

UNIVERSIDAD POLITÉCNICA DE MADRID

**ESCUELA TÉCNICA SUPERIOR
DE INGENIEROS DE TELECOMUNICACIÓN**



**GRADO EN INGENIERÍA DE TECNOLOGÍAS Y
SERVICIOS DE TELECOMUNICACIÓN**

TRABAJO FIN DE GRADO

**DISEÑO DE UN SISTEMA DE DETECCIÓN DE
PRESENCIA Y MOVIMIENTO BASADO EN LOS
CAMBIOS DE LAS SEÑALES DE
RADIOFRECUENCIA**

**ÁLVARO GÓMEZ PAVÓN
JULIO 2023**

UNIVERSIDAD POLITÉCNICA DE MADRID

**ESCUELA TÉCNICA SUPERIOR
DE INGENIEROS DE TELECOMUNICACIÓN**



**GRADO EN INGENIERÍA DE TECNOLOGÍAS Y
SERVICIOS DE TELECOMUNICACIÓN**

TRABAJO FIN DE GRADO

**DISEÑO DE UN SISTEMA DE DETECCIÓN DE
PRESENCIA Y MOVIMIENTO BASADO EN LOS
CAMBIOS DE LAS SEÑALES DE
RADIOFRECUENCIA**

AUTOR

ÁLVARO GÓMEZ PAVÓN

TUTOR: LUIS FELIPE VÉLEZ FLORES

PONENTE: ALVARO ARAUJO PINTO

JULIO 2023

GRADO EN INGENIERÍA DE TECNOLOGÍAS Y SERVICIOS DE
TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

Título: DISEÑO DE UN SISTEMA DE DETECCIÓN DE PRESENCIA Y
MOVIMIENTO BASADO EN LOS CAMBIOS DE LAS SEÑALES DE
RADIOFRECUENCIA

Autor: D. Álvaro Gómez Pavón

Tutor: D. Luis Felipe Vélez Flores

Ponente: D. Alvaro Araujo Pinto

Departamento: Departamento de Ingeniería Electrónica

MIEMBROS DEL TRIBUNAL

Presidente: D.

Vocal: D.

Secretario: D.

Suplente: D.

Los miembros del tribunal arriba nombrados acuerdan otorgar la calificación
de:

Madrid, a de de 20...

Resumen

Con el avance de las tecnologías inalámbricas y las metodologías de detección, muchos estudios han demostrado el buen funcionamiento de las señales inalámbricas para detectar movimiento y actividades humanas. Esto ha permitido el desarrollo de diversas aplicaciones, como por ejemplo la detección de intrusos o el reconocimiento de actividad. Los movimientos del cuerpo humano afectan a la propagación de la señal inalámbrica, lo que brinda oportunidades para capturar los movimientos humanos mediante el análisis de las señales inalámbricas recibidas.

Este proyecto se ha realizado en la cátedra Securitas Direct Verisure, dentro de un programa de becas en áreas de interés para la empresa y que complementan el proceso formativo de los estudiantes en la Escuela Técnica Superior de Ingenieros de Telecomunicación perteneciente a la Universidad Politécnica de Madrid (ETSIT-UPM).

El principal objetivo del proyecto es analizar las distintas posibilidades para la detección de presencia y movimiento basado en cambios en la señal de radio, haciendo uso de diferentes algoritmos de *machine learning* para poder inferir resultados.

Palabras clave: detección de presencia, técnicas de detección, señales de radio-frecuencia, RSSI, algoritmos de decisión no supervisados, *k-means*, dispositivos de detección.

Abstract

With the advancement of wireless technologies and sensing methodologies, many studies have shown the success of re-using wireless signals to sense human activities and thereby realize a set of emerging applications, ranging from intrusion detection to daily activity recognition. The movements of the human body impact the wireless signal propagation, which provide great opportunities to capture human motions by analyzing the received wireless signals.

The main objective of this project is to analyze the different possibilities for the detection of presence and movement based on changes in the radio signal, using different machine learning algorithms to be able to infer results.

This project has been carried out in the Securitas Direct Verisure cathedra, within a scholarship program in areas of interest to the company and that complement the training process of students at Escuela Técnica Superior de Ingenieros de Telecomunicación belonging to the Universidad Politécnica de Madrid (ETSIT-UPM).

Keywords: presence detection, detection techniques, radio frequency signals, RSSI, unsupervised decision algorithms, k-means, detection devices.

Agradecimientos

Aquel niño que entró a la universidad con 18 años jamás habría imaginado encontrarse en este punto. Y así es, después de cuatro años de duro trabajo, esfuerzo y con una pandemia de por medio, llega el punto final. Han sido cuatro años que me han cambiado por completo tanto físicamente, como mentalmente. Por supuesto, no ha sido un camino de rosas, los momentos de angustia y estrés no han faltado, pero ahora puedo decir con orgullo, que todo esto ha merecido la pena.

En primer lugar, agradecer a mis padres por haberme brindado la oportunidad de recibir una educación universitaria. Siempre han sido un pilar fundamental en mi vida, y sobre todo en el ámbito académico. A mi madre María, por toda la atención y preocupación puesta en mí, jamás olvidaré las comidas que me preparabas para aguantar los días duros en la universidad. A mi padre Sergio, por todo el interés mostrado en cada proyecto que realizaba y sus ganas de explicarme siempre cada asignatura. Y por supuesto, a mi hermana Laura, por ser la causante de muchas risas y alegrías cada día que volvía a casa.

En segundo lugar, agradecer a mis abuelos. Maribel y Jesús, por acogerme siempre en vuestra casa y por el cariño inmenso que me dais. Maricarmen y Pepe, por centrarme la cabeza y enseñarme a como ser un buen ingeniero. Y por supuesto a mi "tercera abuela", Tuní, por cuidarme y ser siempre mi ángel de la guarda.

Agradezco sinceramente a mi tutor, Luisfe, por su valioso tiempo y conocimientos durante la realización de mi proyecto. Sin su guía y apoyo, este trabajo no habría sido posible. También agradecer a mi ponente Araujo, por compartir su experiencia y conocimiento. Por supuesto, agradecer a mis compañeros de la Cátedra, Rocío, Mikel y Omar, por hacer las jornadas de trabajo más amenas. Tampoco me quiero olvidar de todos los profesores y personal de la universidad por su ayuda y enseñanzas.

Por supuesto, agradecer a "la receta", mis compañeros en esta aventura. Gracias a ellos los días en la universidad han sido mucho más divertidos. Gracias también por las quedadas, viajes y momentos vividos juntos. Tampoco me olvido de mis amigos de toda la vida, personas que me han visto crecer desde los 3 años, y espero que sean muchos años más.

Por último, pero no menos importante, agradecerte a ti, Isabel. Entramos en la universidad como dos desconocidos, y tras 4 años salimos juntos siendo inseparables. Tengo claro que, sin ti, esto no habría sido posible. Eres mi compañera de carrera, de tardes y noches de estudio, de risas incontables en el metro... Gracias por haber estado conmigo en los buenos y sobre todo los malos momentos. Has logrado que sea la mejor versión de mí mismo.

Muchas gracias a todos los que han hecho esto posible, Álvaro.

Índice general

Resumen	VI
Abstract	VII
Agradecimientos	VIII
Índice General	X
Índice de Figuras	XI
Índice de Tablas	XIII
Lista de Acrónimos	XV
1. Introducción	1
1.1. Entorno y Motivación	1
1.2. Objetivo	2
1.3. Metodología	2
1.4. Tareas	3
1.5. Estructura del contenido	3
2. Estado del arte	4
2.1. Detección de presencia basada en señales inalámbricas	4
2.2. Intensidad de la señal recibida (RSSI)	6
2.3. Información del estado del canal (CSI)	8
2.4. Frecuencia modulada de onda continua (FMCW)	10
2.5. Desplazamiento Doppler	11
2.6. Conclusión del estado del arte	12
3. Desarrollo de escenario de pruebas	14
3.1. <i>RAIL-SoC range_test</i>	16
4. Algoritmos de decisión	21
4.1. Tipos de algoritmos	21
4.2. <i>Clustering</i>	22
4.2.1. Algoritmo <i>k-means</i>	23
4.2.2. <i>Hierarchical clustering</i>	24
4.2.3. Comparación entre <i>k-means</i> y <i>hierarchical clustering</i>	25

