

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS  
INDUSTRIALES Y DE TELECOMUNICACIÓN

UNIVERSIDAD DE CANTABRIA



*Trabajo Fin de Máster*

**INFRAESTRUCTURA IoT PARA LA  
UBICACIÓN DE PERSONAS EN  
TERMINALES AEROPORTUARIAS**  
(IoT Infrastructure for Passengers  
Localization in Airports Terminals)

Para acceder al Título de

***Máster Universitario en  
Ingeniería de Telecomunicación***

Autor: Tomás Llorente Cabello

Octubre - 2018



**Sergio Ortega Alba**  
Jefe Ingeniería y Mantenimiento

Maliaño, 19 de septiembre de 2018

Mediante el presente documento se certifica que el alumno de Máster en Ingeniería de Telecomunicación, D. Tomás Llorente Cabello, estuvo realizando su Proyecto Fin de Máster en colaboración con el personal técnico del Aeropuerto Seve Ballesteros –Santander, en el período comprendido entre marzo y septiembre de 2018.

Dicho Trabajo Fin de Máster, cuyo título es “Infraestructura IoT para la ubicación de personas en terminales aeroportuarias”, ha consistido en la investigación y propuesta de nuevo un sistema basado en infraestructura del tipo IoT que permite recabar información sobre la ubicación de personas en el interior de terminales aeroportuarias, y que posteriormente puede ser procesada para ayudar en la toma de decisiones durante el proceso de gestión de emergencias por parte del gestor aeroportuario.

Desde un punto de vista general se han adquirido conocimientos en el funcionamiento y operativa de las terminales aeroportuarios. Desde un punto de vista investigador, se han adquirido conocimientos teórico-prácticos avanzados en el diseño y funcionamiento de dispositivos IoT, con aplicación directa a la ubicación de personas a través de sus teléfonos móviles.

Finalmente, indicar que el trabajo realizado se considera muy satisfactorio por parte de AENA, al proponer nuevas técnicas que en un futuro cercano pudieran ayudar a la mejora de la gestión de emergencias en aeropuertos.

Atentamente,

Sergio Ortega Alba  
Jefe Ingeniería y Mantenimiento



Ctra. del Aeropuerto s/n

39600 Maliaño  
Cantabria

T + 34 942 202 103 / F +34 942 202 152

sortega@aena.es www.aena.es





## Certificado

Mediante el presente documento se certifica que el alumno de Máster Universitario en Ingeniería de Telecomunicación, D. Tomás Llorente Cabello, estuvo realizando su Trabajo Fin de Máster, cuyo título es "Infraestructura IOT para la ubicación de personas en terminales aeroportuarias", mediante un acuerdo de colaboración entre el Grupo de Ingeniería Telemática de la Universidad de Cantabria, AENA Aeropuerto Seve Ballesteros-Santander y la Delegación en Cantabria de la Asociación Española de Ingenieros de Telecomunicación (AEITCAN)

Dicho Proyecto Trabajo de Fin de Máster se enmarca dentro de la colaboración entre la AEITCAN y los alumnos de último curso del Máster de Ingeniería de Telecomunicación, siendo dirigido el mismo por Ingenieros miembros de la Asociación: Luis Muñoz (Catedrático de la Universidad de Cantabria) y Sergio Ortega Alba (Jefe Ingeniería y Mantenimiento AENA).

Lo cual certifico en Santander a 19 de septiembre de 2018,

Eduardo Artal  
Delegado AEIT-Cantabria

Edificio Sngeniería de  
Telecomunicación-  
Plaza de fa Cienciay s/n  
39005 Santander





## ***Agradecimientos***

*En primer lugar, me gustaría dedicar este trabajo a todas las personas que me han ayudado a llegar hasta este punto.*

*A mi familia, quienes han vivido mis momentos de agobio y me han ayudado a alcanzar cada uno de mis logros.*

*A Los Pollinos, por todos estos años juntos, con nuestro particular humor, pero siempre ahí cuando hace falta.*

*A mis compañeros de clase, quienes han compartido conmigo lo mejor y lo peor de la universidad, y de los que me llevo grandes amigos.*

*A Luis Muñoz, por confiar en mí y permitirme llevar a cabo este proyecto.*

*A Sergio Ortega, por darme la oportunidad de realizar un proyecto en colaboración con AENA.*

*A Juan Ramón Santana, con quien he compartido la mayor parte de este trabajo, y sin el que no hubiera sido posible llevarlo a cabo.*

---

## ÍNDICE GENERAL

---

<b><i>I. Introducción y Objetivos</i></b> .....	<b>1</b>
<b><i>II. Breve Descripción del estado del arte</i></b> .....	<b>4</b>
<b>1. Técnicas de Posicionamiento</b> .....	<b>6</b>
1.1. Ángulo de Llegada .....	6
1.2. Tiempo de Llegada.....	7
1.3. Potencia Recibida .....	8
1.4. Identificación de Célula .....	8
1.5. Fingerprint .....	9
<b>2. Algoritmos de Clasificación</b> .....	<b>9</b>
2.1. Modelos de Regresión .....	10
2.2. Clasificador Bayesiano Ingenuo .....	13
<b>3. Productos Comerciales</b> .....	<b>13</b>
3.1. Situm Indoor Positioning .....	14
3.2. Indoor Survey.....	14
3.3. Google Location Services .....	14
3.4. Otros Productos.....	15
<b>4. Marco del Trabajo</b> .....	<b>15</b>
<b><i>III. Desarrollo de un piloto para la detección de concentración de personas en las terminales aeroportuarias</i></b> .....	<b>19</b>
<b>1. Descripción del Hardware de Detección</b> .....	<b>19</b>
<b>2. Algoritmo de Discriminación Dentro/Fuera</b> .....	<b>33</b>
<b>3. Presupuesto del Proyecto</b> .....	<b>40</b>
<b><i>IV. Resultados</i></b> .....	<b>42</b>
<b><i>V. Conclusiones y Lineas Futuras</i></b> .....	<b>55</b>
<b><i>VI. Referencias</i></b> .....	<b>57</b>



---

## ÍNDICE DE FIGURAS

---

FIGURA 1 – INGRESOS MUNDIALES DERIVADOS DEL USO DE LA INTELIGENCIA ARTIFICIAL .....	1
FIGURA 2 – COMPARATIVA DE DISPOSITIVOS IOT Y POBLACIÓN MUNDIAL .....	2
FIGURA 3 – EVOLUCIÓN DEL MERCADO DE WEREABLES .....	4
FIGURA 4 – ESTABLECIMIENTO DE CONEXIÓN EN EL PROTOCOLO WI-FI [9].....	5
FIGURA 5 – ESTRUCTURA DE IBEACON Y EDDYSTONE BEACON.....	6
FIGURA 6 – EJEMPLO DE MEDIDA DE AOA.....	7
FIGURA 7 – EJEMPLO DE APLICACIÓN DE TOA. ....	7
FIGURA 8 – ESQUEMA DE APLICACIÓN DE LA IDENTIFICACIÓN DE CÉLULA. ....	8
FIGURA 9 – ESQUEMA DE FINGERPRINT POSITIONING.....	9
FIGURA 10 – AEROPUERTO SEVE BALLESTEROS-SANTANDER .....	15
FIGURA 11 – PARTES DE UN AEROPUERTO.....	16
FIGURA 12 – FOTOGRAFÍA DEL LADO AIRE DEL AEROPUERTO SEVE BALLESTEROS-SANTANDER .....	17
FIGURA 13 – NÚMERO DE PASAJEROS ANUALES DEL AEROPUERTO SEVE BALLESTEROS-SANTANDER.....	18
FIGURA 14 – NÚMERO DE OPERACIONES ANUALES DEL AEROPUERTO SEVE BALLESTEROS-SANTANDER.....	18
FIGURA 15 – FOTOGRAFÍA SUPERIOR DE RASPBERRY PI 3 MODEL B+ .....	20
FIGURA 16 – DESPLIEGUE DE RP, CON SEGUNDA INTERFAZ Y ALIMENTACIÓN A TRAVÉS DE POE .....	20
FIGURA 17 – INSTALACIÓN DE LOS DOS ENCAMINADORES MIKROTIK .....	21
FIGURA 18 – ESQUEMA DEL DESPLIEGUE DE LOS DISPOSITIVOS .....	21
FIGURA 19 – CAPTURA DEL FUNCIONAMIENTO DE LA HERRAMIENTA HCITool .....	22
FIGURA 20 – CAPTURA DEL FUNCIONAMIENTO DE LA HERRAMIENTA BTMON.....	23
FIGURA 21 – ESQUEMA DEL DESPLIEGUE EN EL LABORATORIO .....	24
FIGURA 22 – DESPLIEGUE DE RP Y SENSORES DE TEMPERATURA EN EL LABORATORIO .....	26
FIGURA 23 – COMPARATIVA DE LA EVOLUCIÓN DE LA TEMPERATURA EN CADA RP .....	26
FIGURA 24 – COMPARATIVA DE OCUPACIÓN DEL PROCESADOR EN CADA RP .....	27
FIGURA 25 – EJEMPLO GRAFICA CAJAS Y BIGOTES. ....	28
FIGURA 26 – COMPARATIVA DE LA FRECUENCIA DE ENVÍO DE LOS BEACONS EN DIFERENTES ESCENARIOS POSIBLES (IZQUIERDA: TODOS LOS VALORES – DERECHA: SOLO VALORES TÍPICOS) .....	30
FIGURA 27 – COMPARATIVA DE LA FRECUENCIA DE ENVÍO DE LOS BEACONS EN DIFERENTES ESCENARIOS POSIBLES DURANTE LOS PRIMEROS 30 MIN (IZQUIERDA: TODOS LOS VALORES – DERECHA: SOLO VALORES TÍPICOS) .....	32
FIGURA 28 – DESPLIEGUE DE DISPOSITIVOS EN LA TERMINAL.....	33
FIGURA 29 – POSICIONES DE LAS PRIMERAS MEDIDAS REALIZADAS EN LA TERMINAL.....	34
FIGURA 30 – SALA EN LA QUE SE REALIZARON LAS MEDIDAS.....	35
FIGURA 31 – DATOS CAPTURADOS EN LAS POSICIONES 1, 4 Y 6. ....	36
FIGURA 32 – DIVISIÓN DE LA TERMINAL.....	37
FIGURA 33 – MATRIZ DE CORRELACIÓN DE LOS DATOS DE ENTRADA AGRUPADOS POR NÚMERO DE SECUENCIA .....	38
FIGURA 34 – MATRIZ DE CORRELACIÓN DE LOS DATOS DE ENTRADA AGRUPADOS POR SEGUNDO.....	38
FIGURA 35 – SALA DE RECOGIDA DE EQUIPAJES .....	39
FIGURA 36 – RESULTADOS DEL CONTEO DE PERSONAS.....	40
FIGURA 37 – PRESUPUESTO DEL EQUIPAMIENTO .....	40
FIGURA 38 – PRESUPUESTO REFERENTE AL DESARROLLO DEL PROYECTO (PROTOTIPO).....	41
FIGURA 39 – PORCENTAJE DE ACIERTO DEL ENTRENAMIENTO DENTRO/FUERA CON LOS DATOS AGRUPADOS POR SEGUNDO. .....	42
FIGURA 40 – PORCENTAJE DE ACIERTO DEL ENTRENAMIENTO DENTRO/FUERA CON LOS DATOS AGRUPADOS POR SECUENCIA. .....	42
FIGURA 41 – PORCENTAJE DE ACIERTO DEL ENTRENAMIENTO CON 3 SALAS CON LOS DATOS AGRUPADOS POR SEGUNDO..	42
FIGURA 42 – PORCENTAJE DE ACIERTO DEL ENTRENAMIENTO CON 3 SALAS CON LOS DATOS AGRUPADOS POR SECUENCIA.	42
FIGURA 43 – MATRICES DE CONFUSIÓN DE LOS CLASIFICADORES DENTRO/FUERA ENTRENADOS CON LOS DATOS AGRUPADOS POR SEGUNDO. (SUP. IZQ.: NB [dBm], SUP. DCHA.: NB [mW], INF. IZQ: RL [dBm], INF. DCHA.: RL [mW]) ...	44
FIGURA 44 – MATRICES DE CONFUSIÓN DE LOS CLASIFICADORES DENTRO/FUERA ENTRENADOS CON LOS DATOS AGRUPADOS POR SECUENCIA. (SUP. IZQ.: NB [dBm], SUP. DCHA.: NB [mW], INF. IZQ: RL [dBm], INF. DCHA.: RL [mW]) .	45
FIGURA 45 – MATRICES DE CONFUSIÓN DE LOS CLASIFICADORES ENTRENADOS CON 3 SALAS CON LOS DATOS AGRUPADOS POR SEGUNDO. (SUP. IZQ.: NB [dBm], SUP. DCHA.: NB [mW], INF. IZQ: RL [dBm], INF. DCHA.: RL [mW]) ...	46

FIGURA 46 – MATRICES DE CONFUSIÓN DE LOS CLASIFICADORES ENTRENADOS CON 3 SALAS CON LOS DATOS AGRUPADOS POR SECUENCIA. (SUP. IZQ.: NB [dBm], SUP. DCHA.: NB [mW], INF. IZQ: RL [dBm], INF. DCHA.: RL [mW])	.47
FIGURA 47 – COMPARATIVA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO CON LOS DATOS AGRUPADOS POR SEGUNDO (SIN ELIMINAR FILAS)	48
FIGURA 48 – COMPARATIVA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO CON LOS DATOS AGRUPADOS POR SEGUNDO (ELIMINANDO FILAS)	49
FIGURA 49 – COMPARATIVA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO CON LOS DATOS AGRUPADOS POR SECUENCIA (SIN ELIMINAR FILAS)	50
FIGURA 50 – COMPARATIVA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO CON LOS DATOS AGRUPADOS POR SECUENCIA (ELIMINANDO FILAS)	51
FIGURA 51 – COMPARATIVA NUMÉRICA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO PARA EL PRIMER VUELO CON LOS DATOS AGRUPADOS POR SEGUNDO (SIN ELIMINAR FILAS)	53
FIGURA 52 – COMPARATIVA NUMÉRICA DEL CONTEO DE PERSONAS MANUAL Y ESTIMADO PARA EL SEGUNDO VUELO CON LOS DATOS AGRUPADOS POR SEGUNDO (SIN ELIMINAR FILAS)	54

---

## LISTA DE ACRÓNIMOS

---

### **A**

AoA: Angle of Arrival, 6, 7  
AP: Access Point, 9

### **B**

BLE: Bluetooth Low Energy, 5

### **G**

GPS: Global Positioning System, 1, 2, 13, 14, 15

### **I**

IA: Inteligencia Artificial, 1, 2, 3  
IATA: International Air Transport Association, 17  
IoT: Internet of Things, X, 2, 3, 15, 55

### **M**

MAC: Media Access Control, 23, 24, 37

### **N**

NB: Naive Bayes, 43, 44, 45, 46, 47

### **O**

OR: Odds-Ratio, 12

### **P**

PoE: Power over Ethernet, 19, 20, 21

### **R**

RP: Raspberry Pi, 20, 21, 23, 24, 25, 26, 27, 32  
RSSI: Received Signal Strength Indicator, 23

### **T**

TDoA: Time Difference of Arrival, 8  
ToA: Time of Arrival, 7, 8

---

## RESUMEN

---

De la mano de la colaboración entre AENA y el Grupo de Ingeniería Telemática de la Universidad de Cantabria, en una fase en la evolución de la tecnología, donde el teléfono inteligente, la inteligencia artificial y la Internet de las Cosas se encuentran en continua evolución, surge este proyecto que persigue ubicar, en tiempo real, las concentraciones de viajeros y usuarios en el interior de una terminal aeroportuaria.

A lo largo del trabajo se realiza el análisis, diseño y posterior validación del uso de dispositivos IoT en el interior de ciertas zonas aeroportuarias, con el objetivo último de poder realizar una actuación inteligente y eficaz en caso de darse una situación de emergencia o anomalía.

Ello ha conllevado abordar diferentes aspectos del ciclo de vida de un proyecto de ingeniería. En particular, la elección de los componentes hardware y la implementación del algoritmo de detección, así como el desarrollo de la componente software encargada de dotar al sistema de la inteligencia que le permita estimar el número de personas presentes. Todo ello con las restricciones impuestas por la necesidad de operar en tiempo real.

---

## **ABSTRACT**

---

From the collaboration between AENA and the Network Planning and Mobile Communications Group of University of Cantabria, in a stage of technology evolution, in which the smartphone, artificial intelligence and Internet of Things (IoT) are constantly evolving, it rises this project that copes with the real time localization of passenger crowds inside an airport terminal.

This project addresses the design, deployment and validation of the use of Internet of Things as the enabling technology supporting passenger's localization inside airport terminals. The final aim is to provide real time monitoring tools which enable to fulfil the needs linked to emergency situations.

This work has dealt with the different issues belonging to the life cycle of an engineering project. In particular, it has addressed the analysis of the best hardware technology fulfilling the requirements imposed by the use case. Furthermore, an in depth analysis of the detection and decision algorithms has been carried out. Last but not least, the integration of such developments as well as its validation have also been included as part of this project.

---

## I. INTRODUCCIÓN Y OBJETIVOS

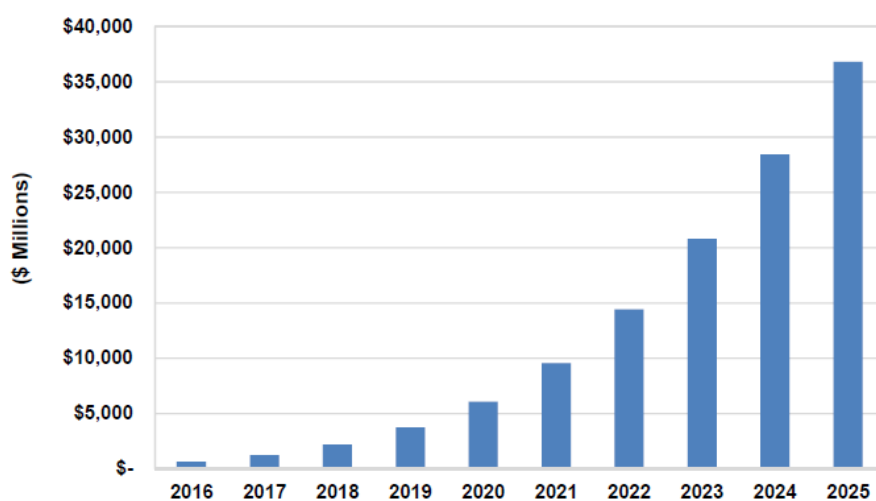
---

En la sociedad actual, el uso del teléfono inteligente ha alcanzado una relevancia tal que ha hecho de él una herramienta indispensable, siendo cada vez más las utilidades disponibles en estos computadores de bolsillo. Al igual que la cámara, el Global Positioning System (GPS) o el reproductor mp3 se hicieron un hueco entre todas las opciones que brindaban los teléfonos inteligentes, en la época actual es la inteligencia artificial (IA) la que trata de consolidarse como una herramienta indispensable.

La IA no solo pretende ganarse un lugar dentro del mundo de los teléfonos inteligentes, si no que ya está presente en multitud de soluciones tecnológicas como ordenadores, cámaras y televisiones. Ampliando un poco su campo de acción, la IA cada vez tiene más relevancia en la vida diaria. Cuando se hace una búsqueda en Google, se realiza una compra en una plataforma online, o simplemente se accede a una red social, hay una entidad dotada de IA procesando todos esos datos para adaptar la experiencia de uso de cada dispositivo al usuario.

Tal es el potencial de esta tecnología, que multitud de empresas y universidades de todo el mundo destinan una parte importante de sus presupuestos a la investigación de diferentes propuestas basadas en la IA. Y es que, como se ha comentado en esta introducción, tiene un gran campo de aplicación del que se pueden obtener herramientas desde la predicción de tráfico, hasta el control inteligente de la calefacción en un hogar, pasando por un sinfín de opciones como el reconocimiento de voz, predicción de búsquedas o la detección de rostros.

**Chart 1.1 Artificial Intelligence Revenue, World Markets: 2016-2025**



(Source: Tractica)

*Figura 1 – Ingresos mundiales derivados del uso de la inteligencia artificial<sup>1</sup>*

---

<sup>1</sup> Fuente: <https://www.top500.org/news/market-for-artificial-intelligence-projected-to-hit-36-billion-by-2025/>

Junto al auge de la IA, en la última década se ha producido otra revolución en el mundo tecnológico, derivado del paradigma de la *Internet de las Cosas* [1] o, por sus siglas en inglés, *Internet of Things* (IoT). Tal ha sido su impacto que, como muestra la Figura 2, a mediados del año 2006, ya había más dispositivos pertenecientes a este conjunto que personas en el mundo, estimando que para el año 2020 esta cifra superará en más de 6,5 veces la población mundial. Este despliegue masivo de dispositivos, con la capacidad de captar información en tiempo real y hacer uso de esta, ha propulsado el desarrollo de la IA. En especial, una de sus ramas conocida como *Machine Learning* [2], la cual engloba técnicas que permiten que, a partir de un conjunto de datos las máquinas con capacidades de cómputo creen conocimiento. En particular, en este proyecto se empleará la IA con objeto de posicionar viajeros y usuarios en el interior de edificios aeroportuarios.

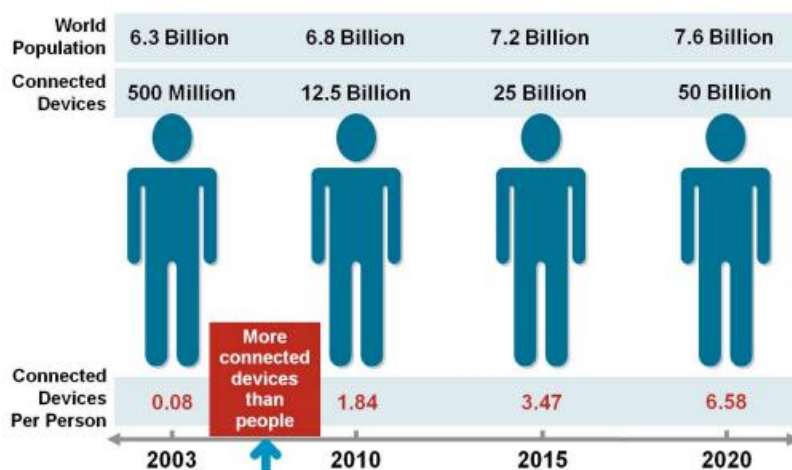


Figura 2 – Comparativa de dispositivos IoT y población mundial<sup>2</sup>

El mundo de los sistemas de posicionamiento en interiores es uno de los temas a desarrollar en el campo de las telecomunicaciones. Si bien, existen herramientas de posicionamiento como el GPS, Glonass [3] o la alternativa desarrollada por la Unión Europea, Galileo [4], las cuales consiguen una gran precisión en la ubicación de dispositivos en el exterior, estas se ven inutilizadas cuando tratan de ser aplicadas en interiores, debido a la atenuación de la señal. Otras alternativas utilizadas en el exterior como el posicionamiento basado en la triangulación [5] a través de las antenas de comunicación sufren la misma suerte cuando el dispositivo se encuentra dentro de un edificio.

Actualmente hay soluciones basadas en distintas tecnologías que proponen varias alternativas. Existen proyectos que utilizan etiquetas del tipo Radio Frequency Identification (RFID) [6] para proporcionar el posicionamiento, otros basados en las tecnologías Wi-Fi o Bluetooth, y así, numerosas combinaciones de las diferentes alternativas de las que disponen los teléfonos inteligentes actuales. También hay propuestas que realizan combinaciones de varias de ellas, pero el estado actual de la temática, sin un sistema tan dominante como el uso del GPS en el posicionamiento en exteriores permite

<sup>2</sup>[https://www.cisco.com/c/dam/en\\_us/about/ac79/docs/innov/IoT\\_IBSG\\_0411FIN\\_AL.pdf](https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FIN_AL.pdf)

investigar diferentes vías para lograr estimar la posición de los dispositivos bajo estas complejas condiciones.

A partir de estas premisas surge el objetivo principal de este trabajo. El cual propone la búsqueda de un sistema que englobe tanto las tecnologías de IoT como la IA para conseguir detectar y monitorizar concentraciones de pasajeros y acompañantes en el interior de las terminales aeroportuarias. Dentro de la gestión de emergencias en entornos aeroportuarios, es una cuestión clave para el gestor aeroportuario el conocer el posicionamiento de las personas o concentraciones de las mismas en el interior de la terminal, con objeto de poder tomar decisiones en tiempo real relativas a las mejores y más seguras vías de evacuación para dichas personas en caso de una situación sobrevenida.

La ubicación del proyecto se da en un emplazamiento singular, como es la terminal del Aeropuerto Seve Ballesteros-Santander, cuyas características únicas definirán la línea de trabajo que se seguirá a lo largo del proyecto.

Con estas condiciones de contorno, se propone el diseño y el despliegue de un sistema basado en infraestructura del tipo IoT que permita recabar información en tiempo real de los dispositivos que se encuentran en el aeropuerto, y que posteriormente pueda ser procesada para dar solución a las metas que se explicaran a continuación. Dentro de este objetivo principal, se buscarán soluciones a diferentes problemas.

En primer lugar, se pretende encontrar, un algoritmo capaz de discriminar los usuarios que se encuentren dentro y fuera de la terminal con una precisión elevada. Entre las personas que se encuentren en el interior de esta, se buscará discernir con el mayor acierto posible, entre las diferentes salas que conforman la terminal.

Finalmente, se utilizará el sistema desplegado para realizar un conteo de personas que permita, no solo posicionar a los usuarios detectados, si no alcanzar unos umbrales de acierto elevados en cuanto al número de personas medidas en tiempo real.

