidad de actuar sobre actuadores rotativos instalados en las filas de ventilación para permitir la restricción de la ventilación en zonas donde no es necesaria. El sistema propuesto ha sido pensado para entornos ya construidos en los que ya existe un sistema de climatización instalado sujeto a una serie de características.

II. SISTEMA DE ALERTA TEMPRANA DE VENTILACIÓN (SATV)

El sistema se ha desarrollado contextualizando el marco del gemelo digital de campus inteligente presentado previamente en [4]. Para demostrar su funcionamiento, el sistema se ha implementado en un caso de estudio: el Internet of Things Institute of Catalonia (IoTICAT), un espacio de trabajo para el personal académico docente e investigador de La Salle Campus Barcelona de la Universidad Ramon Llull, en dónde se han desplegado diferentes sensores (Temperatura, Humedad, CO_2) para monitorizar el espacio (Fig. 1). Se han utilizado las cámaras de seguridad existentes para controlar la ocupación de las diferentes zonas del espacio de trabajo. Toda la información recopilada se muestra en un tablero BIM en tiempo real en la web. El software de previsión desarrollado utiliza esta información para enviar alertas tempranas a los gerentes de las instalaciones.

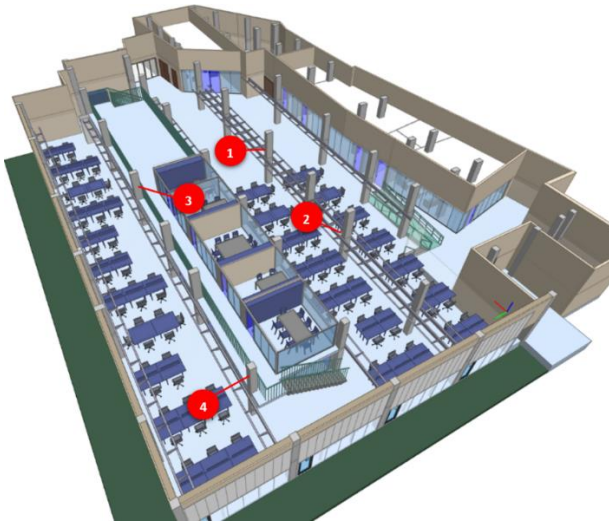


Fig. 1. BIM del espacio IoTICAT indicando la ubicación de los sensores (MSI Studio, 2021).

A. Arquitectura del sistema

Consta de una infraestructura digital para monitorear las condiciones de aire en cada zona del espacio de trabajo (Instalación/BIM, Parte 1), y de un servicio en la nube para almacenar y visualizar la información y simular eventos con el objetivo de, a través del tablero, brindar al administrador una perspectiva del estado del espacio de trabajo y advertir sobre posibles situaciones críticas en la ventilación (Servicio en la nube, Parte 2) (Fig. 2).

La primera parte (Instalación/BIM) incluye la instalación de los componentes físicos (sensores de temperatura, humedad y CO_2, cámaras de video, servidores y brokers) desplegados en el espacio de trabajo y el modelo BIM. El Processing Server (procesado de los datos de Ocupación) y el IoT Broker (gestión de los datos Ambientales) son los módulos Guardian que gestionan los datos de sensores y cámaras de cada Instancia asociada a una zona.

La información sobre la ocupación del espacio de trabajo se extrae de diferentes cámaras de vídeo que enfocan: (1) a cada uno de los accesos al espacio de trabajo, y (2) a las distintas zonas del espacio de trabajo. Las imágenes captadas por las cámaras son procesadas en un servidor que extrae los datos de ocupación sobre el número de personas que se encuentran dentro del espacio de trabajo, y las que habitan cada zona. Este proceso se lleva a cabo a través de algoritmos de IA para el reconocimiento de imágenes que identifican a las personas que entran o salen del espacio de trabajo y a las que se desplazan por las zonas controladas. El IoT Broker se dedica a recopilar los datos ambientales de los sensores instalados, considerando los requisitos de datos y métricas (marca de tiempo, intervalo, frecuencia, etc.). El IoT Gateway agrega los datos de monitoreo (datos ambientales y de ocupación) de cada instancia y los envía a un servicio en la nube (middleware) para Visualización y Simulación con el objetivo de proveer acciones recomendadas al Facility Manager en la toma de decisiones.

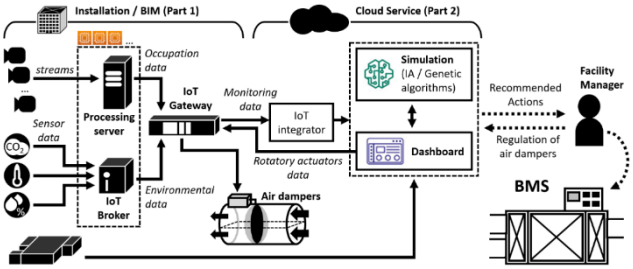


Fig. 2. Arquitectura del SATV.

III. CONCLUSIONES

A pesar de todos los avances realizados hasta la fecha, abordar el problema de la calidad del aire proporcionando una ventilación adecuada en los espacios interiores sigue siendo un tema complejo en la actualidad. El diseño y control del funcionamiento de los sistemas de climatización requiere nuevos enfoques basados en una mayor comprensión de cómo se comporta la ventilación de un edificio para hacer que los espacios interiores sean seguros y confortables sin comprometer la eficiencia energética. El sistema desarrollado contribuye a este objetivo facilitando un uso más adaptado de los sistemas de climatización en espacios de trabajo diáfanos. La investigación realizada pretende obtener un resultado replicable en otros espacios de trabajo, siguiendo la misma filosofía y arquitectura del sistema, también a través del aprendizaje y la búsqueda de patrones.