5. Pruebas y resultados	26
5.1. Escenario 1	28
5.2. Escenario 2	31
5.3. Escenario 3	34
5.4. Escenario 4	37
5.5. Análisis de los falsos positivos	40
5.6. Conclusiones de las pruebas	46
6. Conclusiones y líneas futuras	47
6.1. Conclusiones	47
6.2. Líneas futuras	48
Bibliografía	50
A. Aspectos éticos, económicos, sociales y ambientales	52
A.1. Impacto ético	52
A.2. Impacto económico	52
A.3. Impacto social	53
A.4. Impacto ambiental	53
A.5. Conclusiones	53
B. Presupuesto económico	54
C. Análisis de parámetros del algoritmo <i>k-means</i>	56

Índice de figuras

2.1. Ejemplo de medición RSSI para 0, 3 y 12 personas	7
2.2. Multitrayecto de las señales inalámbricas (elaboración propia)	8
2.3. Ejemplo de medición CSI para 0, 1, 3 y 5 personas	9
2.4. Desplazamiento FMCW (elaboración propia)	10
2.5. Ejemplo desplazamiento FMCW a 80 GHz	11
2.6. Ejemplo de cambio de frecuencia Doppler (elaboración propia)	12
3.1. Dispositivos de Seguridad Direct	14
3.2. Placa <i>EFR32FG</i> de Silicon Labs	16
3.3. Arquitectura <i>software</i> de los nodos (elaboración propia)	16
3.4. Configuración de las placas de Silicon Labs (elaboración propia)	17
3.5. Placas de Silicon Labs configuradas y preparadas para comenzar (elaboración propia)	19
4.1. Comparación de los tipos de algoritmos (elaboración propia)	22
4.2. Ejemplo de <i>clustering</i> o agrupamiento	22
4.3. Ejemplo de <i>hierarchical clustering</i> (elaboración propia)	24
5.1. Esquema del escenario 1 (elaboración propia)	28
5.2. Entrenamiento en el escenario 1	29
5.3. Algoritmo <i>k-means</i> en el escenario 1	30
5.4. Esquema del escenario 2 (elaboración propia)	31
5.5. Entrenamiento en el escenario 2	32
5.6. Algoritmo <i>k-means</i> en el escenario 2	33
5.7. Esquema del escenario 3 (elaboración propia)	34
5.8. Entrenamiento en el escenario 3	35
5.9. Algoritmo <i>k-means</i> en el escenario 3	36
5.10. Esquema del escenario 4 (elaboración propia)	37
5.11. Entrenamiento en el escenario 4	38
5.12. Algoritmo <i>k-means</i> en el escenario 4	39
5.13. Esquema de escenario de detección de falsos positivos (elaboración propia)	40
5.14. Predicciones para una caja pequeña	41
5.15. Predicciones para una caja mediana	42
5.16. Predicciones para una caja grande	43
5.17. Predicciones para una silla de oficina con dos clústeres	44
5.18. Predicciones para una silla de oficina con tres clústeres	45

C.1. Configuración de parámetro $init = k - means++$	57
C.2. Configuración de parámetro $init = random$	57
C.3. Configuración de parámetro n_init	59
C.4. Configuración de parámetro max_iter	61
C.5. Configuración de parámetro tol	62

Índice de tablas

B.1. Costes de personal	54
B.2. Costes de recursos materiales	55
B.3. Costes totales	55

Lista de Acrónimos

BLE: Bluetooth Low Energy.

CSI: Channel State Information.

FMCW: Frequency-Modulated Continuous Wave.

OFDM: Orthogonal Frequency Division Multiplexing.

PHY: Physical Layer.

RSSI: Received Signal Strength Indicator.

RX: Reception.

TX: Transmission.

UART: Universal Asynchronous Receiver / Transmitter.

USRP: Universal Software Radio Peripheral.

Capítulo 1

Introducción

1.1. Entorno y Motivación

Este proyecto se ha realizado en la Cátedra de Securitas Direct Verisure de la Escuela Técnica Superior de Ingenieros de Telecomunicación de la Universidad Politécnica de Madrid; dentro de un programa de becas en áreas de interés para la empresa y que complementan el proceso formativo de los estudiantes de la escuela.

Securitas Direct es una de las mayores empresas dedicadas a la seguridad, registrada y autorizada para la instalación y mantenimiento de dispositivos y sistemas de alarma de última generación. Sus dispositivos utilizan las bandas de frecuencias radio industriales, científicas y médicas (ISM, *Industrial, Scientific, and Medical*), para su funcionamiento. Específicamente, trabajan en la frecuencia de 868 MHz, que se encuentra dentro del espectro asignado para comunicaciones inalámbricas de alto alcance en el ámbito de la seguridad. Gracias a su enfoque en la tecnología inalámbrica y su dedicación a la seguridad, Securitas Direct se posiciona como una empresa líder en el mercado.

La detección de personas basada en los cambios de las señales inalámbricas es importante por las siguientes razones:

- Mejora la seguridad en entornos residenciales ya que permite detectar a personas sin necesidad de una interacción directa ni de llevar dispositivos adicionales.
- Tiene un gran alcance en espacios interiores, lo que permite la detección en grandes áreas como hogares u oficinas.
- Ofrece diversas posibilidades para la detección, lo que permite adaptarse a diferentes escenarios y necesidades específicas.

1.2. Objetivo

El objetivo de este Trabajo Fin de Grado es realizar un análisis de las distintas posibilidades existentes para el cálculo de presencia y movimiento basado en cambios en la señal de radiofrecuencia. Específicamente, el enfoque del proyecto se centrará en el análisis de señales a una frecuencia de 868 MHz. Este proyecto, se va a llevar a cabo en la banda de 868 MHz por las siguientes razones:

- La mayoría de los dispositivos de Securitas Direct Verisure trabajan en la banda de 868 MHz. Esta elección de frecuencia facilita la comunicación con los demás dispositivos, asegurando una integración fluida y confiable.
- Se logra un mayor alcance, ya que los dispositivos que trabajan en la banda de 2,4 GHz tienen una potencia de radiofrecuencia regulada a un máximo de 100 mW. En cambio, la potencia de los dispositivos que trabajan en la banda de 868 MHz es de 500 mW, lo que significa que estos últimos tienen una distancia de transmisión teórica cinco veces mayor que la que podría alcanzar un dispositivo que opera en la banda de 2,4 GHz [1].
- Se producen menos interferencias debido a que la banda de 2,4 GHz es una banda muy solicitada y congestionada al ser utilizada por muchos dispositivos, como por ejemplo routers, móviles o portátiles.
- La comunicación en la banda de frecuencias de 868 MHz consume menos energía en comparación con las comunicaciones en la banda de 2,4 GHz.

Para llevar a cabo este trabajo, se realizará un estudio de las técnicas más importantes de detección de presencia según la variación de la señal de radiofrecuencia. Además, se investigarán los diferentes algoritmos de detección basados en *machine learning*, los cuales nos permiten extraer características y patrones de conjuntos de datos, con el propósito de poder inferir resultados.

1.3. Metodología

La metodología empleada en este Trabajo de Fin de Grado se basó en la combinación de *software* y herramientas Jira para la gestión y organización de tareas e incidencias, junto a una metodología SCRUM. Con esta metodología, se dividió el trabajo y en *sprints* o periodos de tiempos cortos y fijos, de dos semanas. Además, se llevaron reuniones de seguimiento al finalizar cada *sprint* para evaluar las tareas, revisar los resultados obtenidos y planificar los próximos objetivos. Esta metodología proporcionó una gestión ágil y eficiente del proyecto.

1.4. Tareas

Las tareas a realizar en este proyecto son las siguientes:

1. Estudiar el estado del arte en cuanto a las principales técnicas de detección de presencia aplicadas a la explotación de las variaciones de la señal de radiofrecuencia.
2. Simular una transmisión continua de paquetes entre dos puntos.
3. Analizar las variaciones en la banda de radio para la obtención y explotación de datos.
4. Estudiar los diferentes algoritmos basados en *machine learning*.
5. Implementar uno de los algoritmos estudiados previamente para la detección de presencia.
6. Realizar diferentes pruebas para evaluar el desempeño y precisión del sistema.
7. Analizar los resultados y elaborar las conclusiones del proyecto.