En el capítulo siguiente de este trabajo, se abordarán las partes más relevantes en relación con el estado del arte que conforma el fundamento de este proyecto. Se comentarán las principales técnicas de posicionamiento, los algoritmos de clasificación con utilidad para el sistema a desarrollar, y un resumen de algunos productos comerciales y el marco en el que se realiza el trabajo.

En el capítulo III se explicará el proceso de desarrollo del sistema, las diferentes fases hasta alcanzar el despliegue final, y todas las características destacables de los dispositivos instalados en el aeropuerto. Este capítulo concluirá con un resumen del presupuesto necesario para obtener un sistema como el instalado.

En el cuarto capítulo del trabajo se mostrarán los resultados obtenidos de todas las fases comentadas anteriormente, para finalmente presentar las conclusiones y líneas futuras.



## II. BREVE DESCRIPCIÓN DEL ESTADO DEL ARTE

En el ámbito del posicionamiento en interiores existen numerosas técnicas que se basan en tecnologías diversas. Durante el desarrollo de este trabajo se ha prestado especial atención a las que se encuentran relacionadas con Wi-Fi y Bluetooth [7]. La elección de estas dos tecnologías frente a otras se debe a la gran penetración en el mercado de ambas, posibilitando desarrollar un sistema de posicionamiento con un coste reducido y una instalación mínima. Wi-Fi destaca como la principal tecnología en la actualidad en cuanto a comunicaciones inalámbricas se refiere, aunque ante la llegada de los *wereables* que predice la Figura 3, se espera un importante crecimiento del número de dispositivos que utilicen Bluetooth como principal pasarela para comunicarse.

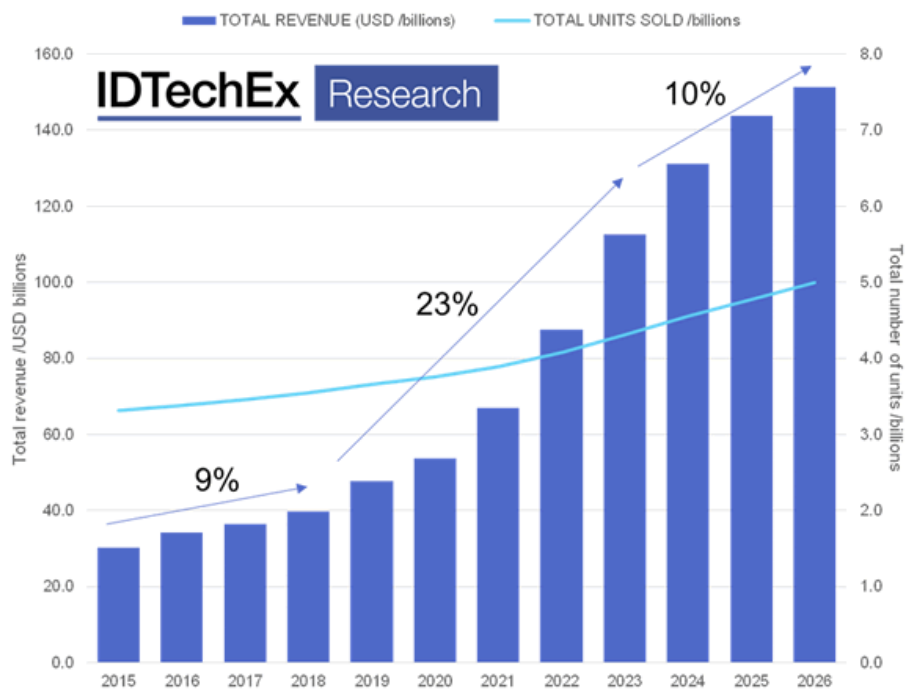


Figura 3 – Evolución del mercado de *wereables*

La aplicación de las diferentes técnicas de posicionamiento es bastante similar para ambas opciones, siendo la principal diferencia el tipo de trama que se captura en cada caso. Cuando se hace uso de la tecnología Wi-Fi, se capturan tramas 802.11 [8], de las que se extrae la información relevante, en función de la técnica aplicada, para poder realizar finalmente una estimación de la posición. Dentro del protocolo 802.11 se diferencian 3 tipos de tramas: gestión, control y datos.

En este trabajo se prestará especial atención a las tramas de gestión, las cuales se utilizan muy especialmente en el proceso de establecimiento de la conexión que se representa en la Figura 4. Este tipo de trama no transporta información de carácter personal de los usuarios, lo que permite respetar las regulaciones relativas a la protección de datos, y se envían de forma continuada en el tiempo.

Como se muestra en la Figura 4, existen dos tipos de búsqueda de puntos de acceso, la pasiva y la activa. En el método pasivo, el dispositivo se mantiene a la escucha de los *beacons* enviados por los puntos de acceso periódicamente, lo que conlleva un consumo energético bastante elevado derivado de la escucha continua. La alternativa a este método

es la búsqueda activa, de esta manera es el dispositivo el que se manifiesta periódicamente mediante el envío de tramas del subtipo *Probe Request*, las cuales se envían en *broadcast* y pueden ser capturadas fácilmente.

El proceso de establecimiento de la conexión concluiría con las fases de autenticación y asociación, las cuales envían diferentes tramas de gestión, pero debido a que este subtipo se envía solamente durante este proceso y no de forma constante, hacen que no sean tan atractivas como las del subtipo *Probe Request*.

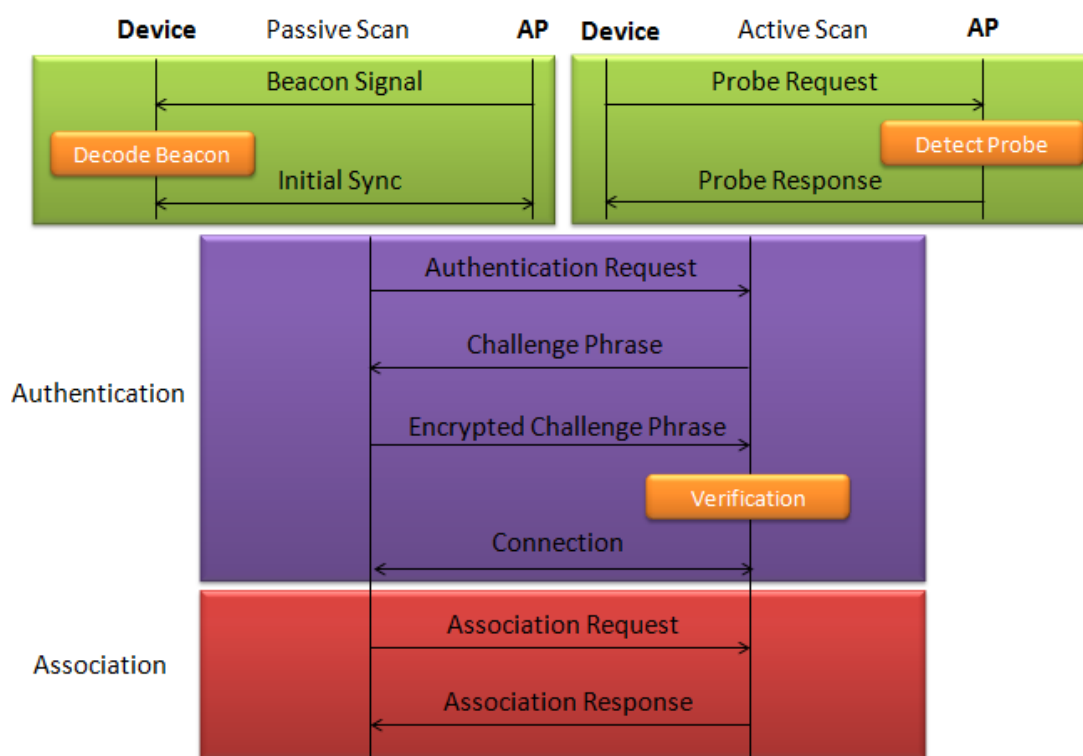


Figura 4 – Establecimiento de conexión en el protocolo Wi-Fi [9]

En el caso del posicionamiento a través de Bluetooth, se utilizan los *beacons* de Bluetooth de bajo consumo, también conocido como Bluetooth Low Energy (BLE), que se implementa a partir del estándar de Bluetooth 4.0. La principal característica de estos es que se envían repetidamente con breves periodos de tiempo entre ellos sin necesidad de un gasto energético tan alto como el que necesita Wi-Fi.

Dentro de las opciones disponibles, las dos alternativas más extendidas en el mercado son el *iBeacon* [10] de Apple y el Eddystone [11] Beacon, con formato abierto, de Google. Como muestra la Figura 5, los *iBeacons* están compuestos por un identificador único universal (UUID) de 128 bits, un número mayor y uno menor de 16 bits cada uno, además de sus correspondientes cabeceras y su campo de potencia de 8 bits. Este tipo de *beacon* tiene un objetivo comercial, y en su especificación se recogen diferentes opciones para enviar los anuncios en *broadcast*. En cuanto a los Eddystone Beacon de Google, su especificación recoge diferentes tipos de trama [12] [13]:

- **Eddystone-UID:** Envía tramas *broadcast* y opacas que pueden ser útiles en el registro de dispositivos. Están formados por un identificador de 16 bytes compuesto por 10

bytes que identifican el *namespace* (grupo de *beacons*) y otros 6 bytes para identificar la estancia (dispositivo dentro del grupo).

- **Eddystone-EID:** Un tipo de *beacon* que varía en el tiempo de modo que puede ser útil durante el proceso de enlace entre dos dispositivos.
- **Eddystone-URL:** Su funcionamiento es parecido al de los iBeacons y su principal propósito es compartir URL.
- **Eddystone-TLM:** Esta trama está destinada a compartir información del estado del dispositivo como carga de la batería, temperatura o el número de paquetes en broadcast enviados.

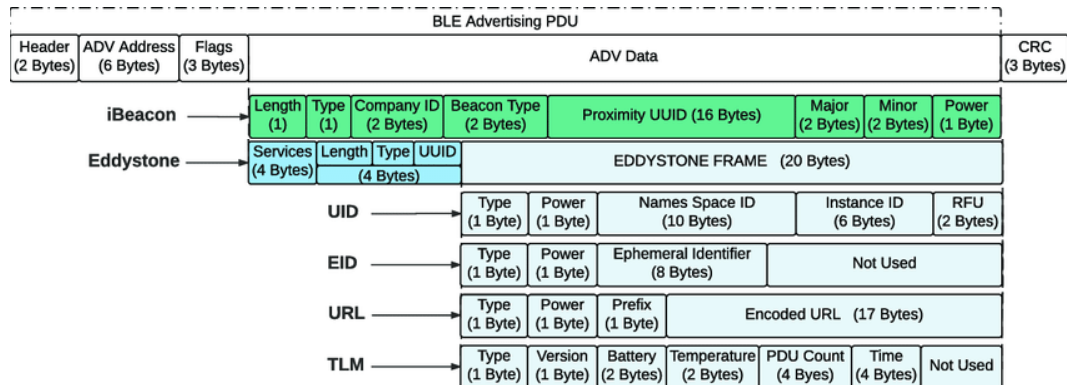


Figura 5 – Estructura de iBeacon y Eddystone Beacon

Una vez constatado el interés en las tecnologías Wi-Fi y Bluetooth, los siguientes apartados del primer capítulo se centrarán principalmente en las diferentes técnicas y algoritmos disponibles con las que llevar a cabo el posicionamiento.

## 1. TÉCNICAS DE POSICIONAMIENTO

En el ámbito del posicionamiento en interiores, existen diferentes técnicas [14] sobre las que se pueden aplicar algoritmos con los que conseguir estimar la posición. En este apartado se recogen algunas de las cuales tienen aplicación tanto en Wi-Fi como en Bluetooth.

### 1.1. ÁNGULO DE LLEGADA

Dentro de las diferentes opciones para obtener la posición de un dispositivo se encuentra la técnica del ángulo de llegada (*Angle of Arrival, AoA*) [15] [16]. Su modo de operación está basado en un conjunto de emisores omnidireccionales (equipos a posicionar) y una estación receptora con un *array* de antenas direccionales que permitirán calcular la posición de este.

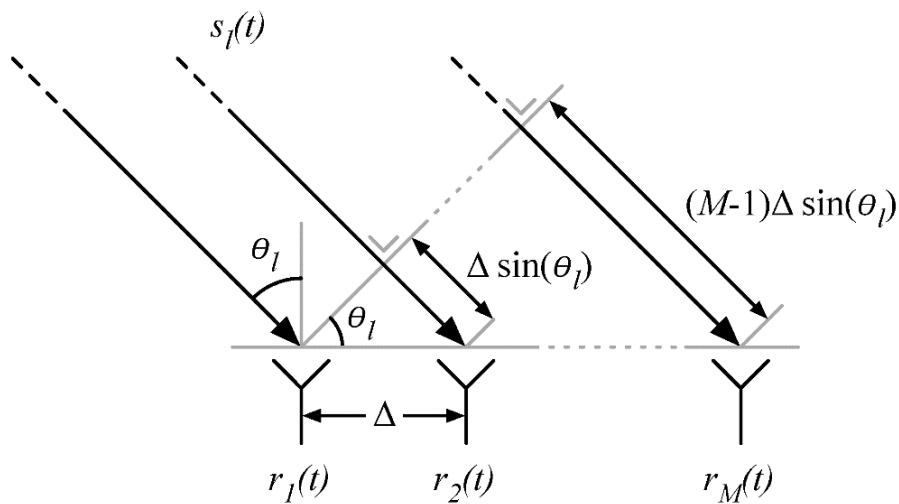


Figura 6 – Ejemplo de medida de AoA

Como se muestra en la Figura 6, el frente de onda llega a las diferentes antenas que conforman el *array* con un cierto desfase entre ellas. A continuación, utilizando este dato, junto a la amplitud de la onda captada, se puede calcular el ángulo de llegada del frente. Por último, utilizando la información aportada por todas las antenas del *array*, es posible estimar la posición de cada uno de los emisores.

## 1.2. TIEMPO DE LLEGADA

En este caso, la técnica del tiempo de llegada (*Time of Arrival, ToA*) [17] utiliza el tiempo de propagación de la señal entre el objeto a posicionar y las estaciones de medida como medida principal en el proceso de estimación. Este dato permite trazar un círculo centrado en torno a cada uno de los receptores de forma que el radio será proporcional al tiempo de propagación. Como se muestra en la Figura 7, a partir de estos círculos, es posible estimar la posición del emisor con determinada precisión.

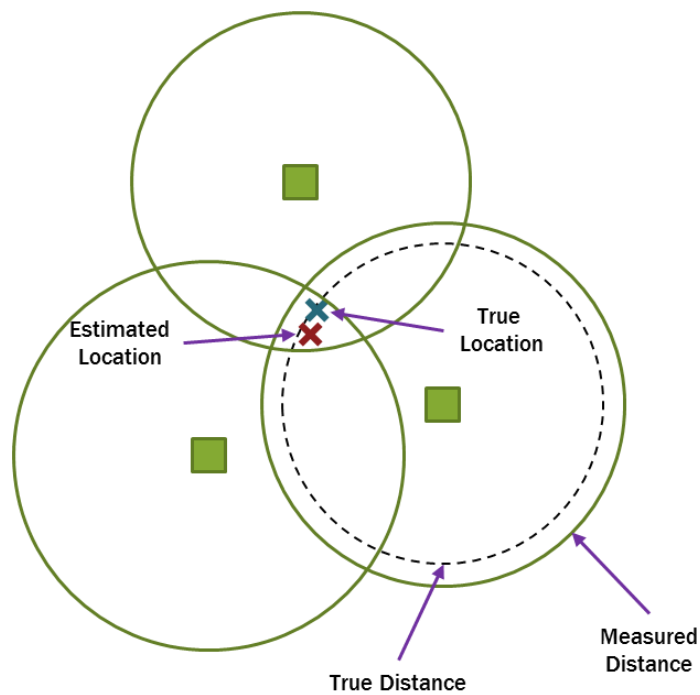


Figura 7 – Ejemplo de aplicación de ToA.

Una variante de esta técnica es la denominada diferencia entre los tiempos de llegada (*Time Difference of Arrival, TDoA*), cuya principal particularidad se encuentra en que, en vez de trazar los círculos en base a los tiempos de propagación de la señal entre el emisor y la estación receptora, se mide la diferencia de tiempos de llegada de los paquetes a cada una de las estaciones de medida. A continuación, de forma similar al ToA se representan unas hipérbolas que permiten aproximar la ubicación del dispositivo a posicionar. Como particularidad de esta técnica, es necesario disponer de 3 o más receptores para poder aplicarla correctamente.

### 1.3. POTENCIA RECIBIDA

Este es uno de los métodos más extendidos entre los sistemas de localización en interiores debido a que su funcionamiento es bastante sencillo. En el primer paso se estima la distancia entre el objeto emisor y la estación receptora a partir de la potencia de señal recibida. Obtenida esta distancia en varias estaciones receptoras, se pueden conseguir varios círculos centrados en cada uno de los receptores con radio la distancia calculada. Finalmente, al igual que se hizo en técnicas anteriores, es posible estimar la posición del objeto como el punto en el que se cortan los diferentes círculos.

Cabe reseñar que la aplicación de este método debe tener en cuenta el estado del canal ya que la potencia recibida puede verse fuertemente afectada por diversos factores que provoquen pérdidas a considerar.

### 1.4. IDENTIFICACIÓN DE CÉLULA

En comparación con las técnicas descritas a lo largo de este capítulo, la identificación de célula (*Cell Identifier*) [18] es un método bastante distinto al resto, y más sencillo de todos. En su despliegue se colocan las diferentes estaciones de medida (células) de forma que su rango de cobertura físico coincida con las celdas lógicas que se quieren utilizar para dividir el terreno. De este modo, los dispositivos que se muevan por el terreno estarán posicionados en todo momento en una de las celdas aplicadas en función de la estación que lo detecte dentro de su rango.

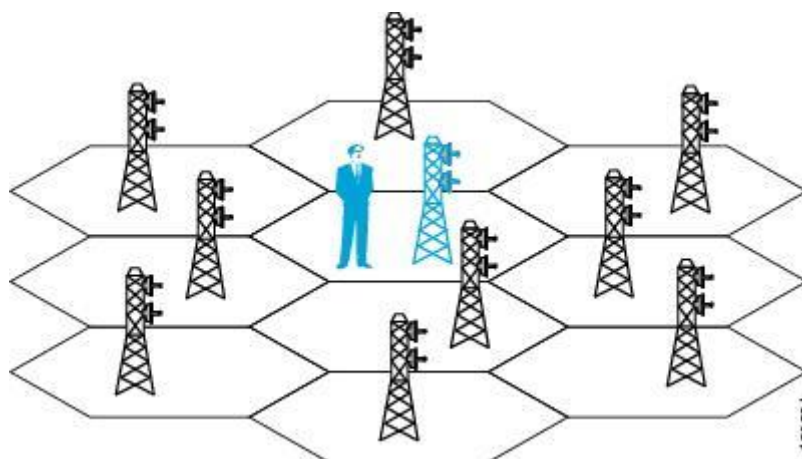


Figura 8 – Esquema de aplicación de la identificación de célula.

Como se puede comprobar, la precisión de este método es notablemente baja, aunque debido a su gran simpleza puede servir como una primera aproximación para combinar con otra de las técnicas descritas en este capítulo.

La técnica Fingerprint [18] [19] es una de las técnicas de posicionamiento basadas en el método de potencia recibida explicado con anterioridad, y comúnmente usado en entornos en los que la propagación con línea de vista no es lo más habitual. A diferencia de las anteriores, la posición no se calcula en base a métodos como la triangulación, sino que se basa en modelos estadísticos.

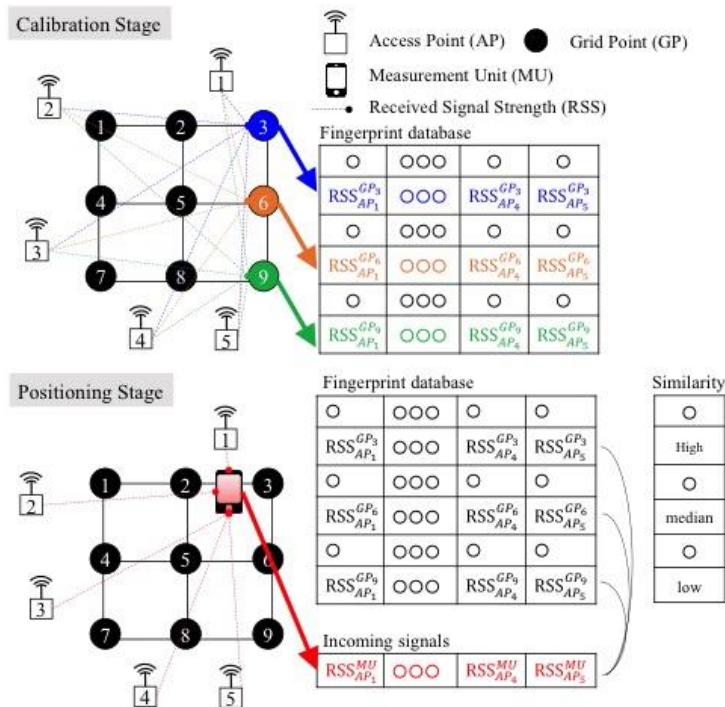


Figura 9 – Esquema de Fingerprint Positioning

Como muestra la Figura 9, el método está compuesto por dos fases: una primera en la que se entrena al sistema y una segunda en la que se estiman las posiciones de los dispositivos. En esta primera fase, el sistema almacena las potencias recibidas en las estaciones de medida desde diferentes dispositivos ubicados en posiciones conocidas previamente, consiguiendo una base de datos desde un gran número de puntos del terreno. Una vez terminada esta etapa, el sistema está listo para comenzar a posicionar. En este proceso se comparan las potencias recibidas en tiempo real en cada uno de los puntos de acceso (*Access Points, APs*) con los datos guardados en la base de datos en la fase de entrenamiento. De este modo se calcula la posición más probable para el dispositivo. Para realizar la aproximación con los datos guardados previamente, se utilizan diferentes métodos, algunos de los cuales se explicarán en el punto 2 de este capítulo.

## 2. ALGORITMOS DE CLASIFICACIÓN

A la hora de desarrollar un sistema de posicionamiento, además de elegir la información que se va a utilizar para obtener los datos que alimentaran al sistema, es necesario definir el algoritmo que, haciendo uso de los datos medidos, marcará la ubicación del objeto que se desea posicionar.

Actualmente existen numerosas soluciones matemáticas que permiten afrontar este problema de diversas formas, pero dadas las características del sistema que se explicarán en el capítulo III de este trabajo, se van a analizar dos soluciones: modelos de regresión logística y clasificadores bayesianos.

## 2.1. MODELOS DE REGRESIÓN

Ante situaciones en las que se desea estimar una variable dependiente de salida ( $Y$ ) a partir de un conjunto de variables independientes de entrada ( $X_1, X_2, X_3, \dots, X_n$ ) los modelos de regresión [20] [21] son una alternativa que plantean una solución más que interesante.

La principal condición para la utilización de estos modelos es que las variables de entrada sean independientes. Cuando estas, además son continuas, es posible introducir su valor de manera directa al sistema o, si se desea, realizando algún tipo de transformación matemática que sea apropiada en cada caso (logaritmo potencia...). En cambio, cuando las variables de entrada son categóricas, con más de dos categorías, es necesario realizar una serie de transformaciones en ellas para poder aplicar el modelo, obteniendo las denominadas variables *dummy*. Estas permiten separar una de  $n$  opciones en  $n - 1$  variables *dummy* distintas que sustituirán a la misma en la entrada al sistema. Un ejemplo clásico que permite facilitar la comprensión de esta transformación es el de la clasificación de fumadores [22]:

Se tiene una variable *FUMA* con 3 categorías que permite discernir entre Fumadores, No Fumadores y Exfumadores. Para poder introducir esta variable al sistema se obtienen sus dos variables *dummy*, de modo que una distinguirá entre Fumadores y No Fumadores y la segunda entre No Fumadores y Exfumadores. Finalmente, se obtiene una tabla como la siguiente:

FUMA	$X_1$	$X_2$
Fumador	1	0
No Fumador	0	1
Exfumador	0	0

*Tabla 1 – Ejemplo variables dummy a partir de variable FUMA*

Realizada la citada transformación, se obtienen dos variables de dos categorías que, sí cumplen los requisitos de las variables de entrada y, por tanto, pueden ser introducidas al sistema.

Además de estas variables *dummy*, también puede ser interesante introducir otras de interacción si se considera que pueden tener relevancia en los resultados de la regresión. Por ejemplo, se pueden tener en cuenta los efectos del tabaco por edad, por género, o incluso todas ellas juntas. El único requisito que se debe cumplir al introducir una variable de interacción es que habrá que introducir también las variables de ordenes inferiores. En el ejemplo anterior, si se introducen los efectos del tabaco por edad y género ( $x * y * z$ ), se deberán introducir también las de segundo orden ( $x * y, x * z, y * z$ ) y las variables independientes ( $x, y, z$ ).

Por otro lado, la selección de las variables que son necesarias se puede realizar de diferentes maneras:

- **Pasos hacia delante (*Forward*):** se basa en ir introduciendo las variables en el modelo, observando si al añadirse, este cumple ciertos requisitos.
- **Pasos hacia atrás (*Backward*):** en este caso se introducen todas las variables y se van eliminando hasta que no pueda suprimirse ninguna más teniendo en cuenta las condiciones fijadas.
- **Por pasos (*Stepwise*):** se trata de una combinación de los dos modelos anteriores en la que se introducen y eliminan variables hasta que no se pueda añadir ni eliminar ninguna.
- **Manual (*Enter*):** en este caso se otorga el poder de decisión a la persona encargada de generar el modelo de regresión. Se introducen todas las variables y se busca la menos significativa, de modo que esta se elimina y se vuelve a iterar hasta que se consigue el mejor modelo posible para las condiciones fijadas.