AGRADECIMIENTOS

La investigación reportada en este artículo se enmarca en el proyecto “GEMELO DIGITAL – Gestión de Smart Building para la era post COVID-19”, que ha recibido financiación de la Agencia para la Competitividad Empresarial de la Generalitat de Catalunya (ACCIO), a través de la subvención ACE012/20/000049. El proyecto ha contado con el apoyo de diferentes miembros del Clúster de Tecnologías Inteligentes para Ciudades, Edificios e Industria: Secartys, SmartTech, Loxone, Siemens, Innova IT, FUNITEC, CT Solutions Group, MSI Studio, NOUMENA, PROSUME y TECNALIA).

REFERENCIAS

[1] [1] Shu Tang, Dennis R. Shelden, Charles M. Eastman, Pardis Pishdad-Bozorgi, y Xinghua Gao. "A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends". *Automation in Construction*, vol. 101, pp. 127-139. 2019. https://doi.org/10.1016/j.autcon.2019.01.020

[2] Christoph Nytsch-Geusen, Werner Kaul, Jörg Rädler, Visvesh Shenoy, Pruthviraj Balekai. "The Digital Twin as a Base for the Design of Building Control Strategies". In *Proceedings of the 16th IBPSA International Conference*, pp. 4141-4148. Rome, Italy. 2019. https://doi.org/10.26868/25222708.2019.210389

[3] Ioannis Brilakis, Yuandong Pan, André Borrmann, Hermann-Georg Mayer, Fabian Rhein, Catharina Vos, Eva Pettinato, y Sigrid Wagner. "Built Environment Digital Twining". In *International Workshop on Built Environment Digital Twinning presented by TUM Institute for Advanced Study and Siemens AG*. Munich, Germany. 2019. Available in: https://publications.cms.bgu.tum.de/reports/2020_Brilakis_BuiltEnvDT.pdf

[4] Agustín Zaballos, Alan Briones, Alba Massa, Pol Centelles, y Víctor Caballero "A smart campus' digital twin for sustainable comfort monitoring". *Sustainability*, 12.21, 9196, 2020. https://doi.org/10.3390/su12219196

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Midiendo la detección temprana de anomalías

Manuel López-Vizcaíno, Francisco J. Novoa, Diego Fernández, Fidel Cacheda
Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC),
Departamento de Ciencias de la Computación y Tecnologías de la Información,
Universidade de A Coruña
Campus de Elviña s/n C.P. 15071 A Coruña
manuel.fernandezl, francisco.javier.novoa, diego.fernandez, fidel.cacheda @udc.es

***Palabras Clave*—early detection, machine learning, time-aware metrics, classification algorithms, network security, social network services.**

I. INTRODUCCIÓN

La detección temprana es un asunto de creciente importancia en múltiples dominios como la seguridad de las redes de comunicaciones, la detección de problemas de salud en redes sociales o la perdición meteorológica relacionada con desastres naturales. En todos estos ámbitos no resulta suficiente realizar una buena decisión sino que además esta debe ser tomada en un tiempo adecuado.

Para asegurar que los sistemas definidos tienen esto en consideración, las métricas utilizadas para su evaluación deben penalizar tanto las decisiones incorrectas como las que se realizan tarde. Por ejemplo, en un análisis de textos publicados en una red social para la detección de depresión entre las personas usuarias de la misma, resulta de interés detectar lo antes posible esta problemática para minimizar posibles consecuencias derivadas de esta situación.

Para ello se presenta la métrica Time Aware Precision (TaP) [1], que ha sido evaluada bajo dos paradigmas de evaluación. Por lotes y en streaming, este segundo se aproxima más al funcionamiento real de sistemas de detección, ya que para cada entidad evaluada, cada item o elemento procesado es añadido de manera individual en lugar de en grupos que representen a un porcentaje del volumen total de items.

En cuanto a los conjuntos de datos utilizados par los experimentos, se han seleccionado tres de diferente naturaleza. Por una parte un conjunto de datos relativo a la detección de depresión en redes sociales (2017 (eRisk 2017 [2]), otro para la detección de ataques en redes de comunicaciones, en particular el escaneo de sistemas operativos (OS Scan Attack, Kitsune [3]) y por último uno relativo a la predicción de inundaciones basadas en datos meteorológicos (extraído de la NOAA [4]).

Para la evaluación de la nueva métrica presentada se ha seguido una aproximación en dos fases, la primera consistente en la evaluación mediante modelos sintéticos: Oracle, Elcaro, Positive, Negative y Random, que proporcionan correspondientemente siempre el valor correcto, incorrecto, positivo, negativo o aleatorio para un número determinado de items procesados. A continuación, se realizaron experimentos para estudiar su comportamiento en la evaluación de modelos de detección basados en aprendizaje máquina (Machine Learning, ML), en particular: LinearSVC, Extra Tree, Ada Boost, Random Forest y Logistic Regression.

Los resultados obtenidos permiten validar la métrica presentada frente a métricas no dependientes del tiempo y a los problemas en otras conscientes del tiempo como Early Risk Detection Error (ERDE) [2] o F-latency [5].

AGRADECIMIENTOS

This work was supported in part by the Ministry of Economy and Competitiveness of Spain and Fondo Europeo de Desarrollo Regional (FEDER) Funds of the European Union under Project PID2019-111388GB-I00; and in part by the Centro de Investigación de Galicia—Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC) Funded by Xunta de Galicia and the European Union (European Regional Development Fund–Galicia 2014-2020 Program), under Grant ED431G 2019/01.