1.5. Estructura del contenido

Este documento se ha estructurado siguiendo los siguientes capítulos:

- Capítulo 1: En el primer capítulo se presenta el proyecto, el entorno y motivación, sus objetivos y las tareas que se llevarán a cabo.
- Capítulo 2: En el segundo capítulo se presenta una revisión detallada de las técnicas de detección las señales inalámbricas.
- Capítulo 3: En el tercer capítulo se describe la simulación de un enlace de dos puntos utilizando dos placas de Silicon Labs.
- Capítulo 4: En el cuarto capítulo se realiza un análisis de los diferentes algoritmos de decisión no supervisados, centrándose especialmente en el algoritmo *k-means*.
- Capítulo 5: En el quinto capítulo se presentan las pruebas y resultados del proyecto.
- Capítulo 6: En el último capítulo se realizan las conclusiones del proyecto.
- Anexo A: En el primer anexo se estudian los aspecto éticos, económicos, sociales y ambientales del proyecto.
- Anexo B: En el segundo anexo se estudia el aspecto económico del proyecto.
- Anexo C: En el tercer anexo se analizan los parámetros configurables del algoritmo *k-means* estudiado en el capítulo 4.

Capítulo 2

Estado del arte

En este capítulo se va a estudiar el desarrollo en los últimos años de la detección de presencia y movimientos humanos a través de señales inalámbricas y las diferentes técnicas que existen para ello en la actualidad.

2.1. Detección de presencia basada en señales inalámbricas

Hoy en día, la detección de presencia humana se ha convertido en un problema desafiante en áreas como la vida asistida y el monitoreo remoto. Además de las metodologías de detección convencionales, basadas en sensores y sistemas de cámaras, las tecnologías basadas en señales inalámbricas están evolucionando recientemente como una solución prometedora para el monitoreo de interiores y el reconocimiento de actividades humanas.

La primera detección de humanos basada en WiFi se remonta a 2007, Youssef et al. [2], usó el indicador de intensidad de la señal recibida (RSSI) como huella digital para inferir la posición de una persona, lo que llamó "localización pasiva sin dispositivo". Los movimientos humanos pueden tener diferentes efectos en las señales recibidas. Por lo tanto, el principio básico de la detección de humanos con WiFi es establecer un mapeo entre las características de las señales recibidas y los movimientos humanos. Sin embargo, los canales del mundo real están lejos de ser limpios, ya que los factores ambientales pueden estar involucrados en el mapeo establecido entre las señales recibidas y los movimientos humanos, lo que limita la comercialización de la detección de presencia humana con WiFi.

Con el avance de las tecnologías inalámbricas y las metodologías de detección, muchos estudios [3] han demostrado el buen funcionamiento de las señales inalámbricas, como el WiFi, para detectar actividades humanas. Esto ha permitido el desarrollo de una serie de aplicaciones, como por ejemplo, la detección de intrusos, el reconocimiento de actividad diaria y la identificación del usuario, que involucra una detección de movimiento aún más detallada. Los movimientos del cuerpo humano afectan la propagación de la señal inalámbrica, lo que proporciona amplias posibilidades para capturar los movimientos humanos mediante el análisis de las señales inalámbricas

recibidas.

Las investigaciones pioneras [4] han explorado la posibilidad de extraer información de movimiento de señales inalámbricas para localizar o rastrear movimientos de todo el cuerpo o incluso gestos, de manera no invasiva. El término "no invasivo", también conocido como pasivo o sin dispositivos, significa que no se supone que las personas lleven ningún dispositivo inalámbrico. La base principal es que el movimiento de una persona puede modular las señales inalámbricas y generar cambios temporales en las señales enviadas, los cuales se pueden observar a partir de las señales recibidas.

Para realizar de manera efectiva el reconocimiento de la actividad humana, varias tecnologías de detección incluyendo sensores de varios tipos se despliegan para inspeccionar diferentes actividades y gestos humanos. Sin embargo, estas tecnologías se enfrentan a diversas dificultades [3]:

- Los enfoques que se basan en las cámaras o sensores de luz solo demuestran un buen funcionamiento bajo ciertas condiciones de iluminación; por lo que podrían verse interferidos por condiciones de poca luz, niebla o humo.
- La estabilidad de los sensores acústicos es altamente sensible al ruido ambiental y a las interferencias de sonido del entorno. Además, su rango de detección está limitado debido a la rápida atenuación de las señales acústicas.

En resumen, todas las técnicas mencionadas anteriormente implican gastos adicionales en cuanto a la instalación *hardware* y necesidades de mantenimiento. Para poder superar las limitaciones, es preferible una solución de bajo coste y no intrusiva para captar los movimientos de los cuerpos humanos. Recientemente, muchos trabajos de investigación utilizan técnicas basadas en radiofrecuencia para la detección de la actividad humana. Un ejemplo de esto es el gran uso de la tecnología WiFi, que conecta prácticamente todos los dispositivos electrónicos en entornos residenciales u oficinas. Asimismo, las señales WiFi tienen un alcance de varias decenas de metros en espacios interiores, permitiendo que estos dispositivos inteligentes se comuniquen entre sí mediante una amplia red de señales reflejadas que se expanden por todos los rincones internos.

Para cuantificar los cambios de la señal recibida, los investigadores [3] podrían medir las propiedades de la capa física sobre el canal inalámbrico, como el indicador de intensidad de la señal recibida, en inglés *Received Signal Strength Indicator* (RSSI) y la información del estado del canal, en inglés *Channel State Information* (CSI). Asimismo, para lograr una detección más precisa, algunos investigadores manipulan la transmisión de señales inalámbricas en una plataforma de radio definida por USRP (periférico de radio de software universal), como la onda portadora modulada en frecuencia (FMCW), para detectar el cambio de frecuencia. Por otro lado, también se puede aprovechar el efecto Doppler para medir el cambio de frecuencia de la señal asociado con los movimientos del cuerpo. Esta técnica también necesita el apoyo de plataformas USRP para controlar la transmisión y recepción de señales inalámbricas.

En los siguientes apartados, vamos a analizar las distintas técnicas utilizadas para la detección de presencia humana a través del cambio de las señales inalámbricas, que van desde la medición de las propiedades de la capa física, hasta el análisis de señales y el uso de herramientas especializadas.

2.2. Intensidad de la señal recibida (RSSI)

La intensidad de señal recibida (RSSI), es una escala de referencia, en relación con 1 mW, utilizada para medir el nivel de potencia de las señales recibidas por un dispositivo en las redes inalámbricas. La escala tiende al valor cero como centro y se representa como *0RSSI* o *0dBm*, y aunque pueden darse valores positivos, generalmente la escala se expresa dentro de los valores negativos. Cuanto más negativos sean los valores, mayor es la pérdida de la potencia de la señal. Es importante destacar que el RSSI indica la intensidad de la señal recibida, no la calidad de la señal, que se determina analizando la intensidad de la señal respecto a la relación señal a ruido (SNR).

La señal recibida está disponible en la mayoría de los dispositivos inalámbricos, la cual indica las pérdidas de la señal inalámbrica respecto a una distancia, y se puede determinar por la ecuación 2.1.

$$P(d) = P(d_0) + 10\gamma \log_{10} \left(\frac{d}{d_0} \right) + X_\delta \quad (2.1)$$

siendo $P(d)$ las pérdidas de la trayectoria, medida en decibelios, a una distancia d , $P(d_0)$ las pérdidas de la trayectoria a una distancia de referencia d_0 , γ el exponente de pérdidas de la trayectoria, y X_δ un ruido de media cero causado por un desvanecimiento plano.

Los métodos de detección basados en RSSI infieren principalmente a los intrusos a través de la detección de perturbaciones humanas en las mediciones de RSSI en las redes inalámbricas. Cuando un individuo entra en un área, los enlaces inalámbricos se interrumpen debido a la presencia o los movimientos del cuerpo del intruso, lo que se convierte en un cambio de RSSI en las señales de radio.

Durante mucho tiempo, una de las aplicaciones más representativas del uso de RSSI ha sido para calcular la posición de usuarios objetivo con dispositivos WiFi [3]. Se ha observado que la presencia de un cuerpo humano dentro del área de detección inalámbrica puede causar atenuación de la señal, lo que resulta en variaciones en las mediciones de RSSI. Por lo tanto, en los últimos años, el uso de RSSI se ha visto ampliamente implementado para la detección de actividades humanas en diferentes aplicaciones. Por ejemplo, se ha empleado para la localización en interiores sin otros dispositivos adicionales, la estimación de la densidad de personas en una sala y el seguimiento de la frecuencia respiratoria.