Al hacer esta selección también es posible encontrar los factores de confusión, es decir, las variables que, al ser eliminadas del modelo, provocan un cambio significativo en los coeficientes del resto de variables. Estas variables son aquellas que, aunque no sean significativas para la variable objetivo o resultado, si lo son respecto al resto y, por tanto, no deberían ser eliminadas del modelo.

Una vez se han preparado las variables para que puedan ser introducidas en el sistema, se debe analizar qué clasificador es el más adecuado, para cada caso, en función de la variable de salida. Si la variable dependiente es continua, se deberá aplicar una regresión lineal, mientras que si esa variable es politómica lo más adecuado es utilizar la regresión logística, especialmente en el caso de las variables dicotómicas, donde es de gran utilidad.

### 2.1.1. Regresión Lineal

El modelo de regresión lineal se basa en una variable de salida dependiente que se obtiene como resultado de la relación lineal de las variables independientes de entrada multiplicadas por un coeficiente que se ajustará en función de su relevancia, además de un coeficiente independiente denominado intercepto.

$$Y = f(x_1, x_2, \dots) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots$$

Para obtener los coeficientes que mejor se adapten al problema, se realiza una fase de entrenamiento con un número suficiente de muestras con los valores de la variable de salida conocidos. Una vez ajustado, los resultados de este modelo se pueden comprobar gracias a la técnica de los mínimos cuadrados, es decir, a la suma de las distancias al cuadrado entre los resultados reales y los estimados por la recta del modelo, siendo el mejor estimador el que minimice esta distancia.

Además, en [21] se mencionan diferentes formas de analizar la bondad del ajuste.

### 2.1.2. Regresión Logística

Cuando se aplica el modelo de regresión logística, al tener una variable politómica, se construye una función basada en la probabilidad de los valores que puede tomar la variable de interés o dependiente, lo que hace que este modelo se base en la técnica de máxima verosimilitud. Esta función puede expresarse de la siguiente manera:



$$Y = \ln\left(\frac{p}{1-p}\right)$$

De este modo, la variable sobre la que se van a realizar los cálculos puede tomar cualquier valor y no está limitado a las categorías de la variable original, lo que permite aplicar métodos de regresión.

Operando, se obtiene un modelo matemático que tiene cierta similitud con el modelo de regresión lineal, ya que este se ajusta en base a un término independiente ( $\theta_0$ ), denominado intercepto, y una serie de coeficientes asociados a cada variable de entrada.

$$p = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x_1 + \theta_2 x_2)}}$$

Al igual que en la regresión lineal, los diferentes coeficientes se ajustan en una primera fase de entrenamiento.

Llegados aquí, los resultados del modelo se evalúan principalmente con el parámetro *odds*, es decir, el cociente entre la probabilidad de presentar una característica y la probabilidad de no presentarla.

$$Odds = \frac{p}{1-p}$$

También, en caso de ser necesario, se pueden comparar las variables con el concepto de *odds-ratio* (OR), que representan el cociente entre ambos *odds*. Además, de cada coeficiente asociado a una variable se puede obtener la probabilidad de que la variable asociada esté presente a través de la siguiente expresión para variables dicotómicas.

$$e^{\theta_i} = OR$$

En caso de variables numéricas, se puede medir el concepto de pasar de un valor  $X_1$  a un valor  $X_2$  a partir de la siguiente expresión:

$$e^{\theta_i*(X_2-X_1)} = OR$$

Adicionalmente, es posible comprobar si el modelo es adecuado analizando su precisión. Para ello, se utiliza una muestra de datos similar a la que se utilizó en el entrenamiento para comprobar si las salidas obtenidas por el modelo son similares a las que se esperan de este a través de los valores reales.

Esos valores se pueden medir de varias formas, por un lado, existe la precisión de cada una de las variables, es decir, el porcentaje de veces que ha clasificado bien el modelo cuando se esperaba una salida determinada. Por otro lado, se encuentra la matriz de confusión, la cual representa los aciertos y fallos para cada uno de los valores de salida que se pueden obtener. En un modelo dicotómico, esta matriz incluirá los verdaderos y falsos positivos (VP y FN), y los verdaderos y falsos negativos (VN y FN), respectivamente.

$$\begin{pmatrix} VP & FP \\ FN & VN \end{pmatrix}$$

---

## 2.2. CLASIFICADOR BAYESIANO INGENUO

---

Los clasificadores basados en el teorema de Bayes [23], a diferencia de otros tipos, miden la probabilidad de una hipótesis. Este tipo de clasificadores destaca por su robustez frente al posible ruido introducido con los datos de entrenamiento o ante la posibilidad de que estos conjuntos de valores estén incompletos.

Su principal debilidad emana de la posible insuficiencia de datos de entrenamiento necesarios para ajustar el clasificador de modo que consiga realizar las estimaciones correctamente.

El funcionamiento del algoritmo bayesiano ingenuo [23] está basado en buscar una función  $f(x)$  que, para un conjunto de entradas independientes, devuelva la probabilidad de una hipótesis. De este modo, una vez se ha ajustado el algoritmo, es posible calcular las probabilidades de todas las hipótesis para un caso concreto. Es decir, para un conjunto de valores  $x = (x_1, x_2, x_3, \dots, x_n)$  y una hipótesis  $c_i$ , la probabilidad de esta se define como:

$$P(c_i|x) = \frac{P(x|c_i)P(c_i)}{P(x)}, \quad \text{con } i = 1, \dots, l$$

Dado que las variables de entrada que marcan cada uno de los casos son independientes, es posible simplificar la expresión como:

$$P(x|c_i) = \prod_{j=1}^n P(x_j|c_i), \quad \text{con } i = 1, \dots, l$$

De modo que la hipótesis más probable para ese caso concreto será la que cumpla la siguiente expresión:

$$\text{argmax}_{c_i \in C} P(c_i|x)$$

Dentro de este algoritmo se diferencian tres tipos de modelos principalmente:

- **Gaussiano:** Se utiliza cuando la hipótesis de salida sigue una distribución normal.
- **Multinomial:** Su uso está dirigido a casos en los que se realizan cálculos discretos. Por ejemplo, cuando se quiere saber la probabilidad de que aparezca una determinada letra en un texto, lo que hace necesario contar la cantidad de veces que se encuentran cada una de las letras.
- **Bernoulli:** Este modelo se aplica cuando la variable de salida se compone de 0s y 1s.

---

## 3. PRODUCTOS COMERCIALES

---

La localización en interiores ha alcanzado un gran interés por parte de las empresas punteras en tecnología. Muchas de ellas compiten por proporcionar una alternativa que permita a los usuarios obtener su posición en zonas donde el GPS no funciona correctamente. Esto ha permitido la entrada al mercado de distintas alternativas que pretenden dar una solución a este problema.

### 3.1. SITUM INDOOR POSITIONING



Situm Technologies [24] es una empresa dedicada a proveer servicios de localización en interiores utilizando una infraestructura mínima. Esta empresa fue fundada por tres ingenieros que a pesar de su pequeño tiempo de vida ya cuenta con su tecnología desplegada en más de 1600 edificios en todo el mundo.

La startup ha sido galardonada con numerosos premios tanto nacionales como internacionales entre los cuales se encuentran, el Smart Ideas Award 2014 en la categoría de Smart City, el National YUZZ Award 2014 del Banco Santander y el Elevator Pitch Singularity University Awards 2015.

Entre las características de su tecnología destacan la localización en interiores con un error máximo de un metro y la capacidad para distinguir mobiliario, ascensores, escaleras y distintos niveles de altura. Para conseguir estas diferencias, el sistema utiliza las tecnologías GPS, red móvil, Wi-Fi y Bluetooth de los dispositivos móviles, además de la información de brújula a través del magnetómetro.

### 3.2. INDOOR SURVEY

Apple lanzó en el año 2015 una aplicación integrada en su servicio de mapas conocida como Indoor Survey [25]. Esta aplicación funciona conjuntamente con Maps Connect [26] y su principal propósito es permitir a los propietarios de locales crear un sistema de localización en el interior de su propiedad.



Para poder hacer uso de esta herramienta el usuario deberá enviar a Apple un plano detallado del edificio que quiere incluir dentro de la aplicación. En caso de que el equipo de validación, tras comprobar la calidad de los planos y el flujo de visitantes del lugar, lo declare apto, el usuario podría comenzar a tomar medidas dentro de su edificio para crear una base de datos de forma similar a como se realizaba en el método de Fingerprint. Realizadas estas medidas, el sistema podría comenzar a posicionar a los usuarios que accedan al servicio de mapas de Apple dentro del edificio.

### 3.3. GOOGLE LOCATION SERVICES

Los servicios de mapas de Google también disponen de un servicio de localización en interiores [27]. En este caso el sistema es muy similar al de Fingerprint pero los datos son recogidos por los terminales, en vez de por los puntos de acceso, y no solo de la tecnología Wi-Fi. El posicionamiento en interiores de Google, al igual que el Situm Indoor Positioning hace uso de todas las tecnologías de posicionamiento disponibles en el terminal.

A partir de todas estas opciones, el sistema utiliza la información recopilada por el teléfono para buscar equivalencias con la base de datos de Google. En ella se encuentran las posiciones de las antenas de telefonía y de numerosos puntos de acceso de modo que

con todos estos datos consigue estimar ubicaciones con una precisión superior a la del sistema GPS en muchos casos.

### 3.4. OTROS PRODUCTOS

Como se menciona en [28], existen otras soluciones comerciales como la propuesta por la empresa Ekahau con su herramienta Ekahau Vision, la cual basa su sistema de posicionamiento en el método de fingerprint, ayudándose de etiquetas RFID sobre WiFi.

Por otro lado, también se explica la herramienta Accuware Indoor Navigation, que al igual que la anterior, utiliza un sistema de fingerprint, pero en este caso, basado en Wi-Fi e iBeacons como datos de entrada al sistema.

## 4. MARCO DEL TRABAJO

Como ya se ha avanzado, este trabajo surge de la colaboración entre AENA, la Asociación Española de Ingenieros de Telecomunicaciones en Cantabria, y el Grupo de Ingeniería Telemática de la Universidad de Cantabria. Debido a las estrictas condiciones de seguridad por las que se rigen las terminales aeroportuarias, se tratará de realizar una mejora en las eventuales situaciones de emergencia, analizando la ubicación donde se concentran las personas.

Para ello, en colaboración con el personal técnico del Aeropuerto Seve Ballesteros-Santander, se ha procedido al diseño, despliegue y posterior validación del uso de dispositivos IoT para el fin de gestionar de forma más eficiente emergencias y situaciones singulares que puedan darse dentro de la terminal.



*Figura 10 – Aeropuerto Seve Ballesteros-Santander*

Seguidamente, y a modo introductorio se incluye una breve explicación relativa a los aeropuertos y sus características más relevantes con especial énfasis en aquellas que impactan más directamente en este trabajo. Se entiende como tal al área definida de tierra o agua, destinada, total o parcialmente, a la llegada, salida y movimiento en superficie de

aeronaves [29]. Por otro lado, esta área debe disponer de las siguientes instalaciones y servicios de carácter público permanentemente:

- Asistir, de modo regular, al transporte aéreo.
- Permitir el aparcamiento de aeronaves.
- Posibilitar la recepción y despacho de pasajeros y carga.

De lo que surgen las siguientes definiciones:

- **Aeronave:** Cualquier aparato capaz de volar, como aviones y helicópteros.
- **Pasajero:** Usuario del transporte aéreo.
- **Equipaje:** Propiedades que el pasajero lleva en el viaje, bien en la cabina del avión (equipaje de mano) o en la bodega del mismo (equipaje facturado).
- **Carga:** Mercancías que se transportan por vía aérea.

Por otra parte, como se indica en [29], un aeropuerto debe cumplir unas características que se explicarán a continuación.

En cuanto a su configuración, en la Figura 11 se explican las partes en las que se divide principalmente un aeropuerto.

<b>Aeropuerto</b>	Lado tierra	Área Terminal		Terminal de pasajeros
				Terminal de carga
				Otros edificios
		Urbanización		Vías de acceso
				Aparcamiento
				Zona industrial
	Lado aire	Área de movimiento	Área de maniobras	Área de aterrizaje
				Área de rodaje
				Plataforma

Figura 11 – Partes de un aeropuerto.

Como se puede observar, existe el lado tierra y el lado aire. El primero de ellos se caracteriza por contener toda la infraestructura dedicadas a satisfacer las necesidades de los pasajeros. Por otro lado, el lado aire se centra en las aeronaves y en todo lo relacionado con ellas.

Dentro del lado tierra, el edificio principal es La Terminal, el cual sirve de enlace entre el transporte aéreo y terrestre y cuyas principales funciones [30] son las siguientes:

- **Tramitación:** Necesaria para iniciar o acabar un vuelo. Las actividades principales son: información, despacho de billetes, facturación, recogida de equipajes, controles (seguridad, aduanas, sanidad, migración) y embarques.
- **Circulación:** A través de itinerarios preestablecidos. Se incluyen las conexiones con los accesos y el avión.
- **Espera:** Incluyendo una serie de servicios y zonas comerciales como tiendas, restaurantes y cafeterías, farmacia, guardería, bancos y cambios de moneda, consigna, correos, alquiler de coches, teléfonos, capilla, salas VIP, aseos, salas de espera, etc.



Figura 12 – Fotografía del lado aire del aeropuerto Seve Ballesteros-Santander

En cuanto al lado aire, se centra en todos los procesos relacionados con el despegue y aterrizajes de las aeronaves. Entre ellos se encuentran algunos como el guiado por la pista, estacionamiento, diferentes tareas de mantenimiento, acceso de pasajeros, gestión de equipajes y carga y descarga de distintas mercancías.

En lo que se refiere al aeropuerto Seve Ballesteros-Santander, este se encuentra en el municipio de Camargo (tercer ayuntamiento, en número de habitantes, de Cantabria), a 4 km al sur de la ciudad. Tiene asignado el código IATA (*International Air Transport Association*), SDR y puede considerarse urbano por su proximidad a la ciudad.

La terminal cuenta con unas dimensiones de 17.112 m<sup>2</sup> de los cuales se puede decir que 13.270 m<sup>2</sup> son de superficie útil. En cuanto a su volumen trabajo, en la Figura 13 y Figura 14 se recogen los datos oficiales ofrecidos por la web de AENA entre los años 2011 y 2015.

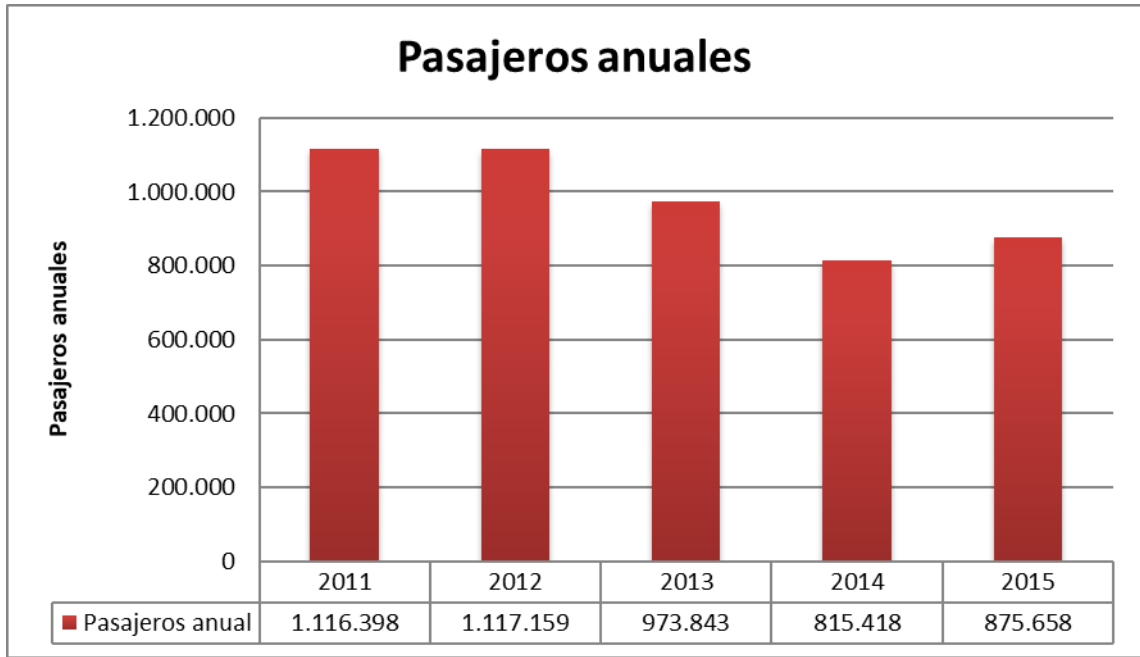


Figura 13 – Número de pasajeros anuales del aeropuerto Seve Ballesteros-Santander

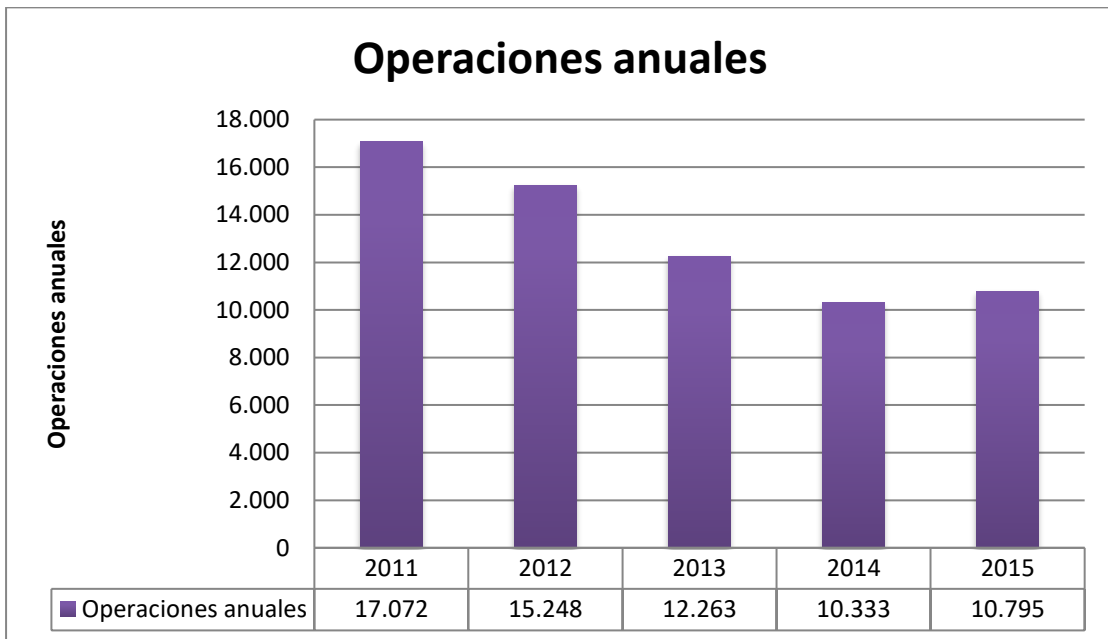


Figura 14 – Número de operaciones anuales del aeropuerto Seve Ballesteros-Santander

---

### III. DESARROLLO DE UN PILOTO PARA LA DETECCIÓN DE CONCENTRACIÓN DE PERSONAS EN LAS TERMINALES AEROPORTUARIAS

---

Para el desarrollo del proyecto se ha desplegado una red de sensores en la terminal que permiten capturar las tramas Wi-Fi y Bluetooth de los dispositivos de los usuarios. El punto de partida sobre el que se ha basado el diseño es el Trabajo de Fin de Grado de Juan Carrasco Alonso [28] y del posterior artículo [31] publicado con los resultados de este. Durante el proyecto se desarrolló un sistema que permitiera capturar las tramas Wi-Fi del tipo *Probe Request* y empleando el método de la potencia recibida explicado en el punto 1.3 de este capítulo, con ayuda del algoritmo del centroide que se explica en la memoria del trabajo, a fin de estimar la posición de los dispositivos.

Las diferencias entre ambos proyectos, que emanan principalmente del enclave, hicieron necesario idear una solución para los siguientes problemas:

- **Envío de datos a través de 3G:** Debido a las condiciones acordadas con el Jefe de Ingeniería y Mantenimiento del aeropuerto, no es posible realizar el envío de los datos capturados a través de la red Wi-Fi del local como ocurría en el caso del “Mercado del Este”, donde se desplegó [28].
- **Problemas de estabilidad:** Se pudo descubrir que el sistema previamente desarrollado no operaba de forma totalmente estable cuando los equipos llevaban un determinado tiempo encendidos sin reiniciarse, debido a la acumulación de datos en la memoria.
- **Necesidad de Bluetooth:** Aunque estaba previsto el uso de Bluetooth como fuente adicional para el posicionamiento, no se había desarrollado completamente el código que permitiera capturar los *beacons* BLE necesarios para posicionar a través de esta tecnología.

Durante este capítulo se explicará el diseño de los dispositivos, tanto HW como SW, la realización del despliegue, y la elección del algoritmo de posicionamiento que permite estimar la concentración de personas. Finalmente, se incluirá un presupuesto breve que recoge el coste de los equipos utilizados.

---

#### 1. DESCRIPCIÓN DEL HARDWARE DE DETECCIÓN

---

Cada uno de los sensores desplegados en la terminal está formado por una Raspberry Pi 3 Model B+ [32] cuyas prestaciones, con el añadido de tener un coste reducido, la hacen apropiada para el desarrollo de este sistema. Algunas de las características más destacadas para este trabajo son:

- Wi-Fi 802.11 ac a 2.4 Ghz y 5 Ghz
- Bluetooth 4.2 BLE
- Power over Ethernet (PoE)
- Gigabit Ethernet hasta 300 Mbps



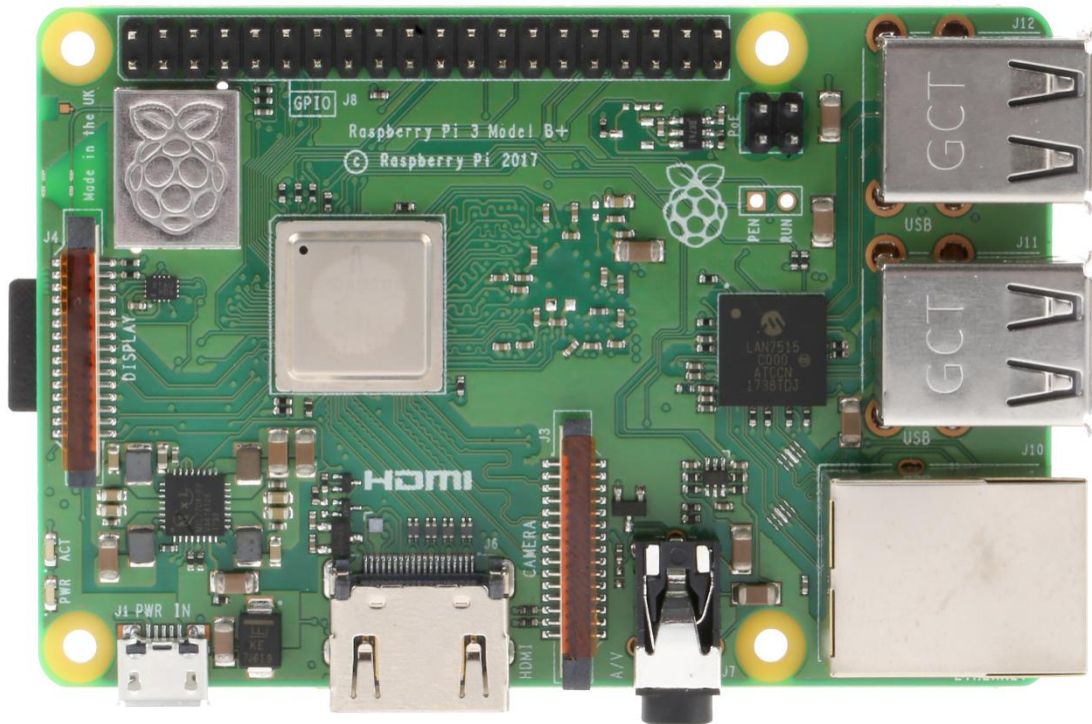


Figura 15 – Fotografía superior de Raspberry Pi 3 Model B+

Debido a la necesidad de disponer de una interfaz Wi-Fi en modo promiscuo [33], a cada uno de los sensores se le ha añadido un adaptador de red Wi-Fi a través de la interfaz USB de la RP de la marca TP-LINK modelo TL-WN722N [34] que amplía sus características iniciales. El disponer de una segunda interfaz para capturar paquetes permite utilizar la interfaz Wi-Fi integrada en la RP para enviar *beacons* periódicamente que capturaran el resto de los sensores, haciendo posible realizar estimaciones del canal.

Con el fin de facilitar la instalación, se ha decidido alimentar todos los dispositivos a través de PoE, lo cual no solo permite eliminar los cables de alimentación en cada una de las localizaciones, sino que, ante la necesidad de un reinicio completo del sistema, simplifica notablemente esta operación al poder realizarse cortando la alimentación del encaminador. Para ello se han utilizado diferentes Splitter TP-LINK modelo TL-POE10R [35] entre el encaminador (*router*) y las RP, permitiendo alimentar cada uno de los sensores a través de PoE.