REFERENCIAS

[1] López-Vizcaíno, M., Novoa, F., Fernández, D. & Cacheda, F. Measuring Early Detection of Anomalies. *IEEE Access*. **10** pp. 127695-127707 (2022)

[2] Losada, D. & Crestani, F. A test collection for research on depression and language use. *International Conference Of The Cross-Language Evaluation Forum For European Languages*. pp. 28-39 (2016)

[3] Mirsky, Y., Doitshman, T., Elovici, Y. & Shabtai, A. Kitsune: an ensemble of autoencoders for online network intrusion detection. *ArXiv Preprint ArXiv:1802.09089*. (2018)

[4] NOAA Storm Events Database 2018. (Centers for Environmental Information NOAA, National, 2019), https://www.ncdc.noaa.gov/stormevents/

[5] Sadeque, F., Xu, D. & Bethard, S. Measuring the latency of depression detection in social media. *Proceedings Of The Eleventh ACM International Conference On Web Search And Data Mining*. pp. 495-503 (2018)

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Tecnologías habilitantes para redes programables y definidas por software: un enfoque sobre control in-band

Elisa Rojas, Juan A. Carral, Isaías Martinez-Yelmo, Jose M. Arco,
Diego Lopez-Pajares, Joaquin Alvarez-Horcajo, David Carrascal
Universidad de Alcalá
Escuela Politécnica Superior, Ctra. Madrid-Barcelona, km. 33,6.
elisa.rojas@uah.es, juanantonio.carral@uah.es, isaias.martinezy@uah.es, josem.arco@uah.es,
diego.lopezp@uah.es, j.alvarez@uah.es, david.carrascal@uah.es

Esta comunicación se centra en la necesidad de profundizar en las tecnologías habilitantes para redes programables y definidas por software, con énfasis particular sobre el control denominado *in-band*, muy poco explorado y con aún notables carencias de cara a los despliegues de red avanzados, como son las redes 5G y futuras 6G

***Palabras Clave*—SDN, *programmable networks*, *in-band control*, IoT, 5G, 6G**

I. RESUMEN

Aunque el ámbito de aplicación de las futuras redes 6G aún está en definición, ciertas tecnologías como son el Internet de las Cosas (del inglés *Internet of Things*, IoT) y la inteligencia artificial (del inglés *Artificial Intelligence*, AI) tendrán un protagonismo innegable.

En este sentido, la compatibilidad de redes programables, definidas por software (del inglés *Software-Defined Networking*, SDN) o *softwarizadas* con dispositivos IoT es clave en los futuros despliegues de red móvil. Sin embargo, este campo está aún relativamente poco explorado, debido a que los dispositivos IoT son más heterogéneos y además suelen estar limitados en potencia computacional y energía, lo que restringe la aplicación directa de tecnologías de softwarización de red sobre ellos. Un claro ejemplo, en el caso del componente clave de las redes 5G denominado como *Multi-Access Edge Computing* (MEC), cuya arquitectura se está actualmente revisando para ser aplicada en este tipo de entornos, dado lugar a lo que se denomina como *constrained* MEC [1].

El objetivo de esta comunicación es presentar el estado del arte de este ámbito, con un enfoque particular en la aplicación de control *in-band*, matizando por qué es relevante este aspecto [2]. A continuación, se presentarán algunos trabajos realizados por el grupo de investigación hasta la fecha, como son ieHDDP [3] e IoTorii [4]. Más concretamente, ieHDDP es un protocolo de encaminamiento y también de control *in-band* para redes limitadas en capacidad o que no soportan el paradigma SDN en su totalidad; mientras que IoTorii es un protocolo de encaminamiento en redes limitadas también, comparable a RPL. La diferencia entre ambos es que el primero utiliza exploración de caminos y generación de árboles en base al camino más rápido, mientras que el segundo genera una jerarquía de etiquetas. Finalmente, se expondrán retos y líneas de investigación futuras.

AGRADECIMIENTOS

Este trabajo ha sido apoyado por la Comunidad de Madrid a través del proyecto MistLETOE-CM (CM/JIN/2021-006), por el Ministerio de Ciencia e Innovación con el proyecto ONENESS (PID2020-116361RA-I00), y por la Universidad de Alcalá gracias a las ayudas "Contratos Predoctorales de Formación de Personal Investigador - FPI-UAH 2022".

REFERENCIAS

[1] E. Rojas, C. Guimaraes, A. d. l. Oliva, C. J. Bernardos, and R. Gazda, "Beyond Multi-access Edge Computing: Essentials to realize a Mobile, Constrained Edge," *IEEE Communications Magazine*, pp. 1–7, 2023.

[2] D. Carrascal, E. Rojas, J. M. Arco, D. Lopez-Pajares, J. Alvarez-Horcajo, and J. A. Carral, "A Comprehensive Survey of In-Band Control in SDN: Challenges and Opportunities," *Electronics*, vol. 12, no. 6, p. 1265, 2023.

[3] J. Alvarez-Horcajo, I. Martinez-Yelmo, E. Rojas, J. A. Carral, and D. Carrascal, "ieHDDP: An Integrated Solution for Topology Discovery and Automatic In-Band Control Channel Establishment for Hybrid SDN Environments," *Symmetry*, vol. 14, no. 4, p. 756, Apr 2022. [Online]. Available: http://dx.doi.org/10.3390/sym14040756

[4] E. Rojas, H. Hosseini, C. Gomez, D. Carrascal, and J. R. Cotrim, "Outperforming RPL with scalable routing based on meaningful MAC addressing," *Ad Hoc Networks*, vol. 114, p. 102433, 2021.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Improving efficiency and security of IIoT using in-network validation of server certificate

Asier Atutxa*, Jasone Astorga*, Marc Barcelo†, Aitor Urbieta†, Eduardo Jacob*
*Department of Communications Engineering, Faculty of Engineering, University of the Basque Country, Alda. Urquijo S/N, 48013 Bilbao, Spain.
{asier.atutxa, jasone.astorga, Eduardo.Jacob}@ehu.eus
†Ikerlan Technology Research Centre, Basque Research and Technology Alliance (BRTA) Po J.M. Arizmendiarrieta 2, Arrasate/Mondragon, 20500 Gipuzkoa, Spain.
{mbarcelo, aurbieta}@ikerlan.es