Como se puede observar en la figura 2.1 [3], para facilitar el control de la ocupación de las habitaciones, los investigadores aseguran que [5]:

- Cuando no hay ningún sujeto en el área de interés, los valores de RSSI se mantienen en un nivel estable.
- Cuando algunos sujetos ingresan a la zona de detección, la lectura de RSSI de algunos enlaces de radiofrecuencia disminuiría drásticamente.
- Cuanto mayor es el número de sujetos, más se ven afectados los enlaces de radio, lo que resulta en caídas significativas en las lecturas de RSSI.

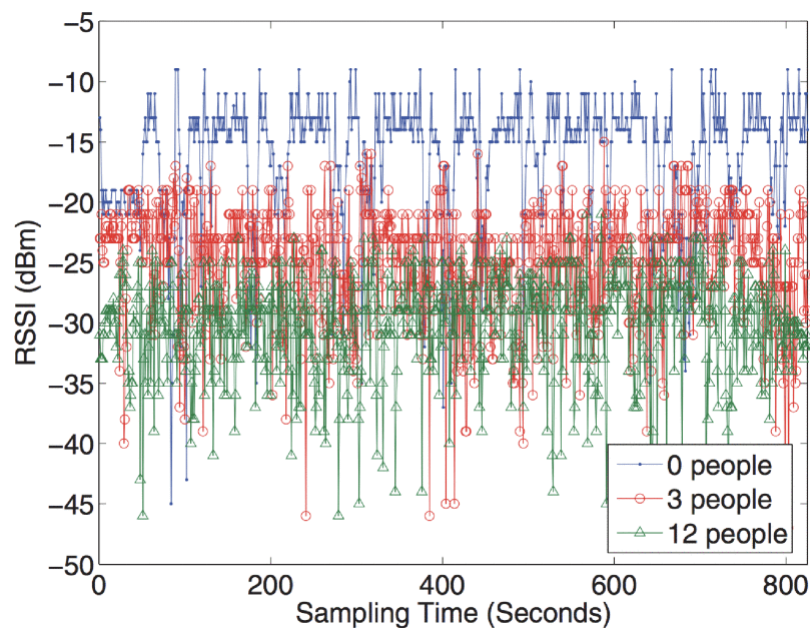


Figura 2.1: Ejemplo de medición RSSI para 0, 3 y 12 personas

Sin embargo, a pesar de la fácil obtención del RSSI en dispositivos inalámbricos sin un hardware adicional, esta técnica solo puede detectar ciertas actividades humanas debido a la información de estado de canal de granularidad gruesa, es decir, el valor de pérdida de ruta única por paquete. Además, se ha demostrado que la estabilidad del RSSI no está garantizada ni siquiera en entornos interiores estáticos, ya que sus resultados dependen de la situación y del entorno; por lo tanto, no es del todo fiable en muchos escenarios.

2.3. Información del estado del canal (CSI)

Otra técnica de gran relevancia es la información del estado del canal (CSI). El CSI es el análisis de las propiedades del canal de comunicación. Esta información describe como una señal se propaga desde el transmisor hasta el receptor y presenta el efecto combinado del espacimiento, la atenuación y la disminución de potencia respecto a la distancia. El CSI permite adaptar las transmisiones a las condiciones actuales del canal, lo cual es crucial para lograr una comunicación fiable con altas tasas de datos. El CSI debe ser estimado en el receptor y generalmente se cuantifica y se retroalimenta al transmisor. Por lo tanto, el transmisor y el receptor pueden tener un CSI diferente.

El CSI de los dispositivos inalámbricos ha sido explorado para las tareas de detección de presencia humana como el reconocimiento de la actividad, gestos, monitorización de interiores y aplicaciones para el cuidado de la salud. Estos avances se benefician de las propiedades de los dispositivos inalámbricos comerciales, como su robustez a las condiciones de iluminación y la detección no intrusiva que no requiere un esfuerzo adicional por parte del usuario [6].

Para lograr una detección precisa y confiable de la actividad humana, es esencial capturar un CSI detallado. Como las señales pueden viajar por cualquier sitio en un entorno interior, la presencia o el movimiento humano altera la propagación de señales inalámbricas. Este fenómeno da lugar a cambios en los rayos reflejados, como se puede apreciar en la figura 2.2. Todos estos rayos de trayectos múltiples contribuyen a los valores de CSI y podrían usarse para detectar y rastrear los movimientos del cuerpo humano.

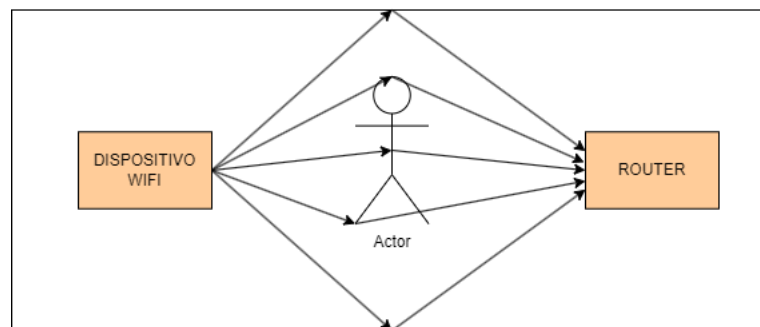


Figura 2.2: Multitrayecto de las señales inalámbricas (elaboración propia)

Al igual que en RSSI, las medidas CSI se pueden obtener en cualquier dispositivo disponible en el mercado con un controlador modificado. En cambio, CSI consta de un conjunto de valores complejos, con amplitud y fase para múltiples subportadoras de multiplexación por división de frecuencia ortogonal (OFDM). Cada subportadora cuenta con una frecuencia central ligeramente diferente que experimenta efectos de desvanecimiento de múltiples rutas, y todas las subportadoras juntas representan el

canal inalámbrico de manera detallada [3]. Esencialmente, CSI permite la estimación de canal de grano fino y se expresa como:

$$H = [H_1, H_2, \dots, H_i, \dots, H_N]^T \quad i \in [1, N] \quad (2.2)$$

donde N indica el número de subportadoras y H_i se puede ver representado según la siguiente ecuación:

$$H_i = |H_i| \cdot e^{j\text{sen}(\angle H_i)} \quad (2.3)$$

donde H_i es la amplitud de CSI en la subportadora i ésima, y $\angle H_i$ denota la fase.

Finalmente, la variación de las mediciones de CSI también se puede extraer para inferir la cantidad de personas que caminan en un ambiente interior. CSI proporciona información de canal más detallada, información de amplitud y fase, con múltiples subportadoras. En la gráfica 2.3 [3] se muestra el impacto de diferentes números de sujetos en las diferencias de amplitud de CSI en 30 subportadoras.

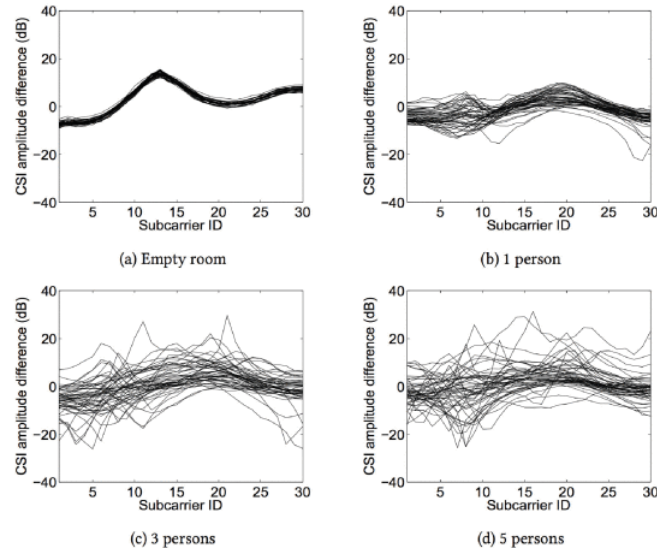


Figura 2.3: Ejemplo de medición CSI para 0, 1, 3 y 5 personas

En resumen, la técnica de detección CSI ofrece varias ventajas como un alto nivel de precisión en la detección del movimiento humano, como la respiración o gestos específicos. Además, en comparación con RSSI, CSI es menos sensible al ruido. Sin embargo, también presenta algunos inconvenientes, como la necesidad de ajustes o calibraciones frecuentemente para garantizar un rendimiento óptimo. Asimismo, la interpretación de la información CSI es más compleja en comparación con RSSI.