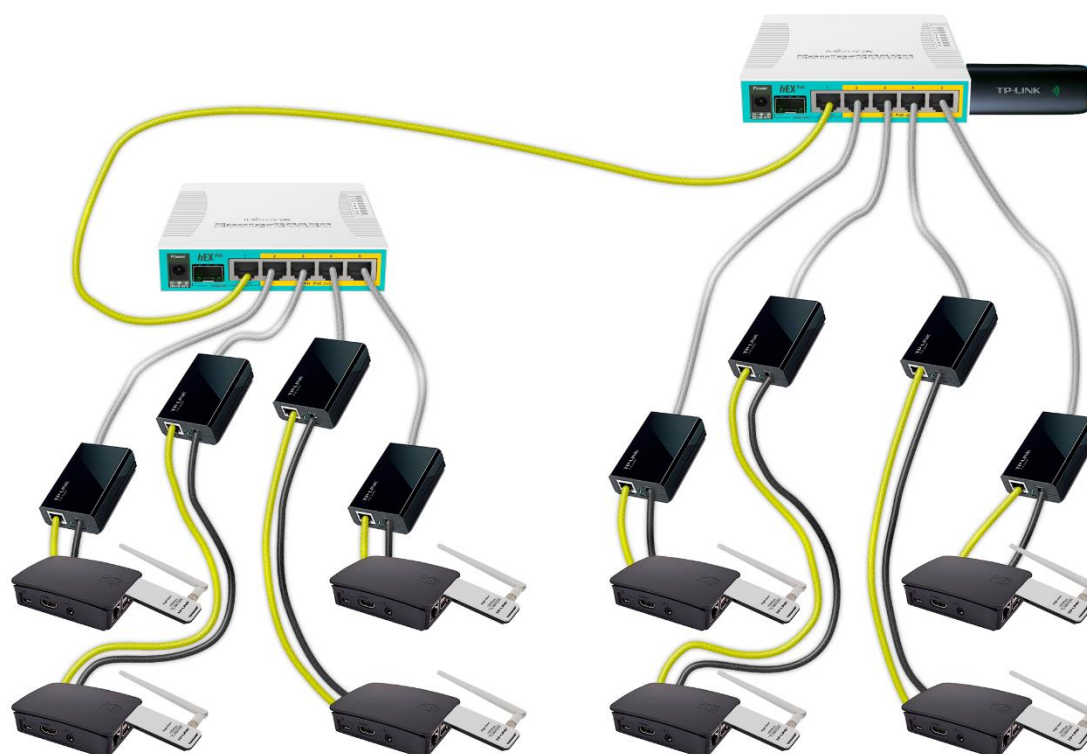
Figura 16 – Despliegue de RP, con segunda interfaz y alimentación a través de PoE

Para crear la red de sensores se han utilizado dos *router* Mikrotik hEX PoE modelo RB960PGS [36] instalados en cascada, los cuales permiten la utilización de PoE para alimentar y conectar en red hasta 4 RP cada uno.



*Figura 17 – Instalación de los dos encaminadores Mikrotik*

Finalmente, como pasarela a internet se ha utilizado un modem 3G TP-LINK modelo MA260 que habilitará el envío de las tramas capturadas por cada uno de los sensores al servidor principal, donde se realizará el procesamiento de estos.



*Figura 18 – Esquema del despliegue de los dispositivos*

Para consolidar el esquema final representado Figura 18 se ha realizado un largo periodo de desarrollo en el que se han ido solventando diferentes problemas que se explicarán a lo largo del capítulo.

El primero se debió al intento de actualización del código a Python 3 con la última versión de las librerías utilizadas. Se pudo comprobar que, en el momento del desarrollo,

la librería pyshark [37] tenía problemas al capturar paquetes haciendo uso de la utilidad *LiveRingCapture* lo que provocó regresar a la versión 0.3.6.2, basada en Python 2. Debido a ello, el resto de las librerías se actualizaron a la última versión disponible que soportara Python 2, no pudiendo completarse totalmente la actualización.

En cuanto a la captura de paquetes Bluetooth, se han utilizado las herramientas Hcitol [38] y BlueZ [13]. La primera, como se puede ver en la figura Figura 19 permite capturar paquetes Bluetooth Low Energy (BLE) de forma sencilla, pero sin aportar información sobre ellos.

```
unican@raspberrypi:~ $ sudo hcitool lescan --duplicates
[sudo] password for unican:
LE Scan ...
7A:58:6D:DC:83:60 (unknown)
F7:D4:FC:1A:26:42 (unknown)
F7:D4:FC:1A:26:42 MI Band 2
7A:58:6D:DC:83:60 (unknown)
F7:16:DE:63:77:5A (unknown)
F7:16:DE:63:77:5A Amazfit Bip Watch
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
F7:D4:FC:1A:26:42 (unknown)
F7:D4:FC:1A:26:42 MI Band 2
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
7A:58:6D:DC:83:60 (unknown)
F7:16:DE:63:77:5A (unknown)
F7:16:DE:63:77:5A Amazfit Bip Watch
```

Figura 19 – Captura del funcionamiento de la herramienta hcitool

Por otro lado, dentro de Bluez se encuentra la herramienta btmon [39], la cual es un fantástico complemento de hcitool [40] ya que a partir de los paquetes BLE capturados permite extraer diferentes campos de información.

```

unican@raspberrypi:~$ sudo btmon
[sudo] password for unican:
Bluetooth monitor ver 5.43
= Note: Linux version 4.9.80-Re4son-v7+ (armv7l) 0.533006
= Note: Bluetooth subsystem version 2.22 0.533016
= New Index: 88:27:EB:6D:AB:5D (Primary,UART,hci0) [hci0] 0.533019
= Open Index: 88:27:EB:6D:AB:5D [hci0] 0.533022
= Index Info: 88:27:EB:6D:AB:5D (Broadcom Corporation) [hci0] 0.533025
@ MGMT Open: bluetoothd (privileged) version 1.14 {0x0001} 0.533030
@ MGMT Open: btmon (privileged) version 1.14 {0x0002} 0.533090
@ RAW Open: hcitool (privileged) version 2.22 {0x0003} 63.669081
@ RAW Close: hcitool {0x0003} 63.669132
@ RAW Open: hcitool (privileged) version 2.22 {0x0003} [hci0] 63.669187
< HCI Command: LE Set Scan Parameters (0x08|0x000b) plen 7 [hci0] 63.669250
    Type: Active (0x01)
    Interval: 10.000 msec (0x0010)
    Window: 10.000 msec (0x0010)
    Own address type: Public (0x00)
    Filter policy: Accept all advertisement (0x00)
> HCI Event: Command Complete (0x0e) plen 4 [hci0] 63.669735
    LE Set Scan Parameters (0x08|0x000b) ncmd 1
    Status: Success (0x00)
< HCI Command: LE Set Scan Enable (0x08|0x000c) plen 2 [hci0] 63.669844
    Scanning: Enabled (0x01)
    Filter duplicates: Disabled (0x00)
> HCI Event: Command Complete (0x0e) plen 4 [hci0] 63.670317
    LE Set Scan Enable (0x08|0x000c) ncmd 1
    Status: Success (0x00)
> HCI Event: LE Meta Event (0x3e) plen 43 [hci0] 63.694860
    LE Advertising Report (0x02)
    Num reports: 1
    Event type: Non connectable undirected - ADV_NONCONN_IND (0x03)
    Address type: Random (0x01)
    Address: 7A:58:6D:DC:83:60 (Resolvable)
    Data length: 31
    Company: Microsoft (6)
    Data: 010920024dc11c73e78f31da27312ff468c60fe78a094ce63b1ef6
    RSSI: -70 dBm (0xba)
> HCI Event: LE Meta Event (0x3e) plen 43 [hci0] 63.778879
    LE Advertising Report (0x02)
    Num reports: 1
    Event type: Connectable undirected - ADV_IND (0x00)
    Address type: Random (0x01)
    Address: F7:D4:FC:1A:26:42 (Static)
    Data length: 31
    Flags: 0x06
    LE General Discoverable Mode
    BR/EDR Not Supported
    Company: Anhui Huami Information Technology Co., Ltd. (343)
    Data: 0022a33c3d689445b9a584796ed8cbd41102f7d4fc1a2642
    RSSI: -80 dBm (0xb0)
> HCI Event: LE Meta Event (0x3e) plen 27 [hci0] 63.779531
    LE Advertising Report (0x02)
    Num reports: 1
    Event type: Scan response - SCAN_RSP (0x04)
    Address type: Random (0x01)
    Address: F7:D4:FC:1A:26:42 (Static)
    Data length: 15
    Name (complete): MI Band 2
    16-bit service UUIDs (partial): 1 entry
    Anhui Huami Information Technology Co. (0xfee0)
    RSSI: -81 dBm (0xaf)

```

Figura 20 – Captura del funcionamiento de la herramienta btmon

Como muestra la Figura 20, con ayuda de ambas herramientas, es posible capturar y obtener datos como la dirección de capa de control de acceso al medio (*Media Access Control, MAC*) del dispositivo transmisor, el nombre del dispositivo (en algunos casos) o el indicador de nivel de señal recibida (*Received Signal Strength Indicator, RSSI*) con el que se ha detectado el paquete. Con toda esta información es posible procesar los datos de forma similar a los paquetes capturados mediante Wi-Fi.

Una vez creada una primera versión de los sensores, resultaba obligado probarlos en una red controlada antes de proceder a un despliegue de campo. Para ello se realizó aquel en el laboratorio con un esquema similar al de la Figura 21, el cual estaba compuesto por 4 RP, un *router* Mikrotik y la pasarela a internet a través de 3G. De forma simultánea se desplegaron otras dos RP con acceso a internet a través de un *router* conectado a la red de la Universidad, con la finalidad de poder observar los posibles problemas que podrían surgir al introducir el modem 3G.





Figura 21 – Esquema del despliegue en el laboratorio

De estas pruebas de laboratorio se pudieron detectar dos problemas a solucionar. En primer lugar, el consumo de datos era excesivo ya que el código estaba diseñado para un despliegue en el que el acceso a Internet se realizaba a través de la propia red del edificio, la cual no precisa de una tarifa de datos limitada. Además, la estabilidad de las RP cuando no se reiniciaban, transcurrido un tiempo, no era la que se precisaba.

Como solución al primer problema, se realizaron varias acciones que permitieron reducir el consumo notablemente:

- **Compresión de la dirección MAC:** La primera medida que se llevó a cabo fue reducir el número de bytes enviados como información de la MAC. El sistema está preparado para realizar un hash de 128 bytes de cada una de las direcciones MAC que no pertenecen a una lista de dispositivos conocidos. Debido a la gran longitud de este campo, se decidió no enviar este hash completo, sino enviar tan solo los 8 primeros bytes del hash. De este modo, a pesar de la importante reducción de tamaño, seguiría habiendo  $16^8 = 4.294.967.296$  posibles opciones, lo que hace altamente improbable que coincidan dos dispositivos con los mismo 8 bytes iniciales en su hash de la dirección MAC, de forma simultánea en la terminal del aeropuerto.
- **Timestamp:** En la versión inicial, se enviaba el *timestamp* de cada una de las RP al enviar los paquetes, pero al observar que este coincidía con el *timestamp* de captura de los paquetes, se vio coherente prescindir de él, para reducir el tamaño. Además, se modificó el envío del *timestamp* de cada paquete para hacerlo en binario y reducir el tamaño a 4 bytes.
- **Tiempo de envíos:** Como se explica en [28], el código está preparado con un umbral donde para un mismo dispositivo no es necesario enviar dos paquetes capturados en un determinado periodo de tiempo. Este tiempo se modificó a 10 segundos, lo que redujo notablemente el envío de tramas sin valor con respecto a los envíos cada 2 segundos programados anteriormente.
- **Aumento de la tarifa:** Además, ante la necesidad de un sistema que presente una gran robustez, se decidió aumentar la tarifa de datos a 4 GB de datos para disponer de un margen superior.

Estas tres medidas permitieron encontrar una solución ante la dificultad de enviar las tramas al servidor a través de 3G, pero aun había un problema de estabilidad derivado de la memoria de las RP. Este problema se debía, por un lado, a que el almacenamiento en las

tarjetas micro-sd estaba configurado en modo *read-only* con el fin de proteger las mismas. Por otro lado, los sensores están configurados para que, al no recibir tramas durante un periodo de dos minutos, se reinicie la parte dedicada a Wi-Fi. De este modo, cuando cada RP sufría un número determinado de reinicios, el espacio habilitado en RAM para el guardado de los diferentes archivos temporales, principalmente los archivos con extensión *.pcap* creados para la captura de tramas, terminaba completándose ante la creación de unos nuevos. Para solucionar este problema se aplicaron las siguientes medidas:

- **Reducción del número de logs:** Cada RP estaba configurada para que almacenara hasta 5 ficheros de depuración con las tramas capturadas, cada uno de ellos de 1024000 bytes máximo. De este modo, cuando se completara el espacio reservado para los 5 ficheros de trazas (*logs*), las nuevas tramas borrarían los datos guardados de las más antiguas de forma cíclica. Al añadir los *logs* de Bluetooth respecto a la versión inicial, la memoria reservada era insuficiente, por lo que se decidió reducir de 5 a 3 logs.
- **Incremento de la memoria reservada:** Dado que las RP tenían capacidad suficiente y se había añadido la función de Bluetooth, se aumentó el espacio reservado para los archivos logs hasta los 50 MB.
- **Borrado de la carpeta tmp:** Como se comentó anteriormente, aparte de los *logs*, cada RP iba creando diferentes archivos temporales derivados de las capturas de la librería pyshark en cada reinicio sucesivo, de modo que la memoria terminaba llenándose en algún momento. Para solucionarlo, se configuró para que, en cada reinicio, no solo volviera a arrancar todos los procesos e interfaces, si no que previamente eliminara estos archivos temporales que no tienen ningún interés en el proyecto.

Implementadas estas soluciones, se prosiguió con las pruebas del sistema y se detectó un nuevo inconveniente; la temperatura de las RP era demasiado alta, por lo que se analizó el comportamiento del sistema. Se pudo comprobar que, en los reinicios comentados anteriormente, no solo se acumulaban archivos temporales, si no que a veces quedaban abiertos procesos de la herramienta *btmon*, explicado su uso en las medidas de BLE anteriormente. Por tanto, se programó el sistema para que en cada reinicio eliminara los procesos abiertos que ya no eran precisos.

A pesar de aplicar esta solución, se detectó que la temperatura seguía siendo alta y se realizó un análisis de su comportamiento con diferentes configuraciones de gestión de energía por parte del microprocesador. Para ello, se configuraron las 4 RP de la red de pruebas de la siguiente manera:

- **RP1:** Se mantuvo la configuración por defecto, en la que el propio sistema operativo elige la frecuencia más adecuada para el procesador, en este caso 600 MHz o 1400 MHz.
- **RP2 y RP4:** Ambas se programaron para que la frecuencia fuera disminuyendo desde los 1400 MHz, en pasos de 100 MHz cada 2 horas, con un límite inferior de 300 MHz.
- **RP3:** En este caso, la configuración se hizo en modo *conservative*, en el cual la frecuencia varía entre 600 MHz y 1400 MHz como en el primer caso, pero esta vez se utiliza el valor inferior siempre que es posible.

Para analizar las diferencias de una manera más precisa, se utilizaron los nodos desplegados en el laboratorio para conocer la temperatura en varios puntos de la sala, de modo que el despliegue final era similar al de la Figura 22.

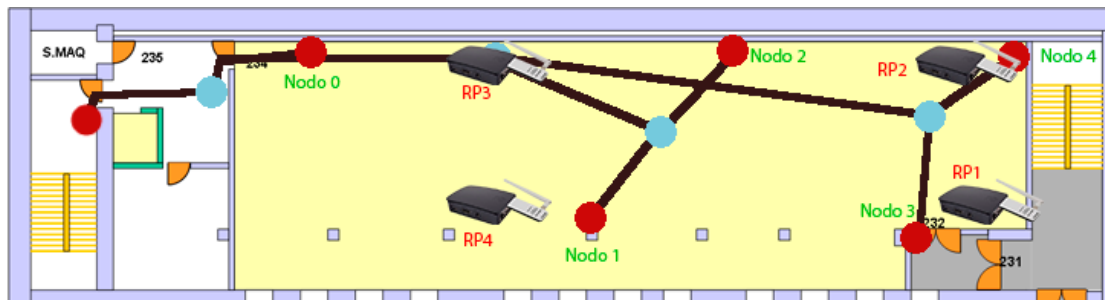


Figura 22 – Despliegue de RP y sensores de temperatura en el laboratorio

Con el fin de que los datos no se vieran afectados por factores externos, se realizaron medidas de temperatura, ocupación y frecuencia cada minuto durante un fin de semana completo. De estas medidas se pudo extraer la siguiente comparativa de temperaturas con datos promediados cada 5 minutos.

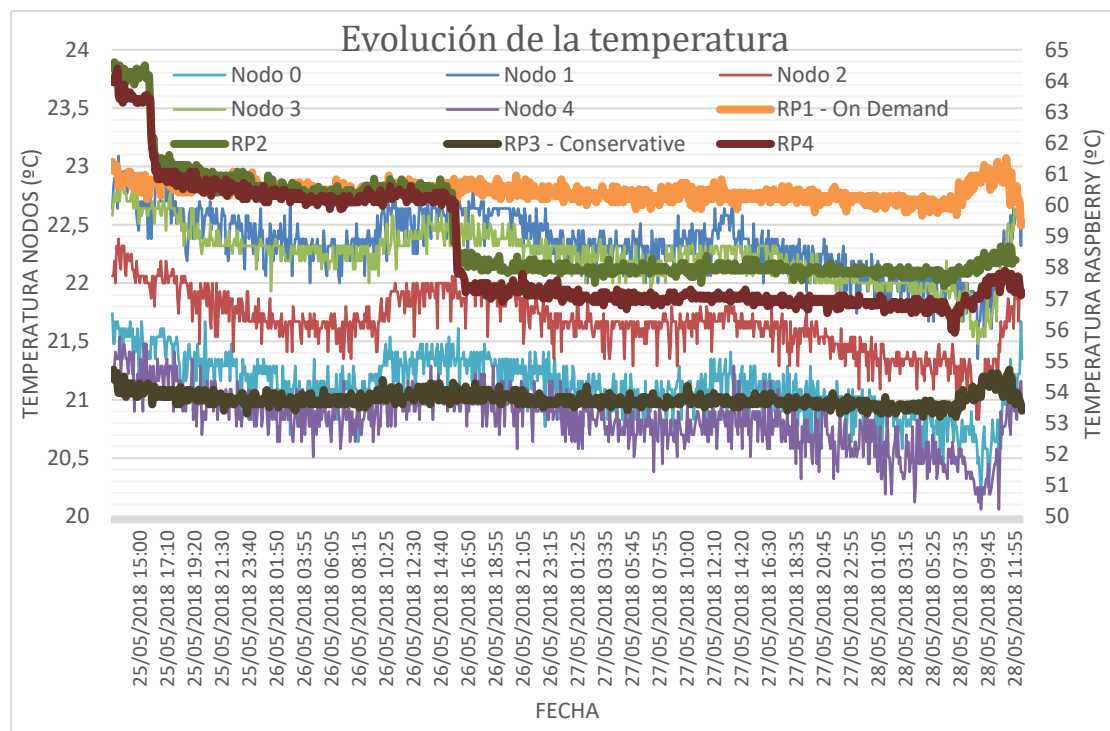


Figura 23 – Comparativa de la evolución de la temperatura en cada RP

Analizando los valores representados en la Figura 23 se pueden apreciar dos instantes de tiempo en los que la temperatura de las RP programadas para ir disminuyendo su frecuencia gradualmente sufren un bajón de temperatura importante. Estos coinciden tanto en la RP2 como en la RP4 con el instante en el que ambas disminuyen su frecuencia a 1200 MHz (primer salto) y 600 MHz (segundo salto). Una vez alcanzado este primer momento, la temperatura se mantiene con unos valores similares a los de la RP1, cuya configuración estaba en modo *on-demand*, configuración que provoca que la frecuencia del procesador ascienda a los 1400 MHz el 58% del tiempo. Estos registros se mantienen hasta el segundo salto, donde la temperatura claramente desciende para mantenerse unos grados por debajo hasta el final del experimento. Esto puede llevar a pensar que mantener la

frecuencia de manera constante a 600 MHz puede ser una buena solución para evitar elevadas temperaturas, pero si se comparan los valores con la RP3, que se encontraba en modo *conservative* se puede ver como a pesar de que este modo eleva la frecuencia a 1400 MHz en determinadas ocasiones (14% del tiempo), la temperatura se mantiene notablemente por debajo en todo momento.

A simple vista, esta valoración puede parecer contradictoria, pero al observar los tiempos de ocupación del procesador mostrados en la Figura 24, se aprecia claramente como la RP2 mantiene el procesador mucho más tiempo inactivo que las RP con frecuencia constante al descender hasta los 600 MHz.

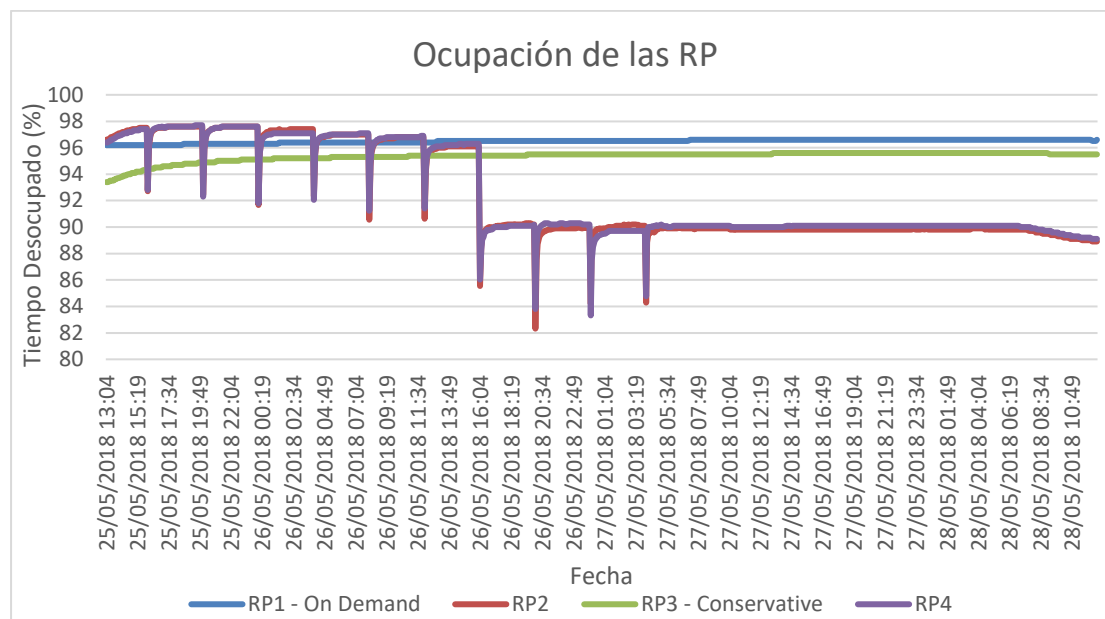


Figura 24 – Comparativa de ocupación del procesador en cada RP

Al analizar todos estos datos en conjunto, se puede concluir con seguridad que para el desarrollo del proyecto, la configuración en modo *conservative* de la RP2 es la más apropiada, ya que mantiene la temperatura a un valor notablemente más bajo que el resto en todo momento, y además consigue unos tiempos de ocupación del procesador del orden de la configuración de la *on-demand* de la RP1, la cual solo se ve superada en cuanto a tiempo con el procesador libre cuando las otras RP mantienen su procesador a frecuencias muy altas continuamente.

Con todas estas modificaciones y ajustes, se consiguió un sistema que presentaba una importante mejora de estabilidad en las pruebas de laboratorio.

Antes de realizar el despliegue final en la terminal aeroportuaria, se decidió someter al sistema a una simulación de medidas controladas en el laboratorio de la universidad. La primera prueba, con el fin de analizar también el funcionamiento de los teléfonos con los que se realizarían las medidas de entrenamiento en la terminal, consistía en un análisis de la frecuencia de emisión de *beacons* por varios teléfonos en diferentes situaciones. Se hicieron medidas con las pantallas apagadas y encendidas y con los teléfonos conectados a una red wifi y desconectados.

En la Figura 26 se puede muestran los datos recopilados para diferentes terminales. Las gráficas de la izquierda incluyen valores atípicos en la representación, mientras que los



de la derecha prescindir de estos, para ver en detalle los valores más habituales. Antes de analizar los datos es importante comprender los datos representados en estas gráficas. Para ello se utilizará el ejemplo [41] de la Figura 25. En ella se diferencian los siguientes puntos:

- **Valor medio:** marcado mediante una X.
- **Valores máximo y mínimo:** marcados en rojos y verde respectivamente los cuales coinciden con los bigotes del ejemplo.
- **Cuartil 1 o 25 percentil:** destacado en naranja y corresponde con el borde inferior de la caja.
- **Mediana o 50 percentil:** señalado en la figura mediante un color amarillo
- **Cuartil 3 o 75 percentil:** en color magenta.
- **Valores atípicos:** marcados por puntos.

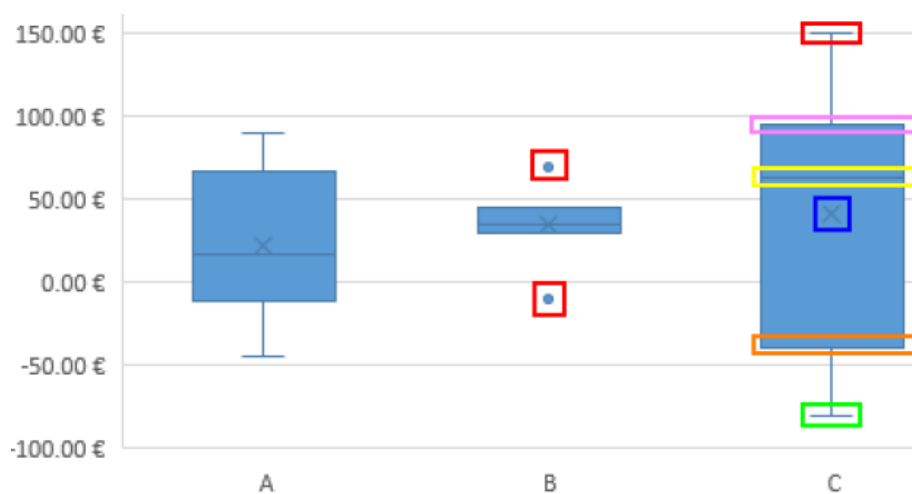


Figura 25 – Ejemplo grafica cajas y bigotes.