***Abstract*—This paper presents the design and implementation of an in-network server certificate validation system that offloads server certificate validation of DTLS handshakes from constrained IIoT devices to resource-richer network elements, leveraging data plane programming (DPP). This approach enhances security as it guarantees that a comprehensive server certificate verification is always performed. Additionally, it increases performance as resource-expensive tasks are moved from IIoT devices to a resource-richer network element.**

***Keywords*—DTLS; In-network Computing; IoT; P4**

I. PROPOSED SYSTEM

The system proposed in this paper, shown in Figure 1, aims at offloading time- and resource-consuming tasks of the DTLS handshake from the IIoT devices to the network. The DTLS handshake starts with a *Client Hello*, and is answered with a *Server Hello* containing the server certificate. When the *Server Hello* message containing the certificate is received in the programmable switch, located between both end-points, it is sent to the controller for further processing. At the controller, the server certificate is extracted and its validity period is verified. If the certificate is within its validity period, the process continues to verify the certificate signature, and if it is correct, the validation request is sent to the server through OCSP. If it is successful, the controller sends to the data plane of the switch the same *Certificate* message initially obtained from it, which forwards it to the client. If on the contrary, the verification of the server certificate fails, the controller returns a DTLS alert message specifying the code *bad_certificate* and ends the handshake.

II. RESULTS AND DISCUSSION

The performed tests measure the time incurred by the handshake process and evaluate the impact of two parameters: (1) the signature algorithm and (2) the length of the certificate chain. When all the processes are performed by the IIoT device, the duration of the whole DTLS handshake varies between 150 ms and 300 ms when RSA keys are used and between 150 ms and 200 ms when ECC keys are used. The use of the proposed mechanism where certificate validations are performed at the controller allows to significantly reduce these times between 50% and 65% depending on the specific case. Additionally, the proposed system allows reducing the energy consumption of the IIoT device by about 40%.

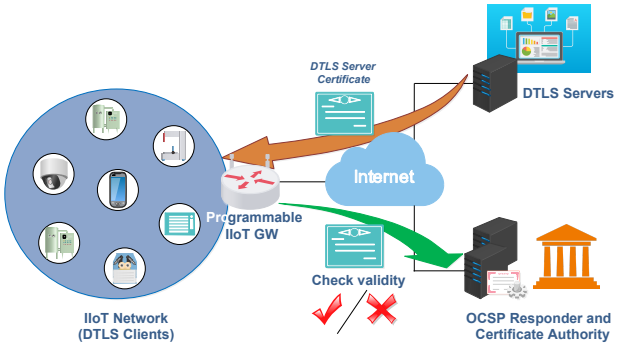


Fig. 1: Proposed architecture of a programmable switch and combination of edge and cloud paradigms.

III. CONCLUSION

The obtained results demonstrate the suitability of the proposed solution for the defined scenario. Additionally, it is noteworthy that the results can be extrapolated to other environments, such as the Internet of Vehicles (IoV) or electronic health systems.

ACKNOWLEDGEMENTS

This work was supported by the TRUE-5G project (PID2019-108713RB-C54/AEI/10.13039/501100011033), the project EGIDA (CER-20191012), and the project REMEDY (KK-2021/00091).

REFERENCES

[1] A. Atutxa, J. Astorga, M. Barcelo, A. Urbieta, E. Jacob, Improving efficiency and security of IIoT communications using in-network validation of server certificate, Computers in Industry, Volume 144, 2023, 103802, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2022.103802.

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Internet of Medical Things y Ciberseguridad: Aplicaciones y retos

Alan Briones, Guiomar Corral, Marc Rivero
Smart Society Research Group
La Salle – Universidad Ramon Llull
08022
Alan.briones@salle.url.edu; Guiomar.corral@salle.url.edu; Marc.Rivero@salle.url.edu

La integración de tecnologías como el IoT, el Big Data y la IA en el sector de la salud ha transformado la atención médica y el acceso a servicios médicos. Sin embargo, esto plantea desafíos en ciberseguridad y protección de datos, ya que la información médica sensible se almacena y transmite electrónicamente. El IoMT, como parte de la eHealth, ha revolucionado la atención médica al conectar dispositivos médicos y recopilar datos de salud en tiempo real. Sin embargo, enfrenta desafíos en interoperabilidad, confiabilidad de datos y seguridad, incluyendo la privacidad del paciente y la integridad de los datos transmitidos. Estos desafíos se plantean abordar en la línea de investigación Blended Network Architectures (Smart Society) para contribuir a una implementación segura y confiable del IoMT en el sector de la salud.

***Palabras Clave*- eHealth, IoMT, Ciberseguridad, Retos**

I. Introducción

La integración de las tecnologías punteras en el sector de la salud, como es el Internet de las Cosas (Internet of Things en inglés), el Big Data o la Inteligencia Artificial, ha transformado la forma en que los profesionales de la salud brindan atención a sus pacientes, así como éstos acceden a los servicios médicos. Este paradigma conocido como eHealth, abarca una amplia gama de aplicaciones, que van desde el uso de registros médicos electrónicos hasta la telemedicina, la monitorización remota de pacientes y las aplicaciones móviles de salud. Sin embargo, esto también plantea desafíos significativos en materia de ciberseguridad y la protección de datos, ya que la información médica sensible se almacena y se transmite electrónicamente. La privacidad del paciente y la confidencialidad de los datos deben ser salvaguardadas para garantizar la confianza. Además, la interoperabilidad entre diferentes sistemas y dispositivos también representa un desafío, ya que es necesario asegurar la integridad confiabilidad de los datos compartidos.

Concretamente, el Internet of Medical Things (IoMT) [1], o Internet de las Cosas Médicas, es un componente integral de la eHealth que ha revolucionado aún más la forma en que se brinda la atención médica. El IoMT se refiere a la red interconectada de dispositivos médicos, sensores y sistemas que recopilan, transmiten y analizan datos de salud en tiempo real. Estos dispositivos incluyen desde wearables y monitores de signos vitales hasta dispositivos implantables y equipos médicos conectados.