2.4. Frecuencia modulada de onda continua (FMCW)

Las actividades humanas también se pueden capturar en función de la reflexión de las señales inalámbricas con el cuerpo humano. En concreto, al estimar el tiempo que tarda una señal en viajar desde el transmisor hasta el cuerpo humano y posteriormente, al receptor. Sin embargo, calcular el tiempo de vuelo directamente es muy difícil ya que las señales se mueven a la velocidad de la luz. Por ello, la frecuencia modulada de onda continua, en inglés *Frequency Modulated Continuous Wave* (FMCW), que asigna diferencias de tiempo a los cambios de frecuencia de la portadora, se puede implementar para medir el tiempo de vuelo.

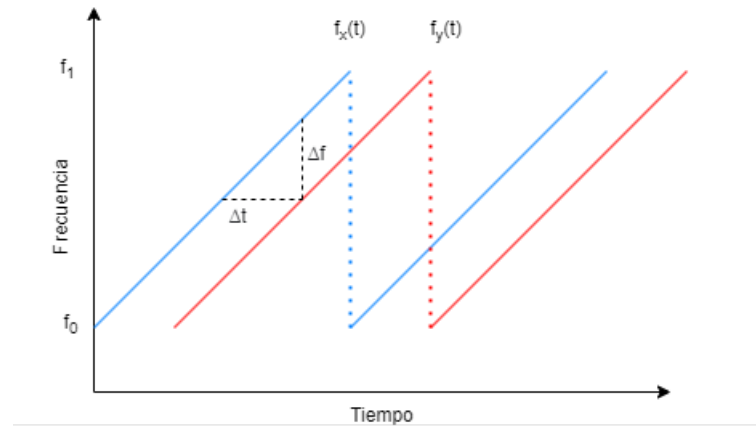


Figura 2.4: Desplazamiento FMCW (elaboración propia)

Como podemos ver en la figura 2.4, la frecuencia de portadora de la señal $f_x(t)$ se barre a través de un ancho de banda específico. Una vez la onda ha sido reflejada en el cuerpo humano, se produce un cambio de frecuencia Δf con una pendiente k en la señal recibida $f_y(t)$, y la variación en el tiempo Δt , respecto a la señal transmitida se puede derivar en función de dicho cambio de frecuencia de la siguiente manera:

$$\Delta t = \frac{\Delta f}{k} \quad (2.4)$$

Se han desarrollado varios sistemas de detección inalámbricos que aprovechan la técnica FMCW para rastrear diferentes actividades humanas. Por ejemplo, los investigadores utilizan señales FMCW generadas por el software de radio USRP con antenas direccionales para capturar la figura humana a través de una pared, calcular la velocidad de movimiento o detectar signos vitales [7].

En la figura 2.5 [8], se puede ver un ejemplo de desplazamiento FMCW a 80 GHz.

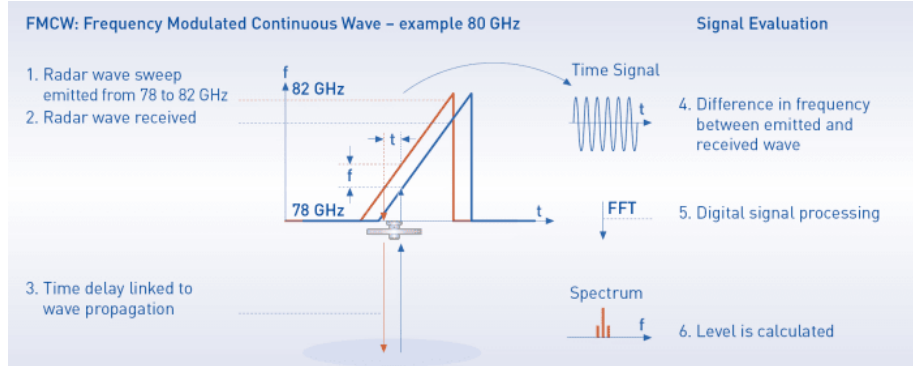


Figura 2.5: Ejemplo desplazamiento FMCW a 80 GHz

En resumen, la técnica de detección FMCW ofrece varias ventajas como su capacidad para obtener mediciones precisas de distancia y velocidad, su inmunidad al ruido y su amplio rango de frecuencias para detectar a diferentes distancias. Sin embargo, también presenta algunos inconvenientes, como la necesidad de un *hardware* específico, por lo tanto, requiere mayor complejidad y un coste elevado, así como el requerimiento de un ancho de banda muy grande para transmitir y recibir las señales.

2.5. Desplazamiento Doppler

El efecto de desplazamiento Doppler es otra propiedad de las señales inalámbricas que se puede emplear para la detección de presencia y movimiento humano. En concreto, el efecto Doppler mide el cambio de frecuencia de la señal recibida a medida que el emisor y receptor se mueven simultáneamente. Como la señal inalámbrica se ve reflejada por el cuerpo humano, cualquier movimiento de este se vería reflejado como un desplazamiento Doppler. Es decir, se produce un cambio de frecuencia positivo si la persona se acerca al receptor, y negativo si la persona se aleja del receptor.

Como se muestra en la figura 2.6, cuando una persona se mueve a una velocidad v a lo largo de una dirección θ con respecto al receptor, provoca un desplazamiento Doppler que corresponde a la ecuación 2.5.

$$\Delta f = \frac{2v \cos(\theta)}{c} \cdot f \quad (2.5)$$

donde c es la velocidad de la luz y f es la frecuencia central de la señal inalámbrica.

Haciendo uso de los efectos de desplazamiento Doppler, se han desarrollado algunos sistemas de detección inalámbrica basados en un dispositivo radio definido por software, como por ejemplo el USRP N210, para detectar el cuerpo o gestos humanos.

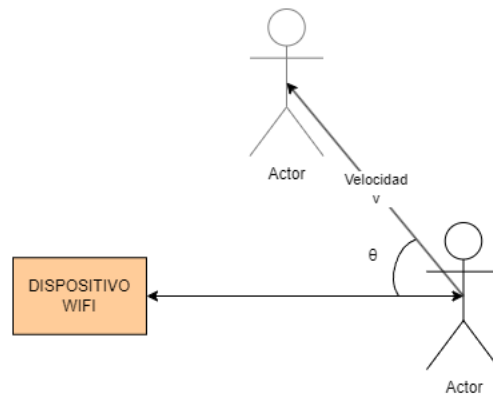


Figura 2.6: Ejemplo de cambio de frecuencia Doppler (elaboración propia)

2.6. Conclusión del estado del arte

En el estado del arte se ha llevado a cabo un estudio a fondo de las técnicas de detección más relevantes en el campo de las señales inalámbricas. Este capítulo ha permitido analizar y evaluar las características, fortalezas y limitaciones de cada una de las técnicas. Además, ha proporcionado una visión de todas las opciones disponibles para llevar a cabo la detección de presencia y movimientos basados en los cambios de la señal de radiofrecuencia.

De acuerdo con las técnicas de detección introducidas en este capítulo, clasificamos el campo en cuatro categorías principales: basada en RSSI, basada en CSI, basada en FMCW y basada en desplazamiento Doppler. Estos estudios de detección inalámbrica han mostrado un futuro prometedor en varios dominios de aplicación.

La intensidad de la señal recibida (RSSI) es una medida utilizada para evaluar el nivel de potencia de las señales recibidas por un dispositivo. Los métodos principales basados en esta técnica se centran en detectar perturbaciones en las mediciones de RSSI causadas por la presencia o los movimientos de personas en las señales inalámbricas.

La información del estado del canal (CSI) en las señales inalámbricas nos proporciona información acerca de cómo las señales inalámbricas se propagan desde el transmisor hasta el receptor. La presencia o el movimiento humano afecta a la propagación de las señales inalámbricas, lo que resulta en cambios en los rayos reflejados. Y estos cambios, se aprecian en los valores de CSI, que contienen información de amplitud y fase de sus distintas subportadoras.

La técnica FMCW se emplea para capturar actividades humanas en función de la reflexión de las señales inalámbricas con el cuerpo humano. Esto se logra estudiando el tiempo que tarda una señal en viajar desde el transmisor hasta el cuerpo humano y posteriormente, al receptor. Esta técnica ha sido aplicada en sistemas de detección inalámbricos para localizar y monitorizar diversas actividades humanas.

Por último, el desplazamiento Doppler es una de las propiedades de las señales inalámbricas que puede ser utilizadas para la detección de presencia y movimiento humano. Este efecto mide el cambio de frecuencia de la señal recibida cuando el emisor y el receptor se mueven de manera simultánea. Por lo tanto, cualquier movimiento de este tipo se vería reflejado como un desplazamiento Doppler.