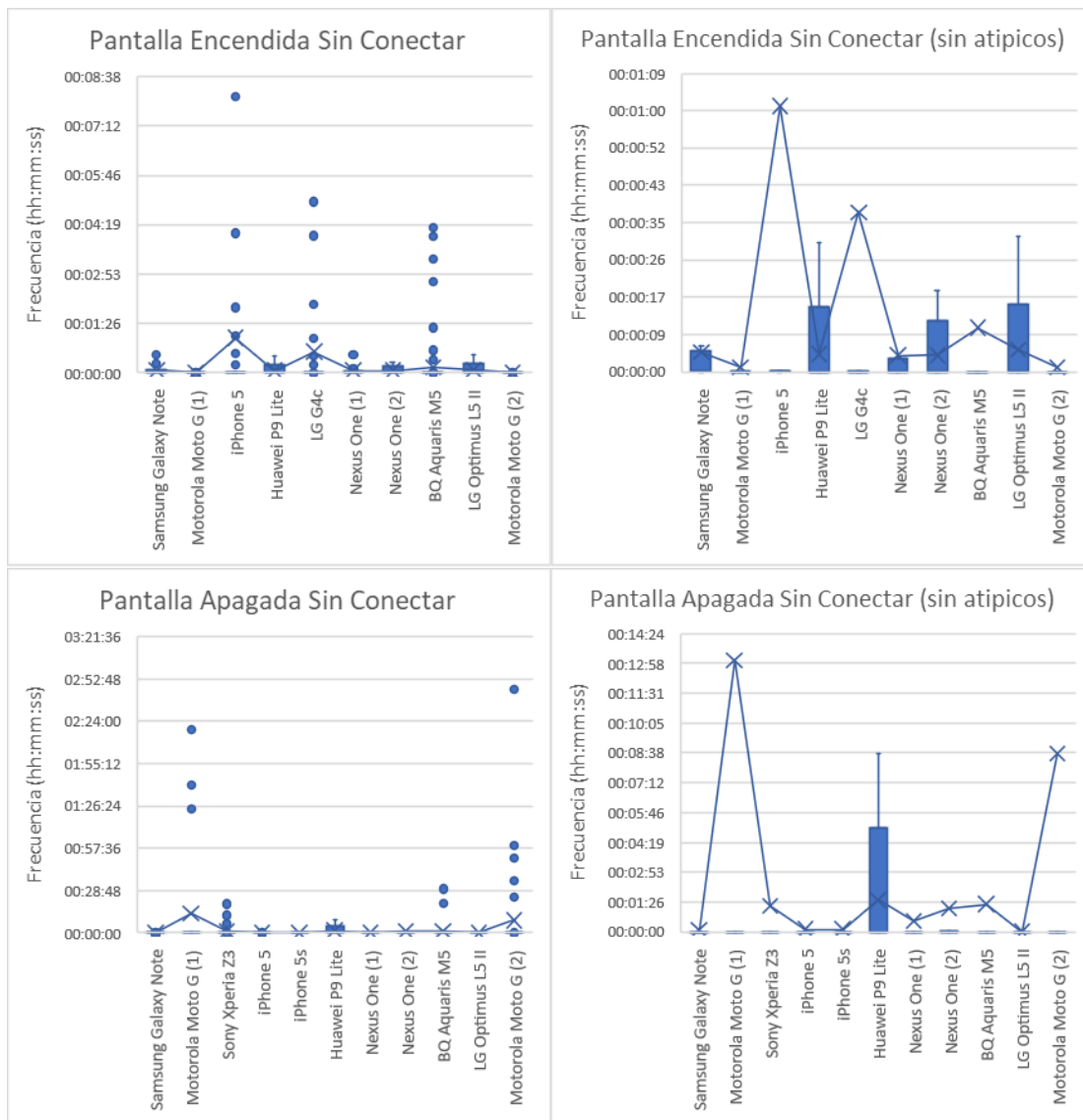
En función de lo compactas o dispersas que sean las distribuciones, será adecuado señalar unos puntos u otros. Pudiéndose dar el caso de que lo más adecuado sea solamente señalar la media

Analizando las figuras es posible observar que cada teléfono tiene unos valores bastante distintos del resto, pero aun así es posible extraer conclusiones globales a todos ellos. La primera valoración que se puede hacer es la fuerte relevancia ligada a tener la pantalla encendida o apagada y estar conectado a una red Wi-Fi o no. Analizando los casos en los que los dispositivos no están conectados a ninguna red, se puede apreciar un importante crecimiento tanto en los valores típicos como en los atípicos, pasando de valores medios de unos segundos en el caso de tener la pantalla encendida, a en muchos casos, cercanos al minuto y medio al apagar la pantalla. Este efecto se nota aún más en terminales como los Motorola Moto G [42], donde en el caso de la pantalla apagada su valor medio alcanza los 8 y 12 minutos. Cuando los dispositivos están conectados a una red, la diferencia entre tener la pantalla encendida es aún más notable, alcanzado valores atípicos de horas y ascendiendo los valores medios desde cerca del minuto hasta los 10 en la mayoría de los dispositivos.

Estos resultados son normales, ya que los teléfonos inteligentes están programados para que, en situación en las que no sea necesaria la búsqueda de redes Wi-Fi, se evite el envío constante de tramas del tipo *Probe Requests* con el objetivo de priorizar el ahorro energético.

También es importante destacar la diferencia de comportamiento entre dos teléfonos del mismo modelo como son los Motorola Moto G, los cuales aportan resultados con ciertas diferencias. Además, el caso del iPhone también muestra resultados contradictorios, ya que sus valores medios son más elevados en el caso de tener la pantalla encendida que en las situaciones en las que está apagada.

Estas observaciones permiten comprobar la dificultad para sacar unas conclusiones globales ya que los comportamientos de los teléfonos no siempre se acercan a lo esperado.



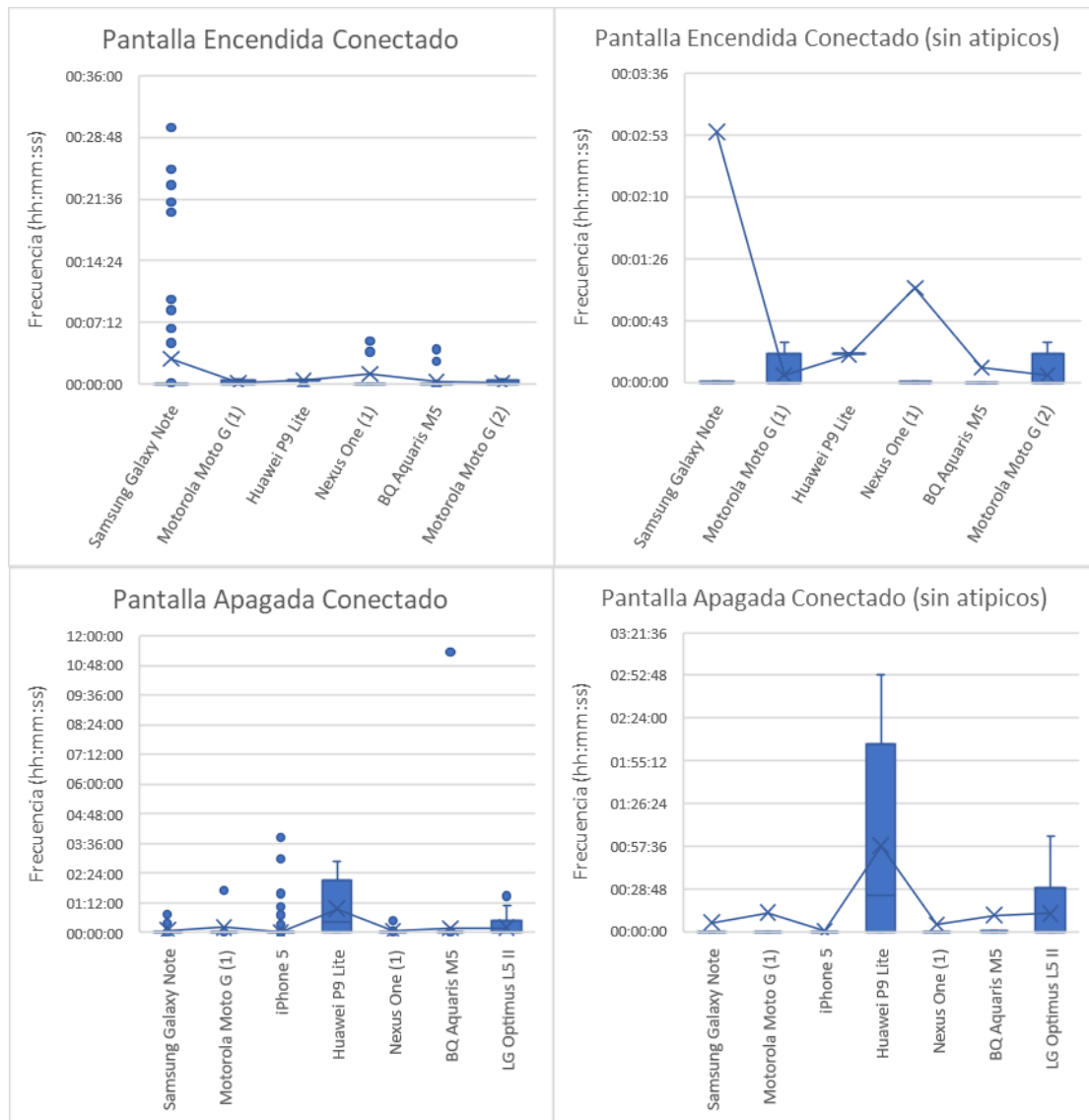


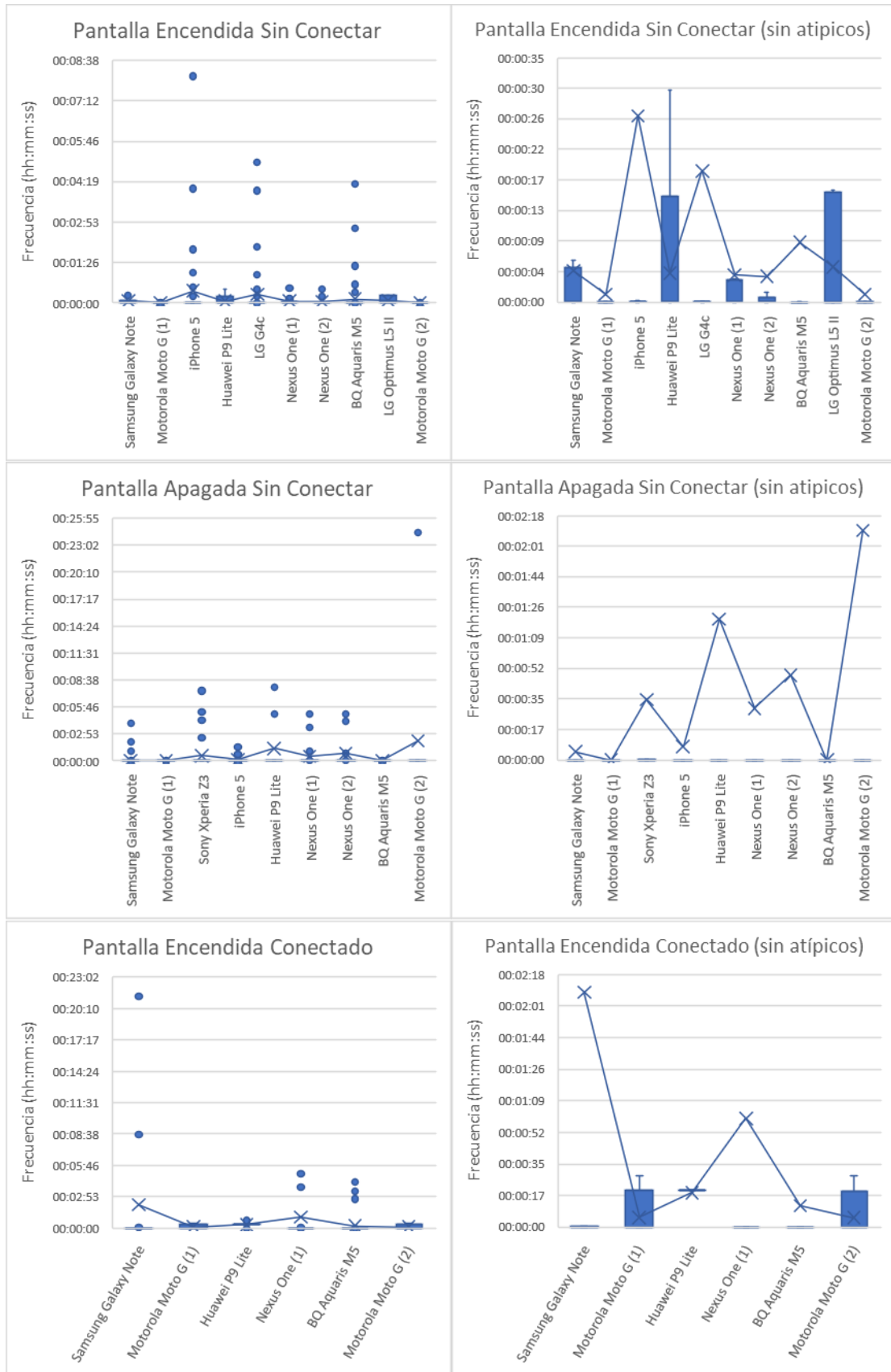
Figura 26 – Comparativa de la frecuencia de envío de los beacons en diferentes escenarios posibles (Izquierda: Todos los valores – Derecha: Solo valores típicos)

Por otro lado, estos resultados se podrían calificar como muy negativos, puesto que los valores parecen altos a priori, pero esto se debe a que están afectados por un estudio en un largo periodo de tiempo, en el cual los móviles con la pantalla apagada aumentan notablemente la frecuencia.

Para disponer de resultados en un caso más real, en la Figura 27 se representan los valores medidos durante los primeros 30 minutos, los cuales pueden ejemplificar a una persona que vaya a la terminal a esperar a un viajero. Centrándose en las 3 primeras comparaciones, los tiempos medios caen hasta valores por debajo del minuto en la mayoría de los dispositivos, lo cual plantea un escenario mucho más positivo que el anterior. Por otro lado, de la comparativa del peor caso en los primeros 30 minutos, se puede observar como cada teléfono se comporta de una manera muy distinta. Observando el grafico con valores atípicos, se puede ver como algunos teléfonos tras unos envíos de *Probe Request* iniciales, se mantiene en reposo hasta casi la media hora.

Finalmente, se puede concluir que teniendo en cuenta la tendencia generalizada al uso del teléfono inteligente durante las esperas, y que, a pesar de tenerlo en reposo, si no se

está conectado a una red la frecuencia sigue siendo aceptable, se prevé probable poder posicionar los dispositivos en la mayoría de las situaciones.



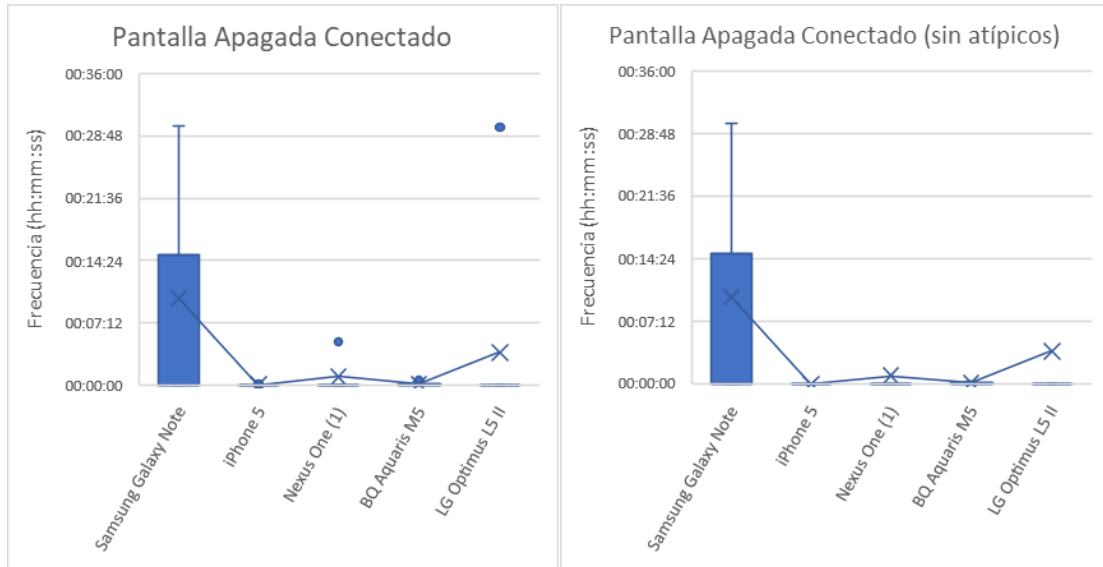


Figura 27 – Comparativa de la frecuencia de envío de los beacons en diferentes escenarios posibles durante los primeros 30 min (Izquierda: Todos los valores – Derecha: Solo valores típicos)

Tras realizar este análisis, se llevó a cabo una simulación de medidas dentro del laboratorio, con el objetivo de encontrar posibles errores antes de antes de las medidas reales en el aeropuerto. En esta simulación se tomaron datos en diferentes puntos del laboratorio con los móviles conectados a una red Wi-Fi y desconectados, con el objetivo de ver posibles alteraciones de la potencia debido a este suceso. Analizando las gráficas se pudo concluir que no había ninguna relación explícita y, por tanto, no hay a priori nada destacable que merezca la pena incluir en la memoria.

Realizados los análisis previos en el laboratorio, se procedió a la instalación en la terminal con un esquema como el de la Figura 18. Para ello, se distribuyeron los dispositivos sobre el área como se muestra en la Figura 28, donde la posición exacta de cada RP está representada con una cruz junto al número asignado. Para realizar este despliegue se utilizó la instalación técnica disponible en la terminal, que estaría representada en la Figura 18 con las conexiones en color blanco.

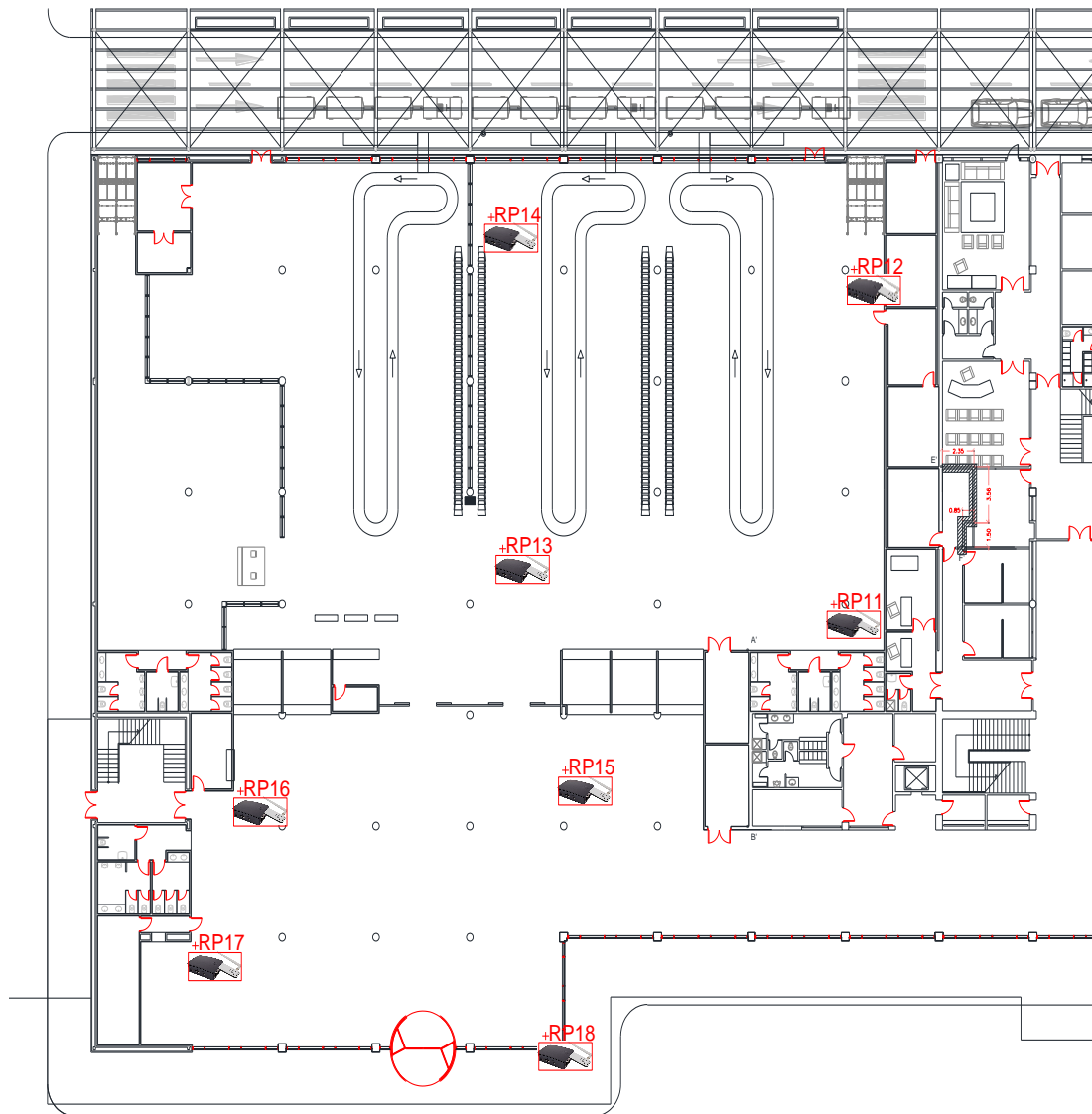


Figura 28 – Despliegue de dispositivos en la terminal

## 2. ALGORITMO DE DISCRIMINACIÓN DENTRO/FUERA

Realizado el despliegue de los dispositivos, el siguiente paso consistió en la elección del algoritmo que permitirá posicionar las concentraciones de personas. Para llevarlo a cabo, se realizó un *script* en Python [43] en el que se utilizó principalmente el paquete de Machine Learning, Scikit-Learn [44], el cual incluye librerías para poder trabajar de forma ágil con distintos estimadores. Previamente, con el fin de corroborar el correcto funcionamiento de los dispositivos sobre el terreno, se llevaron a cabo un conjunto de medidas iniciales en las posiciones marcadas en la Figura 29.

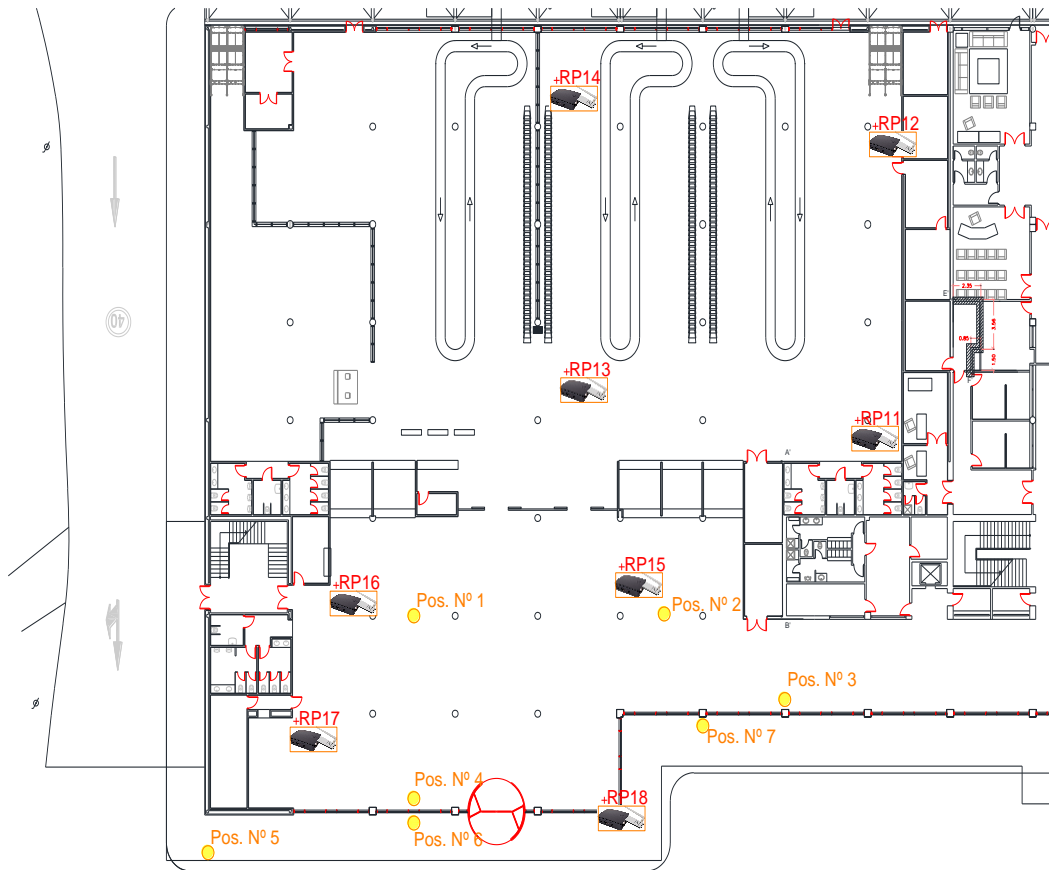


Figura 29 – Posiciones de las primeras medidas realizadas en la terminal.