Este artículo tiene como objetivo presentar las principales aplicaciones y retos del Internet of Medical Things, así como las amenazas a las que se expone en materia de ciberseguridad, siendo este campo de estudio uno de los principales de la línea de investigación Blended Network Architectures dentro del Grupo de Investigación Smart Society.

II. Aplicaciones y retos del IoMT

La integración del IoMT en la atención médica ha abierto nuevas posibilidades en el monitoreo y la gestión de la salud de los pacientes, como es muestra en Fig *1*. Los dispositivos IoMT permiten la recolección continua de datos biométricos, como la frecuencia cardíaca, la presión arterial y los niveles de glucosa, lo que brinda una visión más completa y precisa del estado de salud de los individuos. Además, estos dispositivos pueden transmitir los datos en tiempo real a los profesionales de la salud, lo que facilita una monitorización remota más efectiva, facilitando la toma de decisiones médicas.

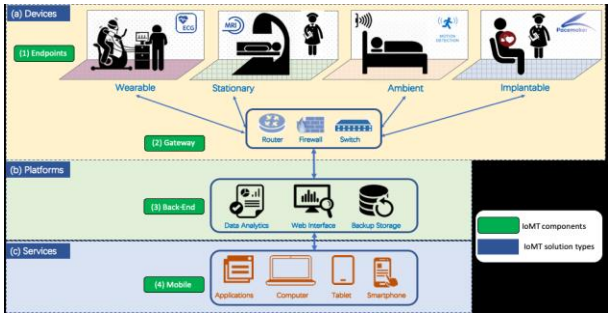


Fig 1. Representación IoMT [2]

Uno de los principales retos es la interoperabilidad y la integración de diferentes sistemas y dispositivos IoMT, ya que se requiere una estandarización efectiva y la capacidad de compartir datos de manera segura entre diferentes plataformas. Otro desafío es la confiabilidad y la precisión de los dispositivos IoMT, ya que es crucial garantizar que los datos recopilados sean exactos y confiables para su posterior uso. El IoMT también enfrenta desafíos éticos y legales, como la responsabilidad de los proveedores de servicios de salud en la gestión y el uso adecuado de los datos de los pacientes, lo cual va ligado a uno de los principales desafíos como es la seguridad de los datos y la protección de la privacidad del paciente.

III. IOMT Y SU SEGURIDAD

Se plantean una serie de retos enfocados a nivel de seguridad del IoMT [2][3]:

- Privacidad y confidencialidad de datos
- Vulnerabilidades de dispositivos
- Seguridad de redes
- Autenticación y control de acceso
- Integridad de datos
- Amenazas internas
- Seguridad de aplicaciones móviles y dispositivos IoMT
- Cumplimiento normativo
- Concienciación y formación en seguridad

La privacidad y la confidencialidad de los datos se han convertido en una preocupación principal, ya que la recopilación y el almacenamiento masivo de información médica sensible plantean riesgos significativos en caso de accesos no autorizados o robo de estos. Además, las vulnerabilidades de los sistemas pueden ser explotadas por atacantes para obtener acceso no autorizado y manipular el funcionamiento de los dispositivos médicos. La seguridad de la red también es un desafío crucial. La autenticación es un aspecto clave para garantizar que solo las entidades autorizadas puedan acceder y modificar los datos médicos. La integridad de los datos es otro reto importante, ya que es necesario garantizar que los datos transmitidos o almacenados no sean alterados o manipulados de manera malintencionada. La seguridad de las aplicaciones móviles y los dispositivos IoMT también pueden representar puntos de entrada para ataques. Cumplir con los requisitos regulatorios es esencial, ya que se deben cumplir los estándares y las regulaciones establecidas para proteger los datos del paciente. Por último, la conciencia y la formación en ciberseguridad son fundamentales para educar a los profesionales de la salud y al personal involucrado en el uso y manejo de la tecnología IoMT, con el fin de minimizar los errores y prácticas inseguras que puedan comprometer la seguridad de los datos médicos.

El avance de la eHealth ha impulsado la implementación del Internet of Medical Things (IoMT), una red interconectada de dispositivos médicos que recopila y transmite datos de salud en tiempo real. Esta integración ha permitido mejoras en la monitorización y la gestión de la salud de los pacientes, proporcionando una visión más precisa de su estado de salud y facilitando la toma de decisiones médicas. Sin embargo, el IoMT enfrenta desafíos significativos, como la interoperabilidad entre sistemas y dispositivos, la confiabilidad y precisión de los datos, así como problemas éticos y legales relacionados con la protección de datos y la privacidad del paciente.

La seguridad del IoMT es un tema crítico, que incluye la protección de la información médica sensible, la prevención de accesos no autorizados y la garantía de la integridad de los datos transmitidos. Es fundamental cumplir con los requisitos regulatorios, promover la conciencia en ciberseguridad y brindar capacitación adecuada para garantizar una implementación segura y confiable del IoMT en el sector de la salud.

AGRADECIMIENTOS

Part of the work of this article was carried out within the framework of the Advanced Training in Health Innovation Knowledge Alliance (ATHIKA) project, funded by the European Commission Erasmus+ Programme—KA2 cooperation for innovation and the exchange of good practices—Knowledge Alliances (601106-EPP-1-2018-1-ES-EPPKA2-KA) and, also, the Cybersecurity Skills Alliance – A New Vision for Europe funded bt the ERASMUS+ programme of the European Union (621701-EPP-1-2020-1-LT-EPPKA2-SSA-B).

REFERENCIAS

[1] S. Vishnu, S. R. J. Ramson and R. Jegan, "Internet of Medical Things (IoMT) - An overview," *2020 5th International Conference on Devices, Circuits and Systems (ICDCS), Coimbatore, India, 2020*, pp. 101-104, doi: 10.1109/ICDCS48716.2020.243558.