Este estudio ha sido fundamental ya que ha proporcionado una base para la toma de decisiones futuras acerca de la selección de la técnica más adecuada para un proyecto o una aplicación específica. Cabe destacar que este estudio se ha realizado en la banda de frecuencia de 2,4 GHz, la cual es ampliamente utilizada en aplicaciones inalámbricas y esta extensamente estudiada.

Sin embargo, para este proyecto se ha decidido emplear una frecuencia de 868 MHz por las siguientes razones:

- La mayoría de los dispositivos de Securitas Direct Verisure trabajan en la banda de 868 MHz. Esta elección de frecuencia facilita la comunicación con los demás dispositivos, asegurando una integración fluida y confiable.
- Se logra un mayor alcance, ya que los dispositivos que trabajan en la banda de 2,4 GHz tienen una potencia de radiofrecuencia regulada a un máximo de 100 mW. En cambio, la potencia de los dispositivos que trabajan en la banda de 868 MHz es de 500 mW, lo que significa que estos últimos tienen una distancia de transmisión teórica cinco veces mayor que la que podría alcanzar un dispositivo que opera en la banda de 2,4 GHz [1].
- Se producen menos interferencias debido a que la banda de 2,4 GHz es una banda muy solicitada y congestionada al ser utilizada por muchos dispositivos, como por ejemplo routers, móviles o portátiles.
- La comunicación en la banda de frecuencias de 868 MHz consume menos energía en comparación con las comunicaciones en la banda de 2,4 GHz.

Como conclusión, el estudio de las técnicas más relevantes en el ámbito de las señales inalámbricas ha permitido obtener una visión completa de todas las opciones disponibles y sus características. La elección de la frecuencia de 868 MHz para este trabajo se basa en las condiciones revisadas anteriormente. Asimismo, se busca emplear la técnica más adecuada para lograr el máximo rendimiento en cuanto a la detección de presencia y movimiento basados en los cambios de la señal de radiofrecuencia.

Capítulo 3

Desarrollo de escenario de pruebas

El objetivo de este Trabajo Fin de Grado es realizar un análisis de las distintas posibilidades existentes para el cálculo de presencia y movimiento, basado en cambios en la señal de radiofrecuencia. Para ello, en este capítulo, se va a simular un enlace entre dos puntos para observar cómo varía la señal en diferentes escenarios y condiciones. La elección de realizar una simulación se debe a su capacidad para hacer pruebas en múltiples escenarios sin asumir costes adicionales de *hardware*. Además, se obtienen beneficios en términos de flexibilidad, ya que se pueden ajustar parámetros y probar diferentes configuraciones, lo que permitirá llevar a cabo una evaluación exhaustiva del proyecto.

Asimismo, se pretende integrar este proyecto en una unidad central de la empresa Securitas Direct, mostrada en la figura 3.1a [9]. Cuando un detector de acceso, representado en la figura 3.1b [9], ubicado en una puerta o ventana, detecte que se está intentando forzar, establecerá una comunicación continua con la unidad central, creando así el enlace simulado en este capítulo. Mediante esta conexión, será posible verificar si el intruso ha logrado entrar a la vivienda, proporcionando una capa adicional de seguridad.

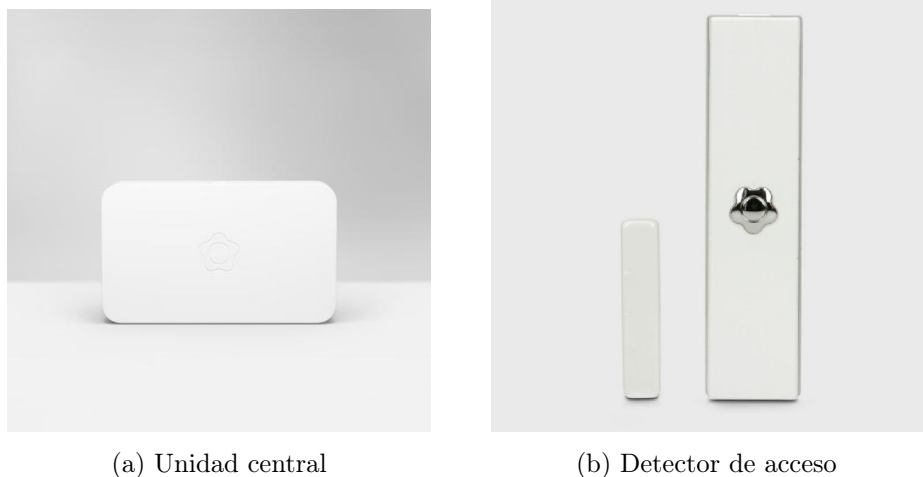


Figura 3.1: Dispositivos de Seguridad Direct

En el capítulo 2, se ha llevado a cabo un estudio de las diferentes técnicas de detección más relevantes en el campo de las señales inalámbricas. Estas técnicas representan los métodos más utilizados para el cálculo de presencia y movimiento humano basado en los cambios de la señal de radiofrecuencia. Se han analizado y evaluado las técnicas: RSSI, CSI, FMCW y Doppler. Una vez realizado este estudio, analizando las diferentes características y aplicaciones, se han llegado a las siguientes conclusiones:

- La primera técnica estudiada para la detección de personas fue RSSI. Es una técnica ampliamente estudiada tanto en frecuencias 2,4 GHz como en 868 MHz y se adapta a los requerimientos específicos que se buscan alcanzar en este proyecto.
- En cuanto CSI, cabe mencionar que es una técnica muy poco estudiada en el contexto de la banda de 868 MHz. Debido a las limitaciones de tiempo para desarrollar este proyecto, se ha decidido desestimar esta opción, ya que no se ha explorado lo suficiente y su desarrollo e implementación requeriría más esfuerzo de investigación adicional.
- Respecto a la técnica FMCW, es importante destacar que es una técnica que utiliza un *hardware* específico en la unidad central, el cual no disponíamos para realizar este proyecto. En consecuencia, se ha tomado la decisión de descartar esta opción debido a la falta de recursos para su implementación.
- Con relación a la técnica *Doppler*, se llevó a cabo un estudio en profundidad durante el desarrollo de este trabajo. Sin embargo, no se lograron obtener los resultados ni el rendimiento deseado para la aplicación en este proyecto. Por lo tanto, se decidió desestimar esta técnica.

Dado todo esto, se ha determinado que la opción más adecuada para llevar a cabo la simulación es utilizar la técnica de RSSI o intensidad de la señal recibida.

Además, se utilizaron dos placas *EFR32FG* de Silicon Labs junto con un *Flex Gecko* capaz de trabajar a una frecuencia de 868 MHz (ver figura 3.2 [10]). Estas placas ofrecen un conjunto de características y funcionalidades óptimas para el proyecto. Asimismo, el rango de frecuencias de las placas se adapta a la perfección al que se va a desarrollar en este trabajo. Una de las placas actuará como transmisor y la otra como receptor. Para implementar la comunicación continua de paquetes, se estudió y ejecutó en ambas placas un programa llamado *RAIL-SoC range_test*.

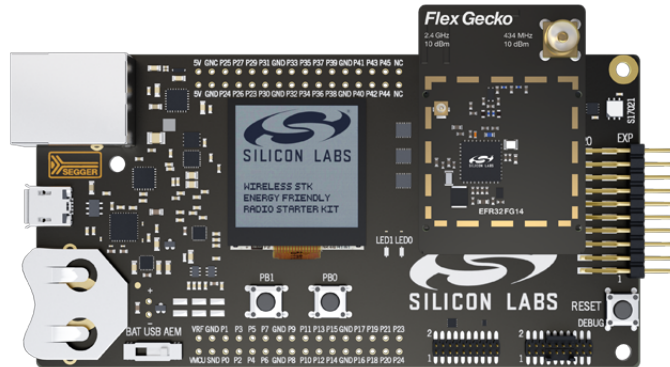


Figura 3.2: Placa *EFR32FG* de Silicon Labs

3.1. *RAIL-SoC range_test*

Este programa ofrece una manera sencilla de evaluar el enlace de los dispositivos inalámbricos *Gecko EFR32* utilizando la capa de interfaz de abstracción de radio (*RAIL*) de Silicon Labs (ver figura 3.3 [11]).