La realización de este grupo de medidas permitió comprobar que los dispositivos operaban correctamente. Además, del análisis de los datos obtenidos, se pudieron extraer diferentes conclusiones.

Para comprobar el escenario de decisión al que se iba a enfrentar el algoritmo, en la Figura 31 se muestran los datos capturados en los puntos 1, 4 y 6 de las posiciones marcadas en la Figura 29. En estos se pueden apreciar importantes diferencias en función del lugar donde se realice la medida.





*Figura 30 – Sala en la que se realizaron las medidas*

En primer lugar, en la posición número 1, al ser bastante céntrica teniendo en cuenta cómo se han distribuido las RP, casi todas obtienen los mismos valores en torno a -70 dBm. Sin embargo, se puede apreciar las RP11, RP12 y RP14, al encontrarse en las posiciones más alejadas de la sala de recogidas de equipajes, los niveles de potencia recibida descienden por debajo de -80 dBm. Además, se puede apreciar claramente cómo capturan un número de tramas menor.

Analizando el punto 4, los datos obtenidos anteriormente se hacen más acentuados. Las RP de la sala en la que se realiza la medida obtienen valores notablemente superiores a los de la zona de recogidas de equipajes y, como ocurrían en el caso anterior, el número de tramas capturadas es inferior en las RP colocadas en zonas donde la señal recibida es menor. Además, la RP17, al estar situada cerca del punto donde se realizó la medida, alcanza unos valores por encima del resto de RP.

Finalmente, analizando el punto 6, el cual se encuentra fuera de la terminal, se observa que no sólo baja la potencia recibida en todas las RP, sino que las RP11 y RP12 ni tan siquiera son capaces de capturar las correspondientes tramas.



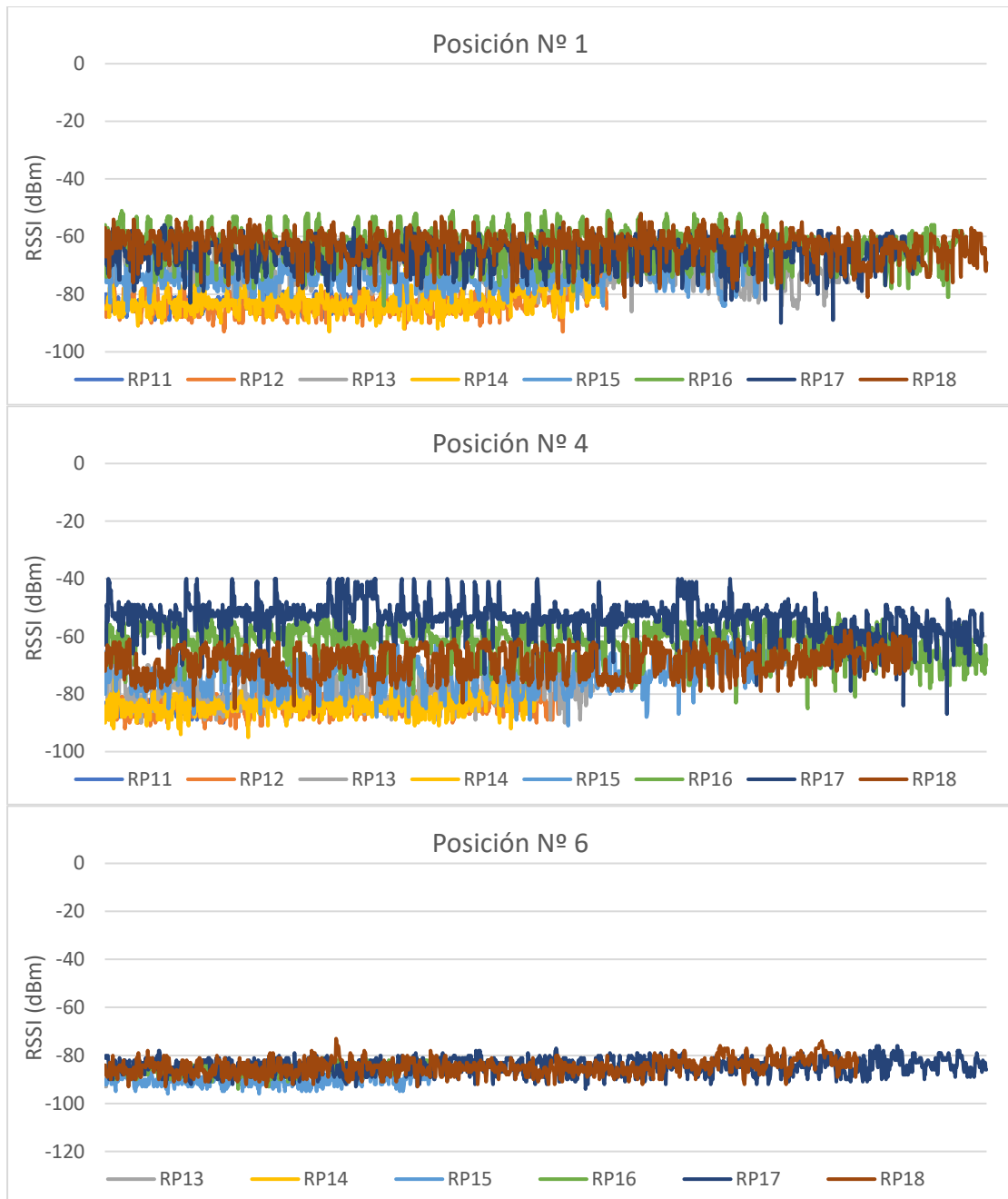


Figura 31 – Datos capturados en las posiciones 1, 4 y 6.

Comparando las valoraciones de cada punto, se puede concluir que, a pesar de que WiFi es una tecnología en la que hay importantes fluctuaciones del nivel de señal en función del estado del canal, es posible discriminar cuando un dispositivo se encuentra dentro o fuera de la terminal, y en qué sala en caso de encontrarse en el interior de aquella. Por tanto, parece oportuno utilizar clasificadores basados en regresiones logísticas y bayesianas.

Como estos algoritmos precisan de una fase de entrenamiento, se realizaron medidas desde numerosos puntos de la terminal. Con el objetivo de obtener muestras de todo el terreno de la mejor manera posible, se dividieron las salas como se muestra en la Figura 32. Desde cada círculo amarillo representado en la Figura 32, se tomaron medidas con los móviles utilizados anteriormente en el laboratorio. De este modo, se pudieron capturar tramas desde cada zona de las salas, y desde el exterior de la terminal.

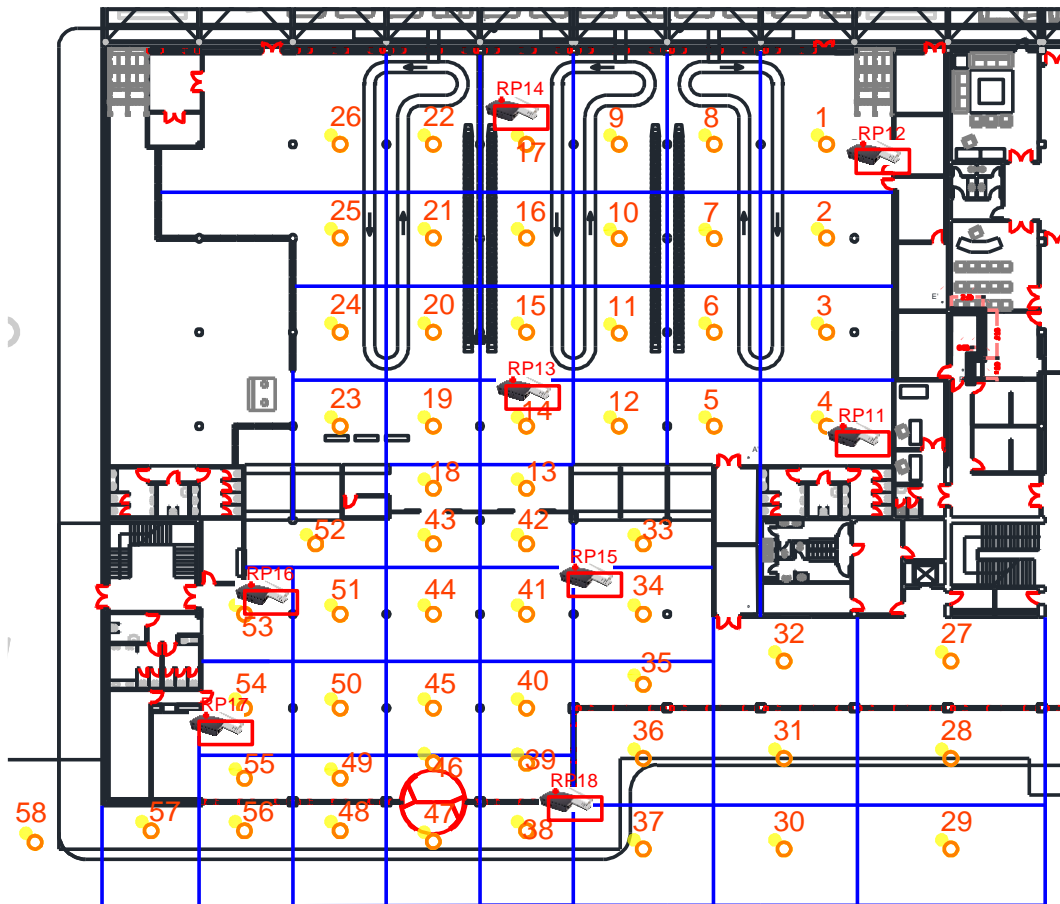


Figura 32 – División de la terminal

Con los datos capturados de estas medidas, se pudieron entrenar los algoritmos de clasificación. Para introducir estos datos al sistema, fue necesaria una etapa previa de preparación de estos. De esta fase se obtuvieron dos conjuntos de datos:

- **Datos agrupados por secuencia:** El primero de los conjuntos consistía en agrupar los datos por el número de secuencia y dirección MAC. De este modo, en cada fila se introducían los valores de potencia de cada RP para cada trama.
- **Datos agrupados por segundo:** En este conjunto se agruparon los datos capturados por cada dirección MAC en cada segundo. Es decir, en cada fila podía haber valores de potencia de diferentes tramas, con la condición de que se capturaran dentro del mismo segundo.

En ambos casos, las celdas vacías de cada fila se han completado con valores de -999 dBm. Con estos se pretendía fijar una potencia sobradamente baja que refleje la incapacidad del sistema para detectar esas tramas en las RP correspondientes.

Como se comentó en el capítulo II, los datos que se introducen deben cumplir una serie de condiciones. Entre ellas, se encuentra la independencia de las entradas. Gracias a las librerías de Python, Seaborn [45] y Pyplot [46] es posible representar la matriz de correlación de las entradas. En la Figura 33 y la Figura 34 se muestra ésta a partir de los datos de las diferentes RP. Las gráficas muestran como la correlación entre los datos es muy cercana a 0 en la mayoría de los casos, aunque existe la excepción cuando se comparan RP de las mismas salas, provocando un aumento de este valor.

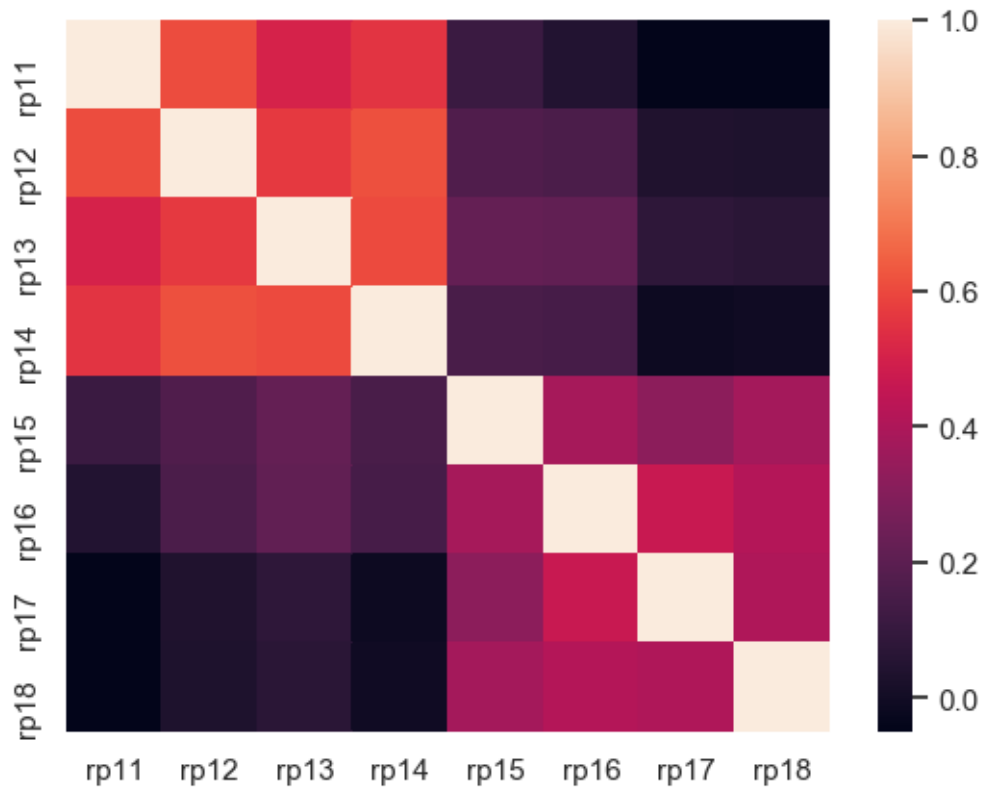


Figura 33 – Matriz de correlación de los datos de entrada agrupados por número de secuencia

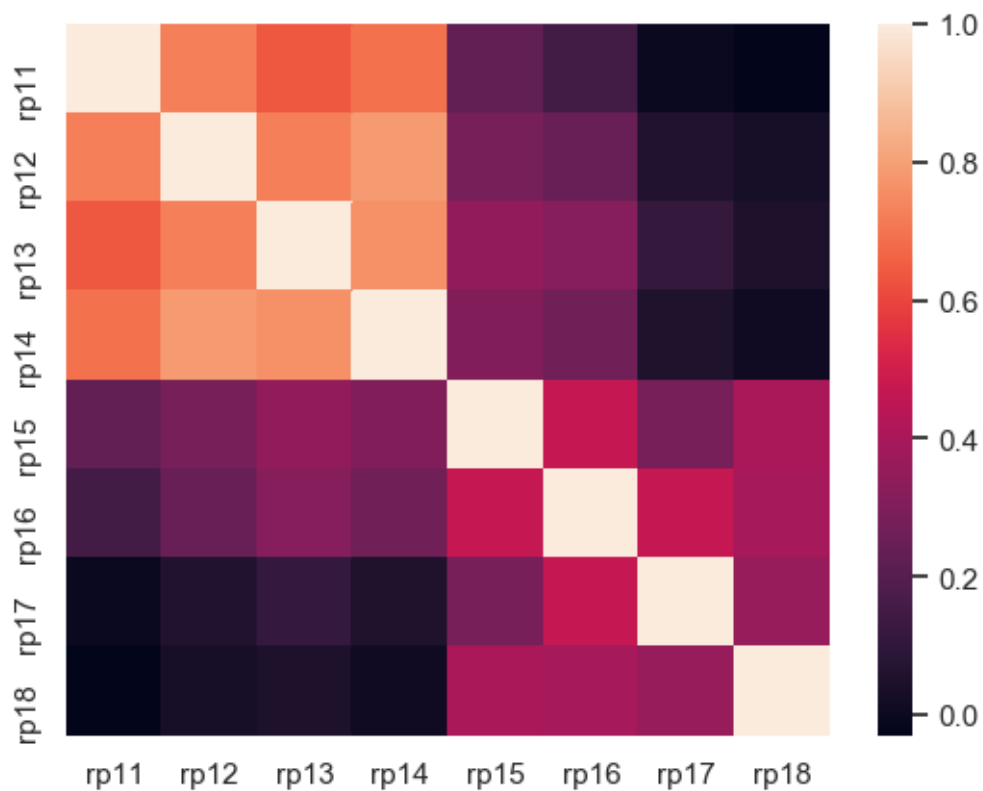
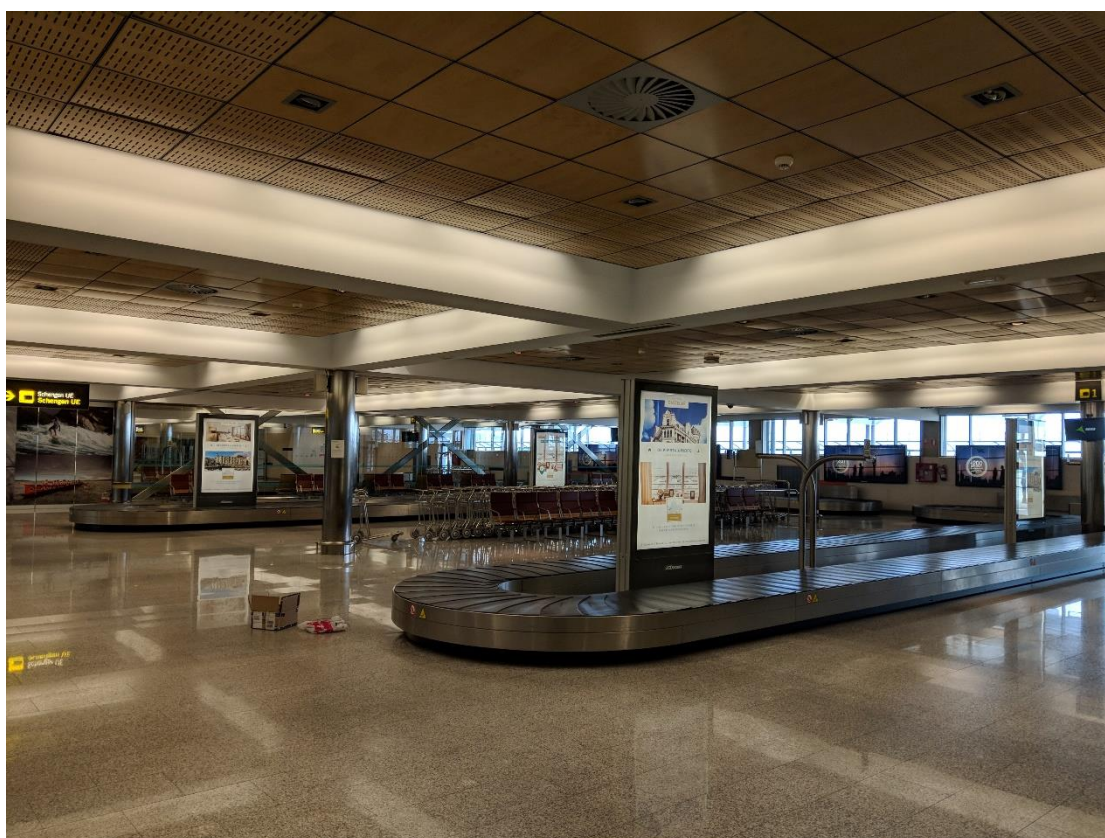


Figura 34 – Matriz de correlación de los datos de entrada agrupados por segundo

De estas agrupaciones de datos, se extrajo de forma aleatoria un subconjunto de aquellos con los que se realizó el entrenamiento de los algoritmos. Después, con el objetivo

de comprobar la precisión alcanzada, se testeó cada clasificador con el conjunto completo de los datos de entrada. De este modo se pudieron obtener una serie de resultados que serán analizados en el siguiente capítulo.

Posteriormente, con el fin de analizar el funcionamiento del algoritmo se realizó un conteo de personas en la sala de recogidas de equipajes, la cual se muestra en la Figura 35, durante diferentes vuelos con la intención de compararlo con los resultados del algoritmo entrenado, y poder concluir si su funcionamiento era el deseado. Para realizar este proceso, se esperó a que entrara el último viajero a la sala de recogidas de equipajes y, a partir de ese momento se comenzaron a contabilizar las salidas de pasajeros de la sala. De este modo, realizando la suma de todos los pasajeros que habían salido, se conseguía saber el número de personas en la sala en el momento del inicio del conteo, de forma que es posible compararlo con los datos estimados por los clasificadores.



*Figura 35 – Sala de recogida de equipajes*

En la Figura 36 se muestran los datos obtenidos del proceso de conteo de personas. En la tabla aparecen la fecha y hora de cada uno de los vuelos, el número máximo de viajeros desde el inicio del conteo y la hora a la que accedió a la sala el último pasajero. En el caso del último vuelo, al tratarse de un origen fuera de la zona Schengen, la recogida de equipajes se realizó en la cinta izquierda de la Figura 28, con unas medidas de seguridad superiores a las del resto de casos. A diferencia de los otros vuelos, los pasajeros no entraban en bloque a la sala de recogida de equipajes, si no que debían pasar un control policial anteriormente. Esto hizo que fueran entrando de uno en uno a medida que iban superando el control, y como ya se habían comenzado a entregar los equipajes, cada viajero recogía el suyo y abandonaba la sala. Estas circunstancias impidieron aplicar el sistema de los otros vuelos, ya que la entrada de gente fue constante durante todo el tiempo de medida. Ello conllevó que no se fijara un instante concreto como momento de máximo número de personas, si

no que se utilizaran los datos para, teniendo en cuenta estas condiciones, analizar el comportamiento del sistema.

Vuelo	Max. Personas	Hora entrada ultimo pasajero
23/08/2018 13:30	156	13:35
24/08/2018 13:05	65	13:16
24/08/2018 18:50	180 <sup>3</sup>	- <sup>4</sup>

Figura 36 – Resultados del conteo de personas

### 3. PRESUPUESTO DEL PROYECTO

En este tercer apartado del capítulo, se pretende recoger parte de la información relacionada al coste del equipamiento, con el objetivo de presentar a AENA un sistema que permita controlar la concentración de personas en el Aeropuerto Seve Ballesteros-Santander. Para analizar los costes se ha tenido en cuenta simplemente la inversión de los dispositivos.

Dispositivo	Coste Unidad	Unidades	Total Dispositivo
Raspberry Pi model 3 B+	29,47 €	8	235,76 €
TL-WN722N	8,47 €	8	67,76 €
TL-POE10R	13,21 €	8	105,68 €
Mikrotik RB960PGS	83,53 €	2	167,06 €
TL-MA260	28,93 €	1	28,93 €
		<b>TOTAL</b>	<b>605,19 €</b>

Figura 37 – Presupuesto del equipamiento

En la tabla se recogen la tabla de precios de cada uno sin incluir IVA. Para este ejemplo se han tenido en cuenta las cantidades utilizadas en el proyecto, aunque en función del caso de aplicación, estas cifras podrían verse alteradas.

Por otro lado, para la alimentación a través de PoE se han utilizado los adaptadores TL-POE10R, para los cuales hubo que realizar modificaciones en los cables de alimentación que llegan hasta las RP debido a los conectores de estas. Por tanto, de cara a un proyecto comercial, sería interesante sustituirlos por el Raspberry Pi POE Hat [47], el cual ha sido lanzado durante la realización de este trabajo, por lo que no pudo utilizarse en el mismo.

Además del coste de los equipos, también se debe tener en cuenta el tiempo invertido durante todo el desarrollo del proyecto. En la siguiente tabla se recogen estos datos divididos en tres periodos:

<sup>3</sup> Número de personas total en el vuelo.

<sup>4</sup> No se estableció la hora del ultimo pasajero.

Fase	Horas	€/Hora	Total Fase
Diseño	300	50,00 €	15.000,00 €
Despliegue	30	50,00 €	1.500,00 €
Puesta en funcionamiento	100	50,00 €	5.000,00 €
<b>TOTAL</b>			<b>21.500,00 €</b>

*Figura 38 – Presupuesto referente al desarrollo del proyecto (prototipo)*

Por lo tanto, el coste total del proyecto ascendería a los 22.105,19€. En caso de aplicar esta solución a otro aeropuerto, los costes se reducirían notablemente, ya que en ese caso se trataría más de una tarea de eventual escalado y ajuste de umbrales partiendo del sistema ya diseñado y validado.

## IV. RESULTADOS

Finalizada la etapa de desarrollo y pruebas de funcionamiento, se caracterizó el comportamiento del sistema. A pesar de que los datos de entrenamiento se obtienen de forma aleatoria en cada ejecución, los resultados son bastante parecidos y se pudo obtener información sobre la calidad de los clasificadores. En primer lugar, se analizó la precisión de estos al diferenciar si las personas se encontraban dentro o fuera de la terminal. Los resultados fueron los siguientes:

Tipo	Unidad	Porcentaje de Acierto	C	Min. Celdas
Naive Bayes	dBm	92.89 %	-	5
Naive Bayes	mW	95.15 %	-	5
Regresión Logística	dBm	97.82 %	100	3
Regresión Logística	mW	97.51 %	0.001	3

Figura 39 – Porcentaje de acierto del entrenamiento dentro/fuera con los datos agrupados por segundo.

Tipo	Unidad	Porcentaje de Acierto	C	Min. Celdas
Naive Bayes	dBm	86.23 %	-	4
Naive Bayes	mW	96.22 %	-	4
Regresión Logística	dBm	98.77 %	0.001	3
Regresión Logística	mW	98.77 %	0.001	3

Figura 40 – Porcentaje de acierto del entrenamiento dentro/fuera con los datos agrupados por secuencia.

En las tablas se muestran el tipo de clasificador, la unidad en la que se han introducido los datos, la precisión alcanzada para cada clasificador, el valor de C con el que se han configurado las regresiones logísticas, y el número mínimo de celdas con valor -999 que debe tener una fila para ser eliminada.

Observando los resultados, todos los clasificadores parecen adecuados para estimar si la gente se encuentra dentro o fuera de la terminal, pero posteriormente se verá como algunos de ellos no serán válidos.