[2] Alsubaei, F., Abuhussein, A., Shandilya V., Shiva, S., "IoMT-SAF: Internet of Medical Things Security Assessment Framework" Internet of Things, Volume 8, 2019, 100123, ISSN 2542-6605, doi: 10.1016/j.iot.2019.100123

[3] Papaioannou, M, Karageorgou, M, Mantas, G, et al. "A Survey on Security Threats and Countermeasures in Internet of Medical Things (IoMT)". *Trans Emerging Tel Tech. 2022; 33:e4049*. doi: 10.1002/ett.4049

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Metodología REWIRE para la creación de cursos en ciberseguridad

Alan Briones, Ramon Martin de Pozuelo, Julia Sánchez, Guiomar Corral, Marc Rivero, Agustín Zaballos
Smart Society Research Group
La Salle – Universidad Ramon Llull
08022
Alan.briones@salle.url.edu, ramon.martinpozuelo@salle.url.edu, j.sanchez@salle.url.edu, marc.rivero@salle.url.edu, agustin.zaballos@salle.url.edu

El proyecto REWIRE tiene como objetivo abordar la escasez de expertos en ciberseguridad mejorando la disponibilidad, accesibilidad y calidad de los cursos y certificaciones en ciberseguridad. Para lograr esto, se desarrolló una metodología de selección de cursos que utiliza seis criterios (A-F) para evaluar los niveles educativos, la demanda del mercado laboral, los cursos existentes, la opinión de los interesados, los ejercicios prácticos y las certificaciones. Estos criterios se ponderaron en función de los objetivos del proyecto y se definió una fórmula de puntuación final. Aplicando esta metodología, el proyecto seleccionó los perfiles de Chief Information Security Officer (CISO), Incident Responder, Cyber Threat Intelligence Specialist y Penetration Tester para la creación de cursos. La metodología ofrece un enfoque integral para cerrar la brecha entre la educación en ciberseguridad y las demandas de la industria, asegurando materiales de entrenamiento de alta calidad y amplia accesibilidad para satisfacer las necesidades de los estudiantes y el mercado laboral.

***Palabras Clave*- educación, ciberseguridad, perfiles, cursos**

I. Introducción

La escasez mundial de expertos en ciberseguridad ha alcanzado aproximadamente los 3.4 millones, superando la escasez de 2.7 millones observada en 2021[1]. Esto destaca la necesidad de mejorar la disponibilidad, accesibilidad y calidad de la educación en ciberseguridad para abordar la brecha de habilidades y cerrar la brecha entre las demandas del mercado y el conocimiento de los graduados. En respuesta, la Agencia de la Unión Europea para la Ciberseguridad (ENISA)[2] introdujo el Marco Europeo de Habilidades en Ciberseguridad (ECSF) en 2022[3]. El ECSF condensa las posiciones de ciberseguridad en 12 perfiles de roles y tiene como objetivo establecer una comprensión compartida de los roles esenciales, competencias, habilidades y conocimientos necesarios para los profesionales de ciberseguridad europeos. Apoya el reconocimiento de la experiencia y la creación de programas de capacitación. Se proporciona un manual de usuario para ayudar en la adopción y uso del ECSF, ofreciendo consejos prácticos y ejemplos.

El proyecto REWIRE[4] tiene como objetivo mejorar la educación y capacitación en ciberseguridad mediante el desarrollo del framework propuesto por REWIRE[5], que aborda la escasez de expertos en ciberseguridad y proporciona un punto de referencia común para los interesados. El proyecto considera los desafíos únicos de la educación en ciberseguridad, como la falta de marcos regulatorios, la rápida evolución tecnológica y los currículos y entrenamientos limitados. Trabaja ampliamente en los 12 perfiles del Marco Europeo de Habilidades en Ciberseguridad (ECSF) para mejorar dicho marco al abordar las brechas en competencias, habilidades y conocimientos. El proyecto también se enfoca en la creación de cursos de capacitación y presenta la metodología de selección de cursos para contribuir al desarrollo de programas de educación en ciberseguridad efectivos que satisfagan las necesidades del mercado laboral y los estudiantes.

La metodología para seleccionar perfiles ocupacionales en la educación en ciberseguridad desempeña un papel crucial en cerrar la brecha entre el estado actual de la educación y las demandas del mercado laboral. Los autores proponen un método de selección multicriterio que se centra en tres objetivos principales. El primer objetivo es garantizar materiales de capacitación de alta calidad que brinden información completa y actualizada. El segundo objetivo es alinear los cursos con las necesidades actuales y futuras de la industria de ciberseguridad. El tercer objetivo es hacer que los cursos sean accesibles para una amplia gama de participantes. Para lograr estos objetivos, se utilizan seis criterios en la metodología de selección de cursos, que incluyen niveles educativos, demanda del mercado laboral, disponibilidad de cursos existentes, aportes de los interesados, ejercicios

prácticos y certificación. El artículo proporciona una explicación detallada de cada criterio.

II. Criterios de la metodología REWIRE

El Criterio A en la metodología de selección de cursos se centra en evaluar los niveles educativos de los Perfiles Ocupacionales de ENISA. El propósito de este criterio es asegurarse de que los materiales de capacitación se ajusten a los diferentes niveles definidos en el Marco Europeo de Cualificaciones (EQF, por sus siglas en inglés). Al considerar los niveles EQF, la metodología tiene como objetivo proporcionar capacitación adecuada y relevante para los estudiantes en diferentes etapas educativas, acomodando así a un grupo más amplio de participantes.