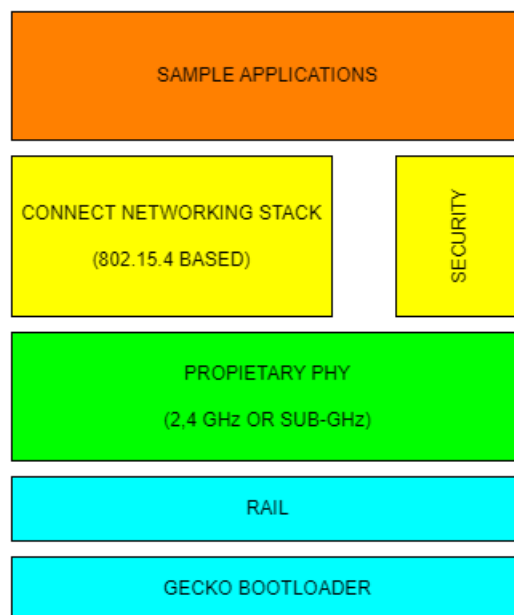
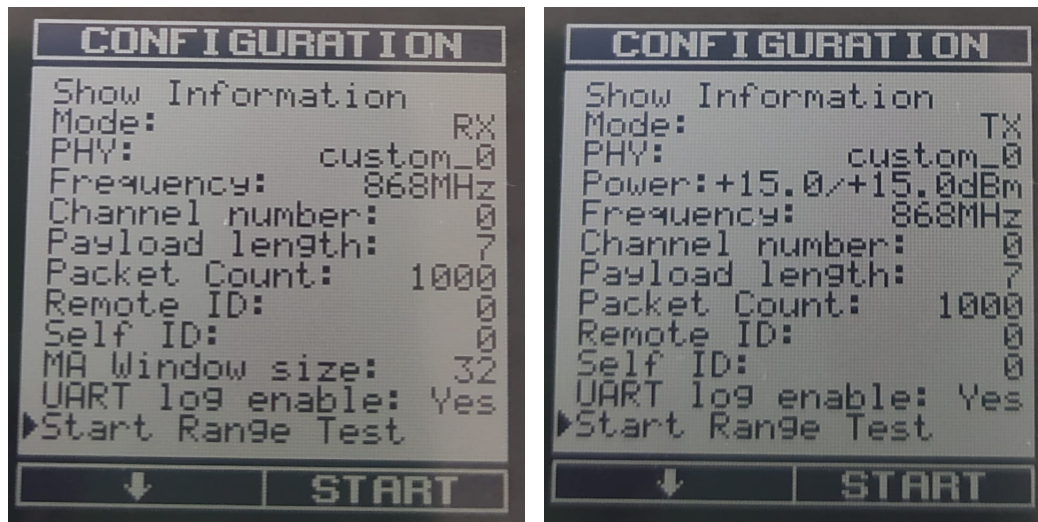


Figura 3.3: Arquitectura *software* de los nodos (elaboración propia)

En este programa se realiza una prueba de rango que crea un enlace de radio entre dos kits de evaluación, uno actuando como transmisor y otro como receptor, y envía una cantidad predefinida de paquetes. La demostración de *Range Test* [12] implementa la medición de la tasa de error de paquetes, en inglés *Package Error Ratio* (PER), siguiendo la ecuación 3.1. La tasa de error de paquetes es una técnica ampliamente utilizada para evaluar la calidad de los enlaces de radiofrecuencia en sistemas inalámbricos bajo condiciones particulares. Este programa, proporciona resultados de medición con respecto a la calidad del enlace de radiofrecuencia.

$$PacketErrorRate(PER)[\%] = \frac{P_{Tx} - P_{Rx}}{P_{Tx}} \cdot 100 \quad (3.1)$$

Al ejecutar el programa en ambos dispositivos, el receptor calcula la tasa de error de paquetes (PER) y muestra el nivel actual de RSSI en unidades *dBm*, además de dibujar un gráfico de los datos históricos de RSSI. En la pantalla de las placas, se mostrará un menú interactivo, reflejado en la figura 3.4, donde se muestran los diferentes parámetros y funciones disponibles. Para navegar por el menú, se utilizan los dos botones de la placa. El botón 1 (*BTN1*) permite desplazarse hacia abajo a través de los elementos del menú; mientras que el botón 0 (*BTN0*) se utiliza para configurar el elemento de menú seleccionado mediante el puntero. Estos elementos permiten ajustar diferentes parámetros según las necesidades del usuario, brindando flexibilidad y personalización en la configuración del programa.



(a) Menú de la placa receptora

(b) Menú de la placa transmisora

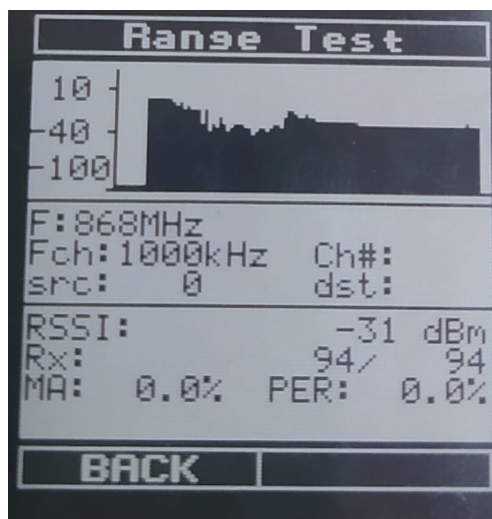
Figura 3.4: Configuración de las placas de Silicon Labs (elaboración propia)

Las diferentes opciones disponibles en el menú son [13]:

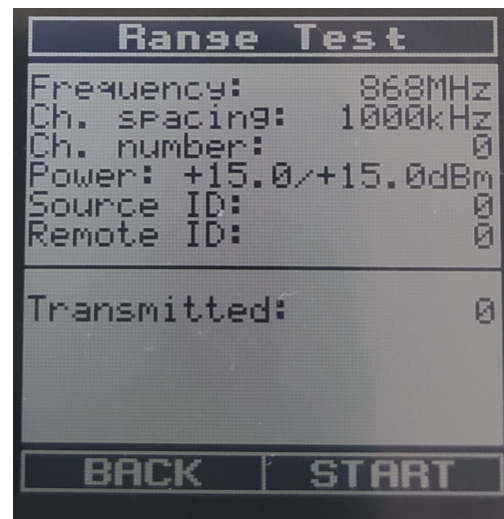
- *Mode* o modo (RX o TX): las placas se inician en modo RX de forma predeterminada. Además, el mismo dispositivo puede actuar como transmisor y luego como receptor sin necesidad de reiniciarlo.
- *PHY*: esta capa solo es configurable en una aplicación PHY estándar, como BLE 1/2 Mbps o 802.15.4. En el caso de una aplicación PHY personalizada con una configuración de radio Multi-PHY, también se puede ajustar la capa PHY de acuerdo con los requerimientos específicos de dicha aplicación.
- *Frecuency* o frecuencia: no es un parámetro configurable, pero muestra la frecuencia del canal base del PHY seleccionado. Por lo tanto, no se actualiza cuando se aumenta el número de canal.
- *Channel number* o número de canal: selecciona el canal en el que el nodo de recepción escuchará o el nodo de transmisión emitirá.
- *Payload length* o longitud de carga útil: configurable entre 7 y 64 en una aplicación PHY personalizada. En una aplicación PHY estándar se puede configurar para 5 y 24 con BLE PHY o entre 5 y 116 con 802.15.4 PHY.
- *Packet count* o cuenta de paquetes: establece el número de paquetes transmitidos o esperados por las placas. Puede variar entre 500, 1000, 2500 o infinito tanto en recepción como en transmisión.
- *Remote/Self ID* o identificación remota/propia: en modo de recepción, la radio escucha en el canal dado e inspecciona los paquetes recibidos. Solo se procesarán los paquetes que se envíen con *Self ID* y *Remote ID* coincidentes. Por lo tanto, los identificadores deben configurarse en el transmisor.
- *MA window size* o tamaño de ventana MA (solo en RX): es el tamaño de ventana en un promedio móvil (MA). Configura cuántos paquetes se contarán en una medición PER más corta durante la prueba. El promedio móvil PER es útil cuando se trata de posicionar los dispositivos hasta obtener un PER específico.
- *Power* o potencia (solo en TX): la potencia de salida se puede configurar en el menú en pasos de 0,5 dBm, que es el punto de ajuste de potencia, entre -15 y +20 dBm.
- *UART log enable* o registro de UART: si está activado, se puede observar un mensaje de estado en la línea recepción y transmisión de UART para cada paquete de radio, formateado en un formato legible por humanos. La asignación de pines predeterminada es el puerto VCOM estándar disponible en la placa base.