Del mismo modo, en las siguientes tablas se muestran los resultados de una ejecución al clasificar por salas para el mejor estimador en cada caso.

Tipo	Unidad	Porcentaje de Acierto	C	Min. Celdas
Naive Bayes	dBm	90.29 %	-	5
Naive Bayes	mW	91.86 %	-	5
Regresión Logística	dBm	80.16 %	100	7
Regresión Logística	mW	56.08 %	0.001	3

Figura 41 – Porcentaje de acierto del entrenamiento con 3 salas con los datos agrupados por segundo.

Tipo	Unidad	Porcentaje de Acierto	C	Min. Celdas
Naive Bayes	dBm	77,41 %	-	5
Naive Bayes	mW	93.12 %	-	3
Regresión Logística	dBm	83.52 %	1	3
Regresión Logística	mW	58.42 %	0.001	3

Figura 42 – Porcentaje de acierto del entrenamiento con 3 salas con los datos agrupados por secuencia.

Lógicamente, en este caso los porcentajes de acierto disminuyen respecto al caso de diferenciar simplemente entre dentro y fuera de la terminal, ya que se incluye una situación adicional. Analizando los resultados, se puede observar como al introducir los datos ordenados por secuencia se obtiene un porcentaje de acierto superior al de los clasificadores entrenados con conjuntos agrupados por segundo. Además, para obtener los mejores resultados en cada caso, se eliminan muchas más filas en el caso de agrupar los datos por secuencia, ya que la condición es notablemente más baja que cuando se ordenan por segundo.

Por otro lado, se puede apreciar como los clasificadores basados en Naive Bayes (NB) operan mejor al introducir los datos en escala lineal (mW), mientras que los basados en regresiones logísticas obtienen mejores porcentajes cuando reciben los datos directamente en escala logarítmica (dBm). También destaca el alto porcentaje de acierto obtenido con el clasificador bayesiano en dBm. Este valor tan alto se debe a una ejecución en la que la aleatorización de los valores de entrenamiento le ha sido favorable. Realizando varias ejecuciones, se observa como esa cifra varía entre valores algo por debajo del 80% a algunos, como este caso concreto, donde supera el 90% de precisión.

Al igual que ocurría con el clasificador dentro/fuera, no todas las opciones tendrán utilidad. Para comprobar esto, de la fase de entrenamiento se pudieron obtener las matrices de confusión para cada uno de los clasificadores. En las siguientes figuras se representan dichas matrices para cada uno de los 4 clasificadores en función del conjunto de datos de entrada elegido y si se entrenan para diferenciar salas o simplemente dentro/fuera.

En estas figuras se muestran comparan los valores reales con los estimados. Los números asignados a cada una de las posiciones son:

- **0:** Exterior
- **1:** Sala de espera
- **2:** Sala de recogida de equipajes



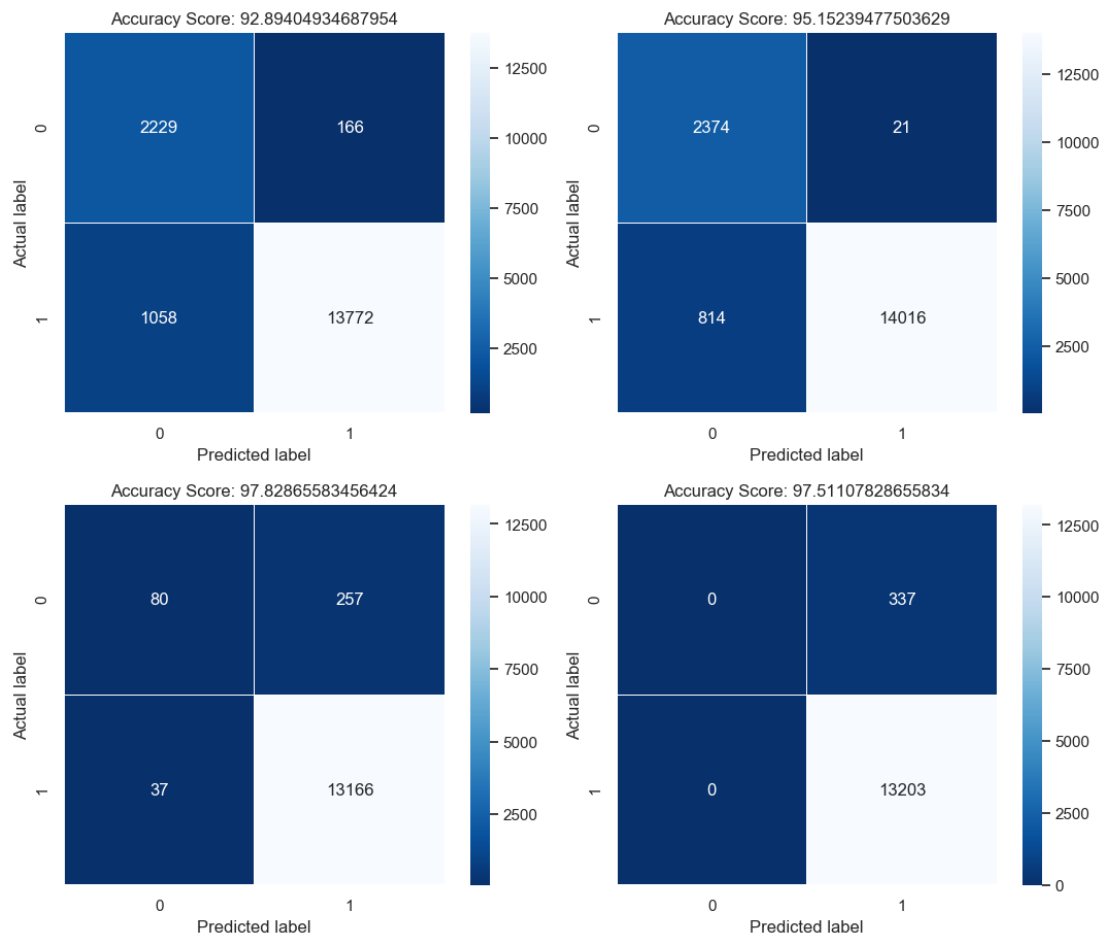


Figura 43 – Matrices de confusión de los clasificadores dentro/fuera entrenados con los datos agrupados por segundo. (Sup. Izq.: NB [dBm], Sup. Dcha.: NB [mW], Inf. Izq.: RL [dBm], Inf. Dcha.: RL [mW])

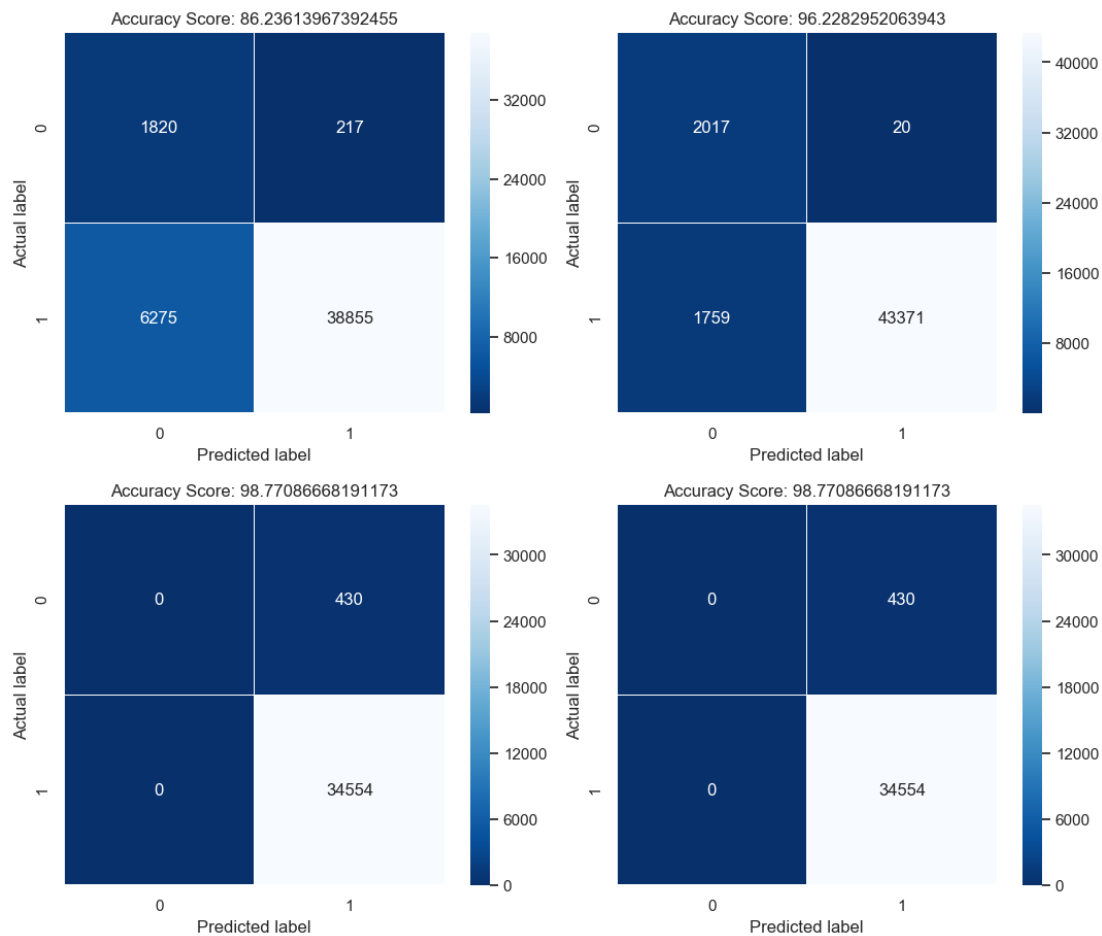


Figura 44 – Matrices de confusión de los clasificadores dentro/fuera entrenados con los datos agrupados por secuencia. (Sup. Izq.: NB [dBm], Sup. Dcha.: NB [mW], Inf. Izq.: RL [dBm], Inf. Dcha.: RL [mW])

Observando ambas figuras, se ve como algunos de los clasificadores que utilizan regresión lineal no estiman que los dispositivos estén fuera en ninguno de los casos. Esto se debe a que el *script* que elige los clasificadores con mejor porcentaje de cada tipo, prueba eliminando filas de los datos de entrada a partir de distintas cantidades de celdas a -999. En el caso de los clasificadores con regresiones logísticas, muchos de ellos comienzan a eliminar a partir de tres celdas, condición que cumplen muchas de las tramas capturadas en el exterior, ya que las cuatro RP de la zona de recogida de equipajes no reciben desde posiciones tan alejadas. Por tanto, estos clasificadores, se adaptan a los datos de entrada y posicionan todo en el interior, ya que apenas hay filas de fuera y, por tanto, así consiguen un mayor porcentaje de acierto.

A continuación, se muestran las matrices relacionadas con los clasificadores entrenados para distinguir entre tres salas distintas.

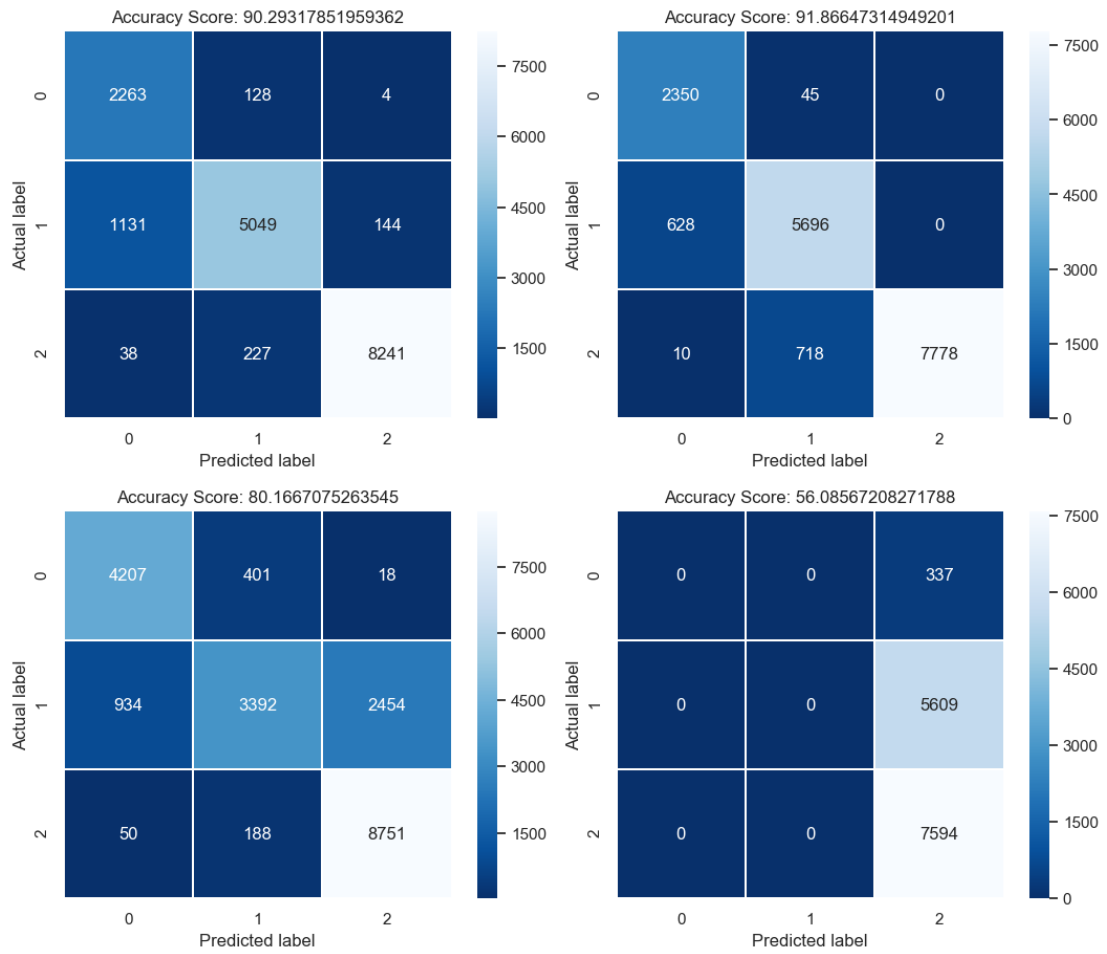


Figura 45 – Matrices de confusión de los clasificadores entrenados con 3 salas con los datos agrupados por segundo. (Sup. Izq.: NB [dBm], Sup. Dcha.: NB [mW], Inf. Izq.: RL [dBm], Inf. Dcha.: RL [mW])

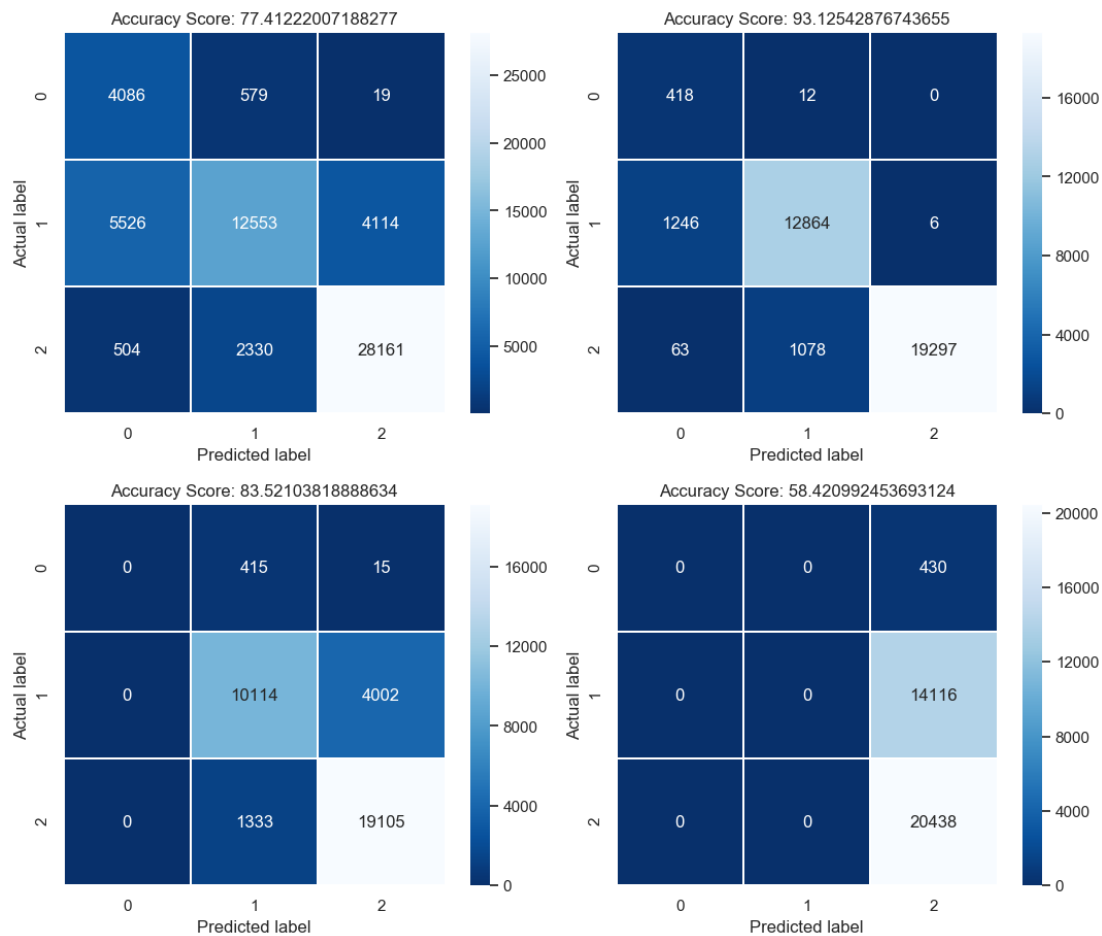


Figura 46 – Matrices de confusión de los clasificadores entrenados con 3 salas con los datos agrupados por secuencia. (Sup. Izq.: NB [dBm], Sup. Dcha.: NB [mW], Inf. Izq.: RL [dBm], Inf. Dcha.: RL [mW])

Analizando los resultados que se representan en estas figuras, del mismo modo que con los clasificadores dentro/fuera, se puede realizar un primer descarte de algunos estimadores. En el caso de realizar el entrenamiento con el conjunto de datos agrupado por los valores recibidos en el mismo segundo, el clasificado basado en Regresión Logística con los datos introducidos en mW, estima que todas las entradas se encuentran en la zona 2, por lo que obtiene un porcentaje de acierto tan bajo y, por tanto, no tiene utilidad.

Si se realiza el mismo análisis con los clasificadores entrenados por el conjunto agrupado por número de secuencia, los dos ejemplos de regresiones logísticas deberían ser descartados, ya que ninguno de los dos clasifica en la zona exterior.

Finalizada esta parte, el siguiente paso fue proceder al conteo de personas explicado en el capítulo III. De los datos obtenidos con el conteo manual en la propia terminal, se pudieron comparar con los estimados por los clasificadores entrenados para distinguir entre las tres salas. En las siguientes figuras se muestran diferentes graficas correspondientes a cada uno de los clasificadores que no fueron descartados en el proceso anterior, tanto para los entrenados con los datos agrupados por segundo como por número de secuencia. Además, se han obtenido las estimaciones sin eliminar filas, y eliminando las que tienen 3 o más celdas a “-999” para los dos primeros vuelos, y con 6 o más para el tercero, ya que las condiciones de este último no permitían eliminar más.

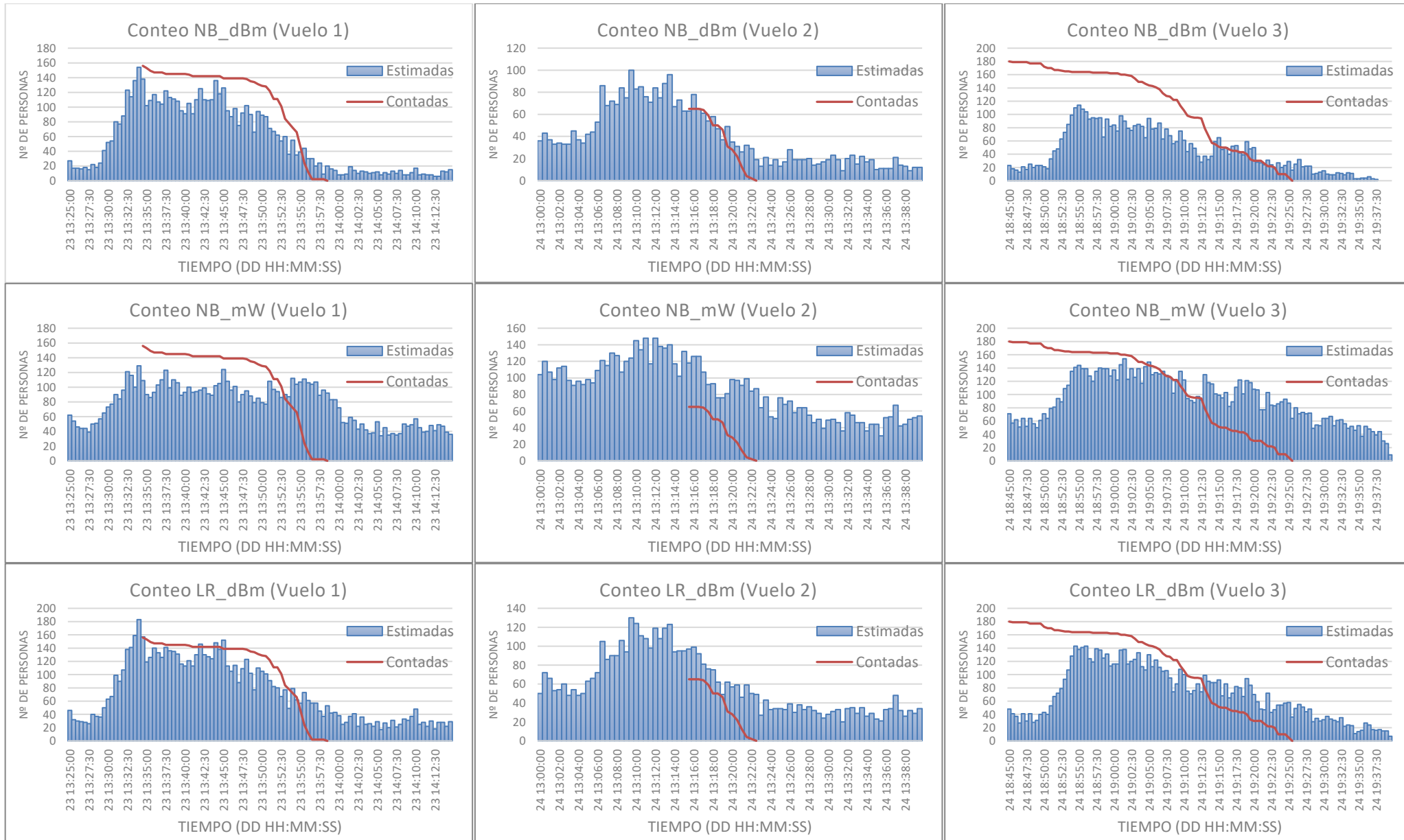


Figura 47 – Comparativa del conteo de personas manual y estimado con los datos agrupados por segundo (sin eliminar filas)

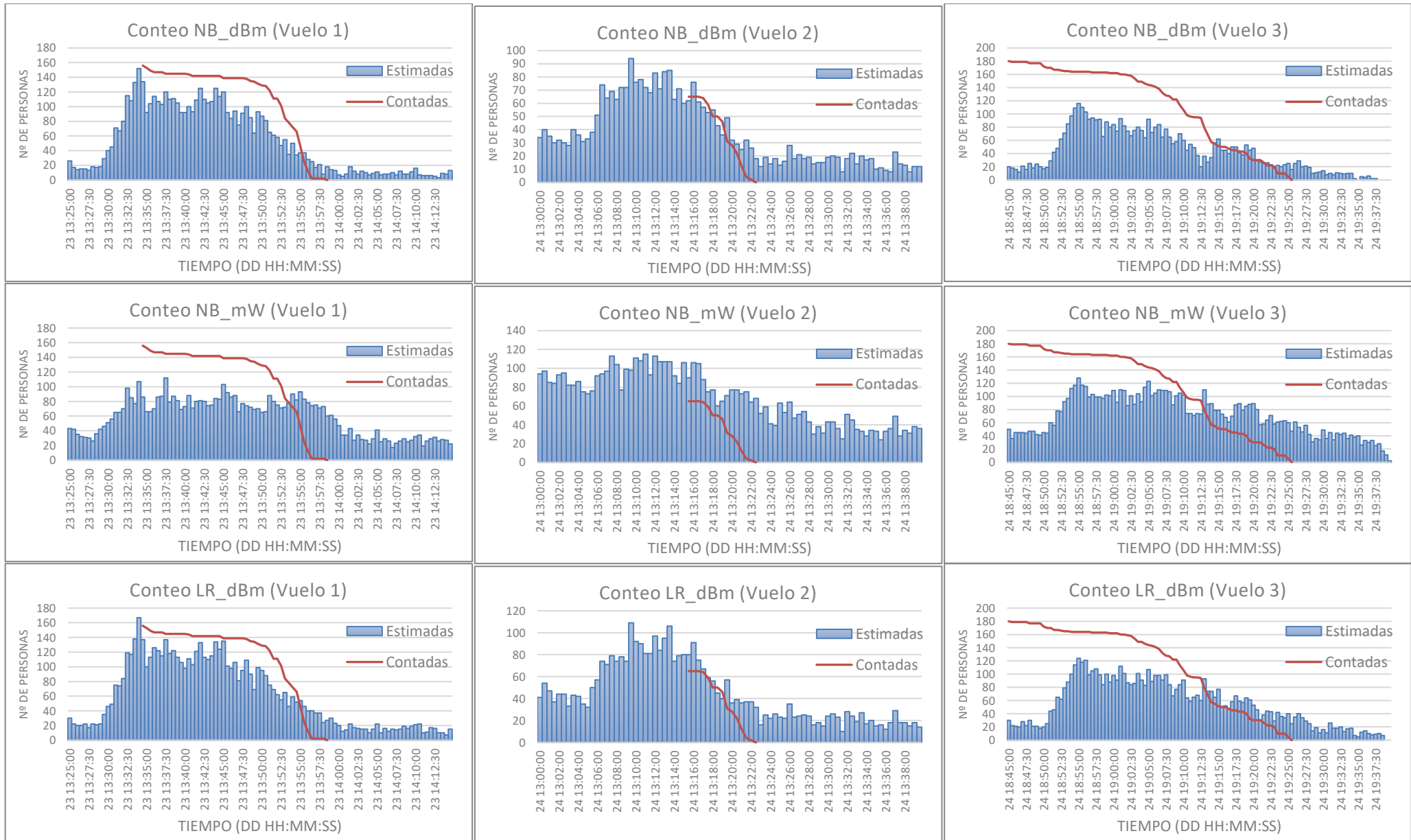


Figura 48 – Comparativa del conteo de personas manual y estimado con los datos agrupados por segundo (eliminando filas)