El Criterio B en la metodología de selección de cursos tiene como objetivo identificar los perfiles de ENISA que tienen alta demanda en el mercado laboral a nivel de la Unión Europea (UE). El objetivo es enfocarse en desarrollar materiales de capacitación que puedan mejorar las habilidades y conocimientos de profesionales y estudiantes, capacitándolos para ocupar estos puestos de trabajo. Al abordar las necesidades específicas del mercado laboral de ciberseguridad, este criterio busca ampliar el grupo de candidatos calificados y satisfacer eficazmente las demandas de la industria.

El Criterio C en la metodología de selección de cursos se enfoca en evaluar la disponibilidad de currículos y capacitación existentes para cada uno de los Perfiles Ocupacionales de ENISA. El objetivo principal es identificar cualquier brecha en la capacitación en ciberseguridad y desarrollar métodos innovadores y efectivos para abordar estas brechas a través de los cursos propuestos. En lugar de duplicar materiales existentes, se hace hincapié en brindar capacitación para Perfiles Ocupacionales que tienen una disponibilidad limitada de capacitación a nivel de la Unión Europea (UE). Esto garantiza que los cursos proporcionados sean específicos y satisfagan las necesidades particulares dentro del campo de la educación en ciberseguridad.

El Criterio D en la metodología de selección de cursos implica aprovechar la experiencia de diversos interesados para evaluar y calificar los 12 Perfiles Ocupacionales de ENISA. Se solicita a los socios, incluidos académicos, proveedores de educación y capacitación vocacional (VET) y el mercado laboral, que evalúen la importancia de cada perfil en función de su conocimiento y percepción. Este enfoque colaborativo asegura que los perfiles seleccionados estén alineados con las necesidades y requisitos reales de la industria de la ciberseguridad, incorporando perspectivas diversas de los principales interesados.

El Criterio E enfatiza la importancia de los ejercicios prácticos en los cursos de REWIRE, aprovechando el REWIRE Cyber Range basado en la plataforma KYPO Cyber Range desarrollada por la Universidad de Masaryk. Este Cyber Range sirve como una herramienta de capacitación distintiva que facilita actividades prácticas y simulaciones basadas en escenarios del mundo real, en línea con los requisitos del mercado laboral y los métodos de enseñanza contemporáneos. El objetivo es proporcionar una experiencia de capacitación interactiva que cautiva a estudiantes de todos los ámbitos, lo que lleva a una mayor participación y compromiso.

El Criterio F se centra en la evaluación de las certificaciones existentes relevantes para los 12 Perfiles Ocupacionales de ENISA. El objetivo es identificar cualquier brecha o deficiencia en el mercado o en el sector académico/VET y desarrollar un programa de capacitación integral que aborde esas necesidades específicas. Al reconocer las áreas donde faltan certificaciones, el criterio busca proporcionar oportunidades de capacitación específicas que cubran las brechas y cumplan con los requisitos de la industria de la ciberseguridad. Una vez que se han definido los seis criterios (A-F) y sus respectivas distribuciones de puntuación e inputs asociados, se establece la Fórmula de Puntuación final para calcular la puntuación ponderada de cada Perfil de ENISA. La Fórmula de Puntuación tiene en cuenta las puntuaciones asignadas a cada criterio y aplica el peso correspondiente para generar una puntuación integral que refleje la idoneidad e importancia de cada perfil. Este sistema de puntuación permite una evaluación justa y sistemática de los Perfiles de ENISA basada en múltiples criterios, lo que facilita la identificación de perfiles ocupacionales clave en el campo de la ciberseguridad.

WScore = [(A x WA) + (B x WB) + (C x WC) + (D x WD) + (E x WE) + (F x WF)]

• A, B, C, D, E y F son las puntuaciones obtenidas para cada criterio.

• WA, WB, WC, WD, WE y WF son los pesos asignados a cada criterio.

III. Perfiles REWIRE

La metodología de selección de cursos empleada en el proyecto REWIRE involucró el uso de seis criterios (A-F) para evaluar diferentes aspectos, como los niveles educativos, la demanda del mercado laboral, los cursos existentes, la opinión de los interesados, los ejercicios prácticos y las certificaciones. Estos criterios se les asignaron pesos basados en los objetivos del proyecto, y se definió una fórmula final de puntuación ponderada para priorizar y seleccionar los perfiles de Chief Information Security Officer (CISO), Cyber Incident Responder, Cyber Threat Intelligence Specialist y Penetration Tester para la creación de cursos.

Al seleccionar estos perfiles, el proyecto REWIRE reconoce su importancia para abordar la brecha de habilidades y satisfacer las demandas del mercado laboral en el campo de la ciberseguridad. Estos perfiles representan roles clave en asegurar la seguridad y resiliencia de los activos digitales de las organizaciones y desempeñan un papel crucial en la lucha contra las amenazas cibernéticas.

IV. Conclusiones

El Marco Europeo de Competencias en Ciberseguridad (ECSF) introducido por ENISA tiene como objetivo definir las tareas, competencias, habilidades y conocimientos necesarios para los profesionales de ciberseguridad en Europa. El proyecto REWIRE se alinea con el ECSF y busca abordar la escasez de expertos en

ciberseguridad mejorando la disponibilidad y calidad de los cursos y certificaciones en ciberseguridad. Han desarrollado una metodología de selección de cursos con seis criterios (A-F) para garantizar materiales de capacitación de alta calidad que satisfagan las necesidades de la industria de ciberseguridad y sean accesibles para una amplia gama de participantes.

Se ha definido la fórmula final de puntuación para la selección de perfiles de ENISA en el proyecto REWIRE, teniendo en cuenta los pesos asignados a cada criterio y los insumos asociados puntuados. Los autores priorizan los perfiles en base a la fórmula de Puntuación Promedio Ponderada (AWScore), lo que lleva a la selección del Chief Information Security Officer (CISO), Cyber Incident Responder, Cyber Threat Intelligence Specialist y Penetration Tester para la creación de cursos. Los siguientes pasos implican definir el plan de estudios para cada perfil, desarrollar los contenidos de capacitación, realizar un programa piloto para su evaluación y realizar los ajustes necesarios antes del lanzamiento final de los cursos.

En general, el proyecto REWIRE tiene como objetivo contribuir a programas efectivos de educación en ciberseguridad que cubran la brecha entre las prácticas educativas actuales y las demandas del mercado laboral. Al seleccionar perfiles clave y utilizar una metodología integral, se esfuerzan por ofrecer capacitación de alta calidad que satisfaga las necesidades de la industria y beneficie a los participantes con diferentes niveles de experiencia en ciberseguridad.

AGRADECIMIENTOS

Part of the work of this article was carried out within the Cybersecurity Skills Alliance – A New Vision for Europe funded bt the ERASMUS+ programme of the European Union (621701-EPP-1-2020-1-LT-EPPKA2-SSA-B).

REFERENCIAS

[1] (ISC)2. 2022. (ISC)2 Cybersecurity Workforce Study. https://www.isc2.org//-/media/ISC2/Research/2022-WorkForce-Study/ISC2-Cybersecurity-Workforce-Study.ashx

[2] ENISA. 2004. European Union Agency for Cybersecurity. https://www.enisa.europa.eu/

[3] ENISA. 2022. European Cybersecurity Skills Framework Role Profiles. https://www.enisa.europa.eu/publications/european-cybersecurity-skillsframework-role-profiles

[4] REWIRE. 2020. REWIRE: Cybersecurity Skills Alliance - A new Vision for Europe. https://rewireproject.eu/

[5] REWIRE. 2022. WP3 Cybersecurity skills Framework. https://rewireproject.eu/wp-content/uploads/2022/11/R3.3.1.-Cybersecurity-Skills-Framework_FINAL.pdf

XVI Jornadas de Ingeniería Telemática
JITEL 2023
La Salle -- Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Reconfigurable and Multiband Antenna Booster Element for IoT Devices with an SP4T

Elena García[1], Aurora Andújar[1], Joan Lluis Pijoan[2], Jaume Anguera[1,2]

[1] Ignion, Alcalde Barnils, 64-68, Mod C, 3rd floor, 08174 Sant Cugat del Vallès

[2] Research Group on Smart Society, La Salle, Universitat Ramon Llull, c/ Quatre Camins 30, 08022 Barcelona

elena.garcia@ignion.io, aurora.andujar@ignion.io, joanlluis.pijoan@salle.url.edu, jaume.anguera@ignion.io

The increase of the IoT (Internet of Things) puts pressure on device's dimensions and energy costs. This is challenging to embed an antenna into such a small device as 70 mm x 65 mm. For this purpose, a further step is taken by optimizing the area required for the integration of a wireless transmission system using the development of smart adaptive networks so that a single antenna system is capable of transmitting at different frequency bands. To validate the procedure, a reconfigurable architecture with a single SP4T (Single-Pole 4-Throw) operating at 698 MHz – 960 MHz and 1710 MHz – 2170 MHz is designed and built with a non-resonant element called an antenna booster element. This procedure opens the window to facilitate the design of tunable antenna solutions systems for IoT designers requiring a multiband operation.

Keywords **- small antennas, reconfigurable antennas, matching networks, multiband, antenna boosters**

I. INTRODUCTION

The use of complex geometries to design small and multiband antennas is one of the most common methods for wireless device designs. The frequency bands of operation depend on the resonant modes of such antenna. To simplify the design, antenna boosters were proposed, where the frequency bands of operation are controlled by the design of a matching network which is easier and faster than designing an antenna based on complex geometries. An antenna booster element features a non-resonant impedance which can be tuned by designing a proper matching network. The difficulty is maintaining good efficiency in all frequency bands. A reconfigurable antenna booster-based solution is proposed to solve this dilemma. This proposal is convenient for IoT devices where the size of the device is small or the number of frequency bands is large. Thus, challenging to be covered with a passive matching network comprising capacitors and inductors.

Reconfigurable antennas are a reality; design methodologies include PIN diodes, varactors, and RF MEMS switches. PIN diodes have the limitation of two states, on and off. Varactors are limited in the range of capacitance values. In the proposal, an RF MEMS switch has been used as it is more versatile due to the number of states and the possibility of adding different components, such as capacitors or inductors.

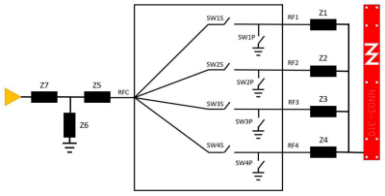


Fig.1 RF SP4T Switch design with all matching networks

As observed in prior art, either the antenna size or the number of switches and lumped components are large. To reduce the antenna element size and the complexity, a reconfigurable architecture has been proposed for operation at 698 MHz – 960 MHz and 1710 MHz – 2170 MHz in a small 70 mm x 65 mm device size. The method relies on using a small antenna booster element (λ/14 @698MHz) with the addition of an RF switch SP4T, including simple matching networks with only one component per branch. The switch allows the common port to be connected to various outputs, increasing flexibility in the impedance-matching process.

After the proposed design process, the measured total efficiency shows competitive efficiency values, and with seven RF switch states, the goal is achieved. The average total efficiency with the implementation design is 32.4% at 698 MHz – 960 MHz and 58.4% at 1710 MHz – 2170 MHz, so the device using NB-IoT passes all the required specifications.

This reconfigurable antenna method with an embedded antenna is easy, simple, and attractive for IoT designers in charge of designing the entire RF chain, including the antenna. Moreover, it solves the problem of having lots of IoT devices at different frequencies. Therefore, this design process simplifies the antenna integration in an IoT design.

More information can be found in the paper from the same authors "Reconfigurable Antenna Booster Element for Multiband Operation in IoT Devices with an SP4T", 17th European Conference on Antennas and Propagation (EuCAP), Florence, 2023, pp. 1-4.