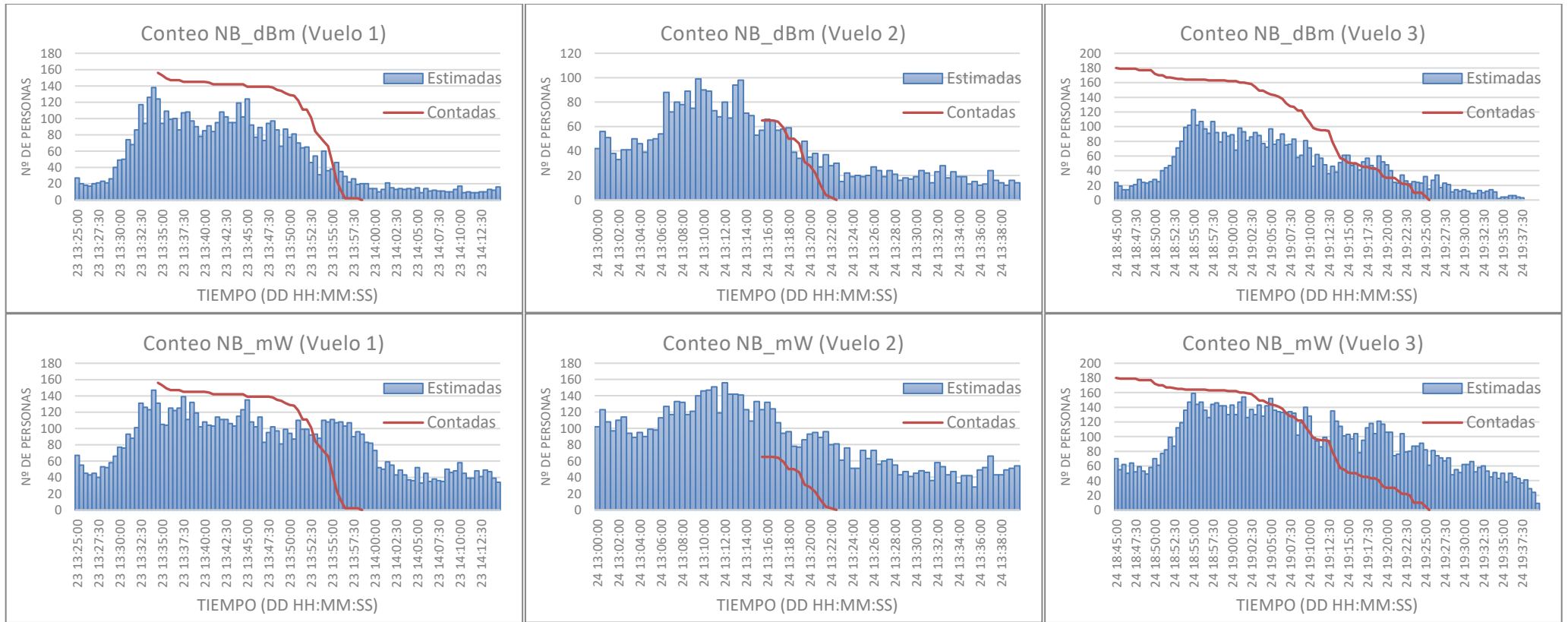


Figura 49 – Comparativa del conteo de personas manual y estimado con los datos agrupados por secuencia (sin eliminar filas)

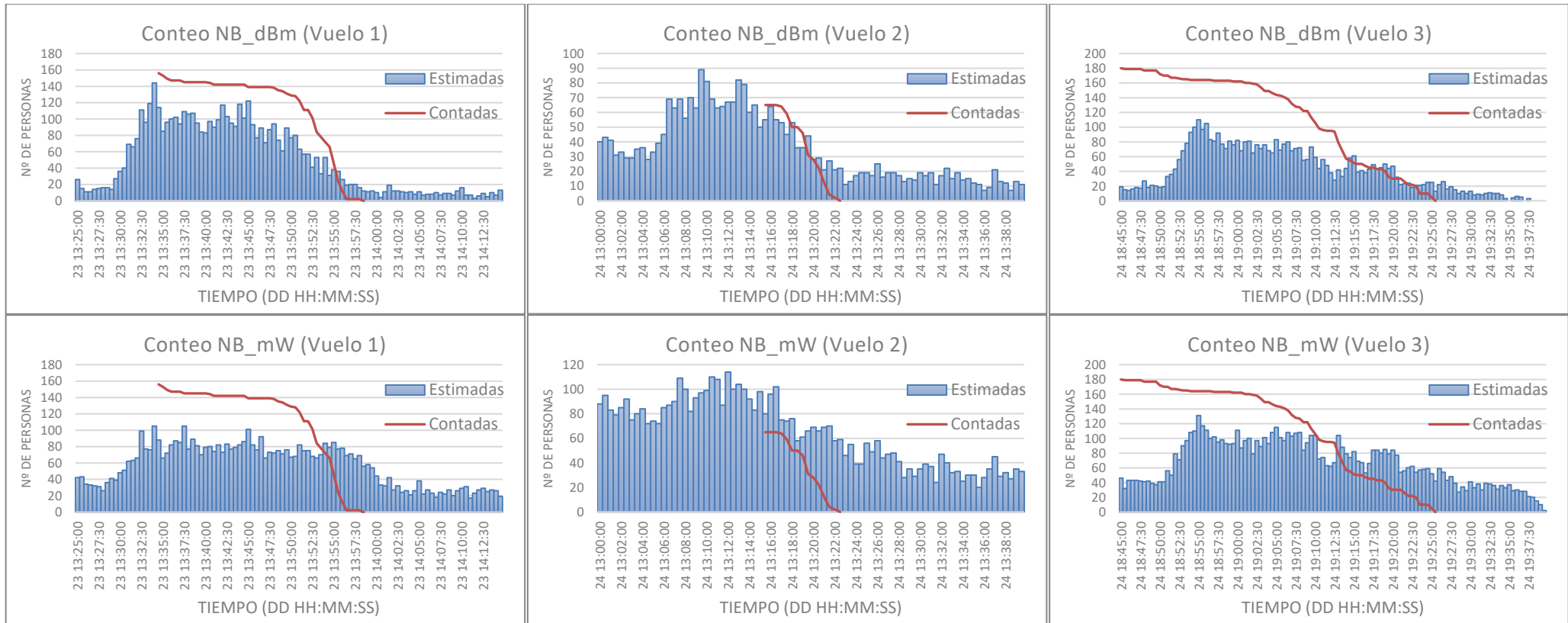


Figura 50 – Comparativa del conteo de personas manual y estimado con los datos agrupados por secuencia (eliminando filas)



En las gráficas se muestran los resultados obtenidos para los vuelos comentados en la Figura 36. En ellas se representan el número de personas estimadas por cada clasificador con un margen temporal anterior y posterior al vuelo, y el número de personas obtenido en el conteo manual. En los dos primeros vuelos, como se comentó anteriormente, se pudo establecer el momento de entrada del último pasajero para contabilizar las personas en la sala a partir de ese momento. En cambio, para el último, al no ser posible esto, se contó como si todos los pasajeros estuvieran en la sala desde el comienzo de las estimaciones.

Como era de esperar, al comparar los clasificadores entrenados por ambos conjuntos, los resultados tienen bastantes similitudes. Para el primer vuelo, todos ellos estiman por debajo del número de usuarios, mientras que, para el segundo, la estimación es superior en mayor o menor medida al valor real.

Por otro lado, también se puede apreciar como la cantidad de personas estimada, desciende ligeramente cuando se eliminan las filas con un número determinado de celdas a “-999”, esto demuestra que los clasificadores distinguen adecuadamente los dispositivos que se encuentran en posiciones alejadas.

Finalmente, con ayuda de los valores de las Figura 51 y Figura 52 se puede concluir que los valores obtenidos por los clasificadores no consiguen estimar la cantidad exacta de personas, debido a los factores a los que se enfrentan, por ejemplo, personas sin teléfono inteligente (niños) o dispositivos en modo avión o reposo que no emiten con la frecuencia necesaria para ser detectados cada treinta segundos. A pesar de esto, sí que se puede apreciar en las gráficas anteriores, como en muchos de los casos, especialmente cuando se introducen los datos en dBm, que el clasificador sigue una pendiente similar a la del conteo manual, por lo que sí permiten conocer de forma aproximada la cantidad de personas en la sala.

Fecha y Hora	Conteo	NB_dBm	NB_mW	LR_dBm
23-8-18 13:34:30	156	138	109	157
23-8-18 13:35:00	153	102	90	119
23-8-18 13:35:30	149	109	86	126
23-8-18 13:36:00	147	117	93	140
23-8-18 13:36:30	147	107	103	133
23-8-18 13:37:00	147	104	110	126
23-8-18 13:37:30	145	122	123	141
23-8-18 13:38:00	145	113	99	136
23-8-18 13:38:30	145	111	110	135
23-8-18 13:39:00	145	108	106	131
23-8-18 13:39:30	145	95	89	116
23-8-18 13:40:00	145	91	93	113
23-8-18 13:40:30	144	105	100	121
23-8-18 13:41:00	142	91	93	113
23-8-18 13:41:30	142	110	94	130
23-8-18 13:42:00	142	125	96	146
23-8-18 13:42:30	142	110	99	130
23-8-18 13:43:00	142	109	91	127
23-8-18 13:43:30	142	110	89	124
23-8-18 13:44:00	142	136	102	148
23-8-18 13:44:30	142	118	105	138
23-8-18 13:45:00	139	126	124	152
23-8-18 13:45:30	139	95	108	113
23-8-18 13:46:00	139	87	96	105
23-8-18 13:46:30	139	98	101	114
23-8-18 13:47:00	139	75	80	88
23-8-18 13:47:30	139	92	90	109
23-8-18 13:48:00	138	102	95	123
23-8-18 13:48:30	135	90	88	102
23-8-18 13:49:00	134	66	79	77
23-8-18 13:49:30	131	94	85	110
23-8-18 13:50:00	129	89	79	105
23-8-18 13:50:30	128	87	77	101
23-8-18 13:51:00	122	71	108	91
23-8-18 13:51:30	111	67	97	82
23-8-18 13:52:00	111	62	94	80
23-8-18 13:52:30	101	54	86	67
23-8-18 13:53:00	84	60	90	77
23-8-18 13:53:30	78	36	87	49
23-8-18 13:54:00	72	55	112	79
23-8-18 13:54:30	66	35	104	62
23-8-18 13:55:00	46	39	107	57
23-8-18 13:55:30	24	44	111	73
23-8-18 13:56:00	12	30	106	61
23-8-18 13:56:30	2	30	104	57
23-8-18 13:57:00	2	20	107	57
23-8-18 13:57:30	2	24	89	45
23-8-18 13:58:00	2	9	96	37
23-8-18 13:58:30	0	20	92	53

Figura 51 – Comparativa numérica del conteo de personas manual y estimado para el primer vuelo con los datos agrupados por segundo (sin eliminar filas)

Fecha y Hora	Conteo	NB_dBm	NB_mW	LR_dBm
24-8-18 13:15:30	65	63	118	97
24-8-18 13:16:00	65	78	126	99
24-8-18 13:16:30	65	65	126	92
24-8-18 13:17:00	64	61	107	81
24-8-18 13:17:30	59	54	92	76
24-8-18 13:18:00	50	58	93	75
24-8-18 13:18:30	50	47	76	62
24-8-18 13:19:00	46	37	76	49
24-8-18 13:19:30	31	49	81	62
24-8-18 13:20:00	28	35	98	57
24-8-18 13:20:30	22	31	97	59
24-8-18 13:21:00	12	26	91	46
24-8-18 13:21:30	4	32	99	59
24-8-18 13:22:00	2	29	84	50
24-8-18 13:22:30	0	19	87	49

*Figura 52 – Comparativa numérica del conteo de personas manual y estimado para el segundo vuelo con los datos agrupados por segundo (sin eliminar filas)*

---

## V. CONCLUSIONES Y LINEAS FUTURAS

---

Tras la realización de las diferentes etapas del proyecto, se puede concluir que el mismo ha servido para caracterizar las prestaciones de los dispositivos IoT en las tareas de monitorización de concentraciones de personas en el entorno de las terminales aeroportuarias.

Aunque a primera vista pueda parecer que algunos de los resultados obtenidos no son satisfactorios, ya que muestran una diferencia importante con los datos medidos, sí que se puede constatar la relación entre las pendientes de las gráficas obtenidas de la estimación y la presencia real de pasajeros. Esto implica un seguimiento de la tendencia de las medidas reales por parte de los valores estimados. Además, de la comparativa de los diferentes clasificadores realizada a lo largo del trabajo, es posible afirmar que entrenando estos con los datos en unidades dBm, se obtienen unos resultados mejores en la fase de conteo de personas. Esto proporciona una indicación en relación con cómo se deberían introducir los datos en una etapa comercial del sistema.

Por otra parte, los resultados logrados con los clasificadores basados en regresiones logísticas son algo más precisos que los que utilizan Naive Bayes, lo que al igual que con la forma de introducir los datos, aporta información acerca de cómo sería el sistema final.

Finalmente, cabe hacer notar que el sistema no está completamente desarrollado para producción, ya que los resultados del conteo no son lo suficientemente precisos como para otorgarle la responsabilidad en la gestión de situaciones de emergencia. A pesar de todo esto, sí que se sientan las bases para lograrlo en un futuro cercano.

Al tratarse de un proyecto práctico, en el que los tiempos pueden verse afectados con facilidad impidiendo llevar a cabo una fase más madura del sistema y a las nuevas ideas que han ido surgiendo durante la realización de este, se establecen las siguientes líneas futuras:

- **Clasificar dispositivos Bluetooth:** El sistema está capacitado para capturar tramas BLE, pero debido a que el número de dispositivos es actualmente inferior al de equipos WiFi, el proyecto se ha centrado en realizar las estimaciones a partir de este otro. Por tanto, sería interesante en un futuro explorar esta otra vía.
- **Inclusión de un coeficiente de estado del canal:** Como se ha comentado, cada RP tiene dos interfaces, lo que permite enviar por una y recibir mediante la otra. Por falta de tiempo no se ha empleado esta opción, pero gracias a ella, es posible realizar envíos periódicos desde cada una de las RP, de modo que el resto de RP podrán capturarlo, y al saber la posición donde está ubicada cada una de ellas, realizar una estimación del estado del canal.
- **Añadir información de vuelos:** Otra de las opciones que se plantearon durante el proyecto, es proporcionar información al sistema en relación con la inminencia de la llegada de un vuelo. Se podría desarrollar un escenario en el que el clasificador se adaptara a situaciones de llegada de vuelos y otras en las que la sala de recogidas de equipajes está completamente vacía.
- **Aprendizaje continuo:** Siguiendo la línea de las anteriores, se plantea una solución a través de otro clasificador en la que, utilizando las medidas del canal, y de la información recogida en tiempo real, este pueda adaptarse continuamente a la situación de la terminal en tiempo real, y de este modo, ser más preciso.

- **Explorar otras técnicas de aprendizaje máquina:** En este proyecto se ha utilizado solamente los clasificadores bayesianos y basados en regresiones logísticas ya que se consideraban los más apropiados. Otra de las opciones sería utilizar redes neuronales de forma que cada una de las RP fuera una de las neuronas y otras alternativas basadas en diferentes tipos de inteligencias artificiales.
- **Despliegue de un puesto de control:** Una vez se desarrolle más en profundidad el modelo presentado en este TFM, una posible línea de investigación es la implementación de este sistema, no sólo con la puesta en explotación de la infraestructura IoT, sino también mediante el despliegue de un puesto de control que a través de la información recogida por los dispositivos muestre la información en tiempo real en el propio centro de gestión de incidencias del aeropuerto. De esta manera se pretende que ayude al gestor aeroportuario a conocer las concentraciones de personas en el interior de la terminal, en cualquier momento del horario de funcionamiento del aeropuerto.
- **Despliegue de campañas comerciales:** Como en el caso anterior, la recogida de los datos por los dispositivos a lo largo del tiempo puede ayudar a la obtención de patrones de concentración de personas, que puede ayudar no sólo a determinar nuevas necesidades en la infraestructura aeroportuaria, sino también puede ayudar a desplegar campañas comerciales enfocadas a determinados tipos de clientes.

---

## VI. REFERENCIAS

---

- [1] K. Ashton, «That "Internet of Thing" Thing,» *RFID Journal*, Junio 2009.
- [2] G. Jayavardhana, B. Rajkumar, M. Slaven y P. Marimuthu, «Internet of Things (IoT): A vision, architectural elements, and future directions,» *Future Generation Computer Systems*, vol. 29, n° 7, pp. 1645-1660, 2013.
- [3] «www.nis-glonass.ru,» [En línea]. Available: [http://www.nis-glonass.ru/en/glonass/technical\\_descript/index.php](http://www.nis-glonass.ru/en/glonass/technical_descript/index.php). [Último acceso: Agosto 2018].
- [4] «GSA Europa,» [En línea]. Available: <https://www.gsa.europa.eu/european-gnss/galileo/galileo-european-global-satellite-based-navigation-system>. [Último acceso: Agosto 2018].
- [5] L. Hui, H. Darabi, P. Banerjee y J. Liu, «Survey of Wireless Indoor Positioning Techniques and Systems,» *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 37, n° 6, pp. 1067 - 1080, 2007.
- [6] L. M. Ni, Y. Liu y Y. C. Lau, «LANDMARC: indoor location sensing using active RFID,» *IEEE*, 2005.
- [7] C. Gao y R. Harle, «Easing the Survey Burden: Quantitative Assessment of Low-Cost Signal Surveys for Indoor Positioning,» *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2016.
- [8] «community.cisco.com,» 28 Agosto 2017. [En línea]. Available: <https://community.cisco.com/t5/wireless-mobility-documents/802-11-frames-a-starter-guide-to-learn-wireless-sniffer-traces/ta-p/3110019>. [Último acceso: Agosto 2018].
- [9] «Sharetechnote.com,» [En línea]. Available: [http://www.sharetechnote.com/html/WLAN\\_Protocol.html](http://www.sharetechnote.com/html/WLAN_Protocol.html). [Último acceso: Agosto 2018].
- [10] «iBeacon Apple,» [En línea]. Available: <https://developer.apple.com/ibeacon/>. [Último acceso: Julio 2018].
- [11] «Google Beacons,» [En línea]. Available: <https://developers.google.com/beacons/>. [Último acceso: Julio 2018].
- [12] «Google Developers,» [En línea]. Available: <https://developers.google.com/beacons/edystone>. [Último acceso: Julio 2018].
- [13] «Bluez,» [En línea]. Available: <http://www.bluez.org>. [Último acceso: Agosto 2018].
- [14] A. A. A. A.-S. M. A. A. A. S. A.-H. M. A. A.-A. y H. S. A.-K. , «Ultra Wideband Indoor Positioning Technologies,» p. 36, 2016.
- [15] R. Peng and M. L. Sichitiu, “Angle of Arrival Localization for Wireless Sensor Networks,” *IEEE*, North Carolina, 2006.
- [16] S. W. y L. D. S. , «Indoor Multipath Assisted Angle of Arrival Localization,» *Sensors*, vol. 17, n° 11, p. 29, 2 Noviembre 2017.
- [17] C. R. y M. , «TDOA Computation Using Multicarrier Modulation for Sensor Networks,» vol. 1, n° 1, p. 6, 2011.
- [18] S. W. S. J. E. M. L. X. C. C. M. P. y . J. H. , «Application of WiFi-based indoor positioning system for labor tracking at onstruction sites: A case study in Guangzhou MTR,» *Automation In Construction*, n° 20, pp. 3-13, 2010.

- [19] L. C. B. L. K. Z. C. R. y Z. Z. , «An Improved Algorithm to Generate a Wi-Fi Fingerprint Database for Indoor Positioning,» vol. 13, nº 8, pp. 11085-11096, 2013.
- [20] I. Moral Peláez, «Modelos de regresión: lineal simple y regresión logística,» *Revista SEDEN*, vol. IV, pp. 195-214, 3 Diciembre 2006.
- [21] J. M. Marín Diazaraque, «Universidad Carlos III de Madrid,» [En línea]. Available: <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/GuiaSPSS/18reglin.pdf>. [Último acceso: Agosto 2018].
- [22] «Hospital Universitario Ramón y Cajal,» [En línea]. Available: [http://www.hrc.es/bioest/Reglin\\_12.html](http://www.hrc.es/bioest/Reglin_12.html). [Último acceso: Julio 2018].
- [23] C. Malagón Luque, «Clasificadores Bayesianos. El algoritmo Naïve Bayes,» 2003.
- [24] «Situm,» [En línea]. Available: <https://situm.es/>. [Último acceso: Julio 2018].
- [25] «Applesfera.com,» [En línea]. Available: <https://www.applesfera.com/aplicaciones-ios-1/indoor-survey-apple-lanza-una-nueva-app-para-posicionamiento-en-interiores>. [Último acceso: Julio 2018].
- [26] «Maps Connect,» [En línea]. Available: <https://mapsconnect.apple.com/>. [Último acceso: 3 Julio 2018].
- [27] «Android Developers,» [En línea]. Available: <https://developer.android.com/training/location/>. [Último acceso: Julio 2018].
- [28] J. C. Alonso, «Desarrollo de una herramienta para la localización de dispositivos en entornos inalámbricos,» Santander, 2016.
- [29] Diseño y operaciones de Aeródromo, vol. 1, Montreal: OACI, 2009.
- [30] N. J. Ashford y S. A. Mumayiz, *Airport Engineering: Planning, Design, and Development of 21st Century Airports*, USA: John Wiley and Sons Inc, 2011.
- [31] J. R. Santana Martínez, J. Carrasco Alonso, J. A. Galache López, L. Sánchez González y R. Agüero Calvo, «A solution for tracking visitors in Smart Shopping environments: A real platform implementation based on Raspberry Pi,» *MONAMI*, 2016.
- [32] «Raspberrypi.org,» [En línea]. Available: <https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/>. [Último acceso: Agosto 2018].
- [33] «WireShark Wiki,» [En línea]. Available: [https://wiki.wireshark.org/CaptureSetup/WLAN#Promiscuous\\_mode](https://wiki.wireshark.org/CaptureSetup/WLAN#Promiscuous_mode). [Último acceso: Septiembre 2018].
- [34] «Tp-Link TL-WN722N,» [En línea]. Available: <https://www.tp-link.com/ar/products/details/TL-WN722N.html>. [Último acceso: Agosto 2018].
- [35] «TP-LINK TL-POE10R,» [En línea]. Available: [https://www.tp-link.com/es/products/details/cat-43\\_TL-POE10R.html](https://www.tp-link.com/es/products/details/cat-43_TL-POE10R.html). [Último acceso: Agosto 2018].
- [36] «Mikrotik hEx PoE,» [En línea]. Available: <https://mikrotik.com/product/RB960PGS>. [Último acceso: Agosto 2018].
- [37] «GitHub Pyshark,» [En línea]. Available: <https://github.com/KimiNewt/pyshark>. [Último acceso: Agosto 2018].
- [38] «GitHub Hcitol,» [En línea]. Available: <https://github.com/aspict/bluetooth-hcitol>.
- [39] «btmon,» [En línea]. Available: <https://manpages.debian.org/testing/bluez/btmon.1.en.html>. [Último acceso: Septiembre 2018].

- [40] «hctool,» [En línea]. Available: <https://www.systutorials.com/docs/linux/man/1-hctool/>. [Último acceso: Septiembre 2018].
- [41] «amby.net,» 15 Abril 2018. [En línea]. Available: <https://amby.net/2016/04/15/graficos-estadisticos-excel-2016-cajas-y-bigotes-12/>. [Último acceso: 10 Agosto 2018].
- [42] «SmartGSM,» [En línea]. Available: <https://www.smart-gsm.com/moviles/motorola-moto-g>. [Último acceso: Septiembre 2018].
- [43] «Python.org,» [En línea]. Available: <https://www.python.org/about/>. [Último acceso: Septiembre 2018].
- [44] «Scikit-learn.org,» [En línea]. Available: <http://scikit-learn.org>. [Último acceso: Septiembre 2018].
- [45] «Pydata.org,» [En línea]. Available: <https://seaborn.pydata.org/api.html>. [Último acceso: Septiembre 2018].
- [46] «Matplotlib,» [En línea]. Available: [https://matplotlib.org/api/plot\\_summary.html](https://matplotlib.org/api/plot_summary.html). [Último acceso: Septiembre 2018].
- [47] «POE-Hat,» [En línea]. Available: <https://www.raspberrypi.org/products/poe-hat/>. [Último acceso: 2018].