Una vez estudiadas las diferentes opciones de configuración, para llevar a cabo la simulación, se deben realizar los siguientes pasos:

1. Conectar las placas de radio, *Flex Gecko*, a las dos placas de Silicon Labs.
2. Conectar ambas placas a un ordenador y compilar el proyecto, programándolo en los dispositivos.
3. Configurar los dispositivos, uno para actuar como receptor y el otro como transmisor. Además, se debe ajustar la configuración para poder trabajar a una frecuencia de 868 MHz y habilitar la comunicación a través del puerto UART. Y por último, determinar el número de paquetes que se quieren enviar.
4. Mantener la placa receptora conectada al ordenador para poder obtener los datos, mientras que la placa transmisora puede ser colocada en cualquier ubicación dentro de la habitación.
5. Ejecutar los scripts en el ordenador para obtener la información de las placas a través del puerto UART.
6. Comenzar la ejecución del programa presionando el botón start (ver figura 3.5).



(a) Placa receptora configurada



(b) Placa transmisora configurada

Figura 3.5: Placas de Silicon Labs configuradas y preparadas para comenzar (elaboración propia)

Por último, para facilitar la recopilación y evaluación de los datos del programa, se han creado dos *scripts* en *Python*. El primero de ellos, se encarga de obtener los datos transmitidos por la UART y almacenarlos en un archivo de texto. Este script establece una comunicación con el programa a través de la UART, recibiendo los datos en tiempo real y registrándolos. Esto permite una fácil recopilación de los datos generados durante la ejecución del programa para su posterior análisis.

El segundo *script* se encarga de tomar el archivo de texto generado por el primer *script* y convertirlo en un archivo Excel. Utilizando las bibliotecas especializadas, este *script* lee el archivo de texto línea por línea, extrae toda la información y la almacena en un archivo tipo Excel. De esta manera, se logra una transición de los datos capturados en la UART a un formato Excel, para poder estudiar los datos más adelante.

Tanto el programa ejecutado en las placas como los scripts utilizados en la recopilación de datos se encuentran disponibles en el repositorio del siguiente enlace. Este repositorio proporciona acceso a todo el código necesario para realizar una reproducción exacta de los resultados presentados en la memoria.

https://github.com/alvarogom3z/868_sensing.git

En resumen, este capítulo se ha enfocado en la simulación de un enlace entre dos puntos para analizar los cambios en la señal de radiofrecuencia. Para la detección de presencia y movimiento humano se ha decidido utilizar la técnica RSSI. Con los datos obtenidos, el siguiente capítulo se centrará en el diseño de un algoritmo no supervisado basado en árboles de decisión para analizar los datos e inferir resultados.

Capítulo 4

Algoritmos de decisión

En este capítulo se van a estudiar los diferentes tipos de algoritmos para la detección de patrones en un conjunto de datos. Se abordarán en especial los métodos no supervisados, con énfasis en los algoritmos de *clustering* como *k-means* y *hierarchical clustering*.

El objetivo del *machine learning* es crear un modelo que nos permita resolver una tarea propuesta, entrenando el modelo con una gran cantidad de datos. El modelo aprende de esta información y es capaz de hacer predicciones. El uso de *machine learning* en este proyecto se justifica por la intención de desarrollar un proyecto ligero y de rápida implementación. Al emplear este enfoque, se evita agregar el nivel de complejidad conceptual que se alcanzaría usando redes neuronales u otros algoritmos más complejos.

4.1. Tipos de algoritmos

Un algoritmo es un conjunto de instrucciones definidas, no ambiguas, ordenadas y finitas que permite, solucionar un problema. Dentro del conjunto de los algoritmos, existen los llamados algoritmos basados en árboles de decisión. Un árbol de decisión es un modelo de predicción utilizado en ámbitos como inteligencia artificial. Dado un conjunto de datos, se fabrican diagramas de construcciones lógicas que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema [14]. Podemos diferenciar dos tipos de algoritmos [15]:

- Los algoritmos supervisados son un tipo de técnica de *machine learning* cuyo objetivo es intentar encontrar una función que, dados unos datos de entrada, les asigne la etiqueta de salida adecuada. Por lo tanto, se le entregan muestras con las variables de entrada (x) y de salida o etiquetas (y).
- Los algoritmos no supervisados son un tipo de técnica de *machine learning* a los cuales se le deja encontrar libremente estructuras de interés en el conjunto de muestras. Las muestras entregadas no están etiquetadas, lo que significa que solo se entregan las variables de entrada (x), sin sus correspondientes variables de salida (y).

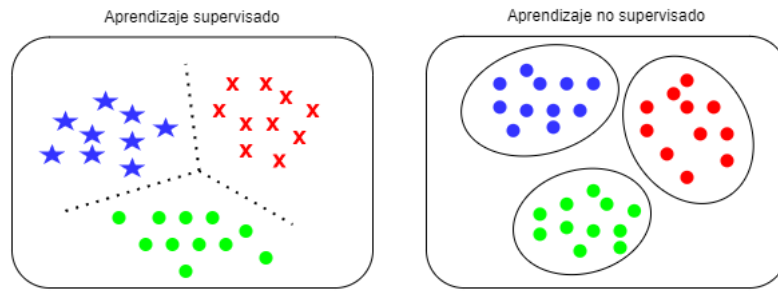


Figura 4.1: Comparación de los tipos de algoritmos (elaboración propia)

En la figura 4.1 se puede apreciar a la izquierda un algoritmo supervisado, donde se usan técnicas de clasificación que se enfocan en clasificar objetos en diferentes categorías. A la derecha un algoritmo no supervisado, donde las muestras son separadas en base a las características y la predicción se basa en el grupo al que pertenece.

El uso de aprendizaje no supervisado en este proyecto se justifica debido a la dificultad de etiquetar gran parte del conjunto de datos. Además, los algoritmos de aprendizaje no supervisado son más flexibles y se adaptan mejor a cambios en los datos o escenarios sin que se requiera una actualización de las etiquetas.

4.2. Clustering

El *Clustering* o agrupamiento es un tipo de aprendizaje no supervisado, cuyo principal objetivo es agrupar objetos por similitud, en grupos o conjuntos de manera que los miembros del mismo grupo tengan características similares. En la figura 4.2, [16], a la izquierda se pueden apreciar los datos sin clasificación; mientras que, a la derecha, se pueden observar las muestras divididas en tres colores o grupos basados en sus características. Cuando el algoritmo obtiene una nueva muestra, este verifica a que grupo pertenece según sus características y realiza la predicción.

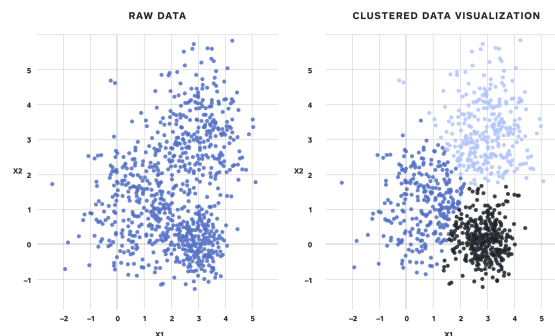


Figura 4.2: Ejemplo de *clustering* o agrupamiento

En este proyecto, se ha decidido emplear *clustering* debido a su capacidad de agrupar las muestras en diferentes conjuntos o grupos basados en sus características. Esta técnica resulta muy útil en casos donde se busque identificar ciertos patrones en el conjunto de datos, donde cada grupo representa una clase específica. Al utilizar *clustering*, se reduce el problema o la decisión a una solución binaria, lo que permite una gran flexibilidad a la hora de interpretar los resultados.

4.2.1. Algoritmo *k-means*

El algoritmo *k-means* se destaca como uno de los métodos más ampliamente utilizados y reconocidos dentro de los algoritmos de agrupamiento. *K-means* o *k-medias* es un algoritmo de agrupamiento iterativo no supervisado que tiene como objetivo encontrar máximos locales en cada iteración. El algoritmo trabaja iterativamente para asignar a cada muestra uno de los “*k*” grupos basándose en sus características [17]. Como resultado de la ejecución del algoritmo, se obtienen:

- Los centroides de cada grupo, que representan las “coordenadas” de cada uno de los *k* conjuntos que se utilizarán para poder etiquetar nuevas muestras.
- Las etiquetas para el conjunto de datos de entrenamiento, donde cada etiqueta indica a que grupo pertenece cada muestra.

Los grupos se definen de manera natural, es decir, van ajustando su posición en cada iteración del proceso, hasta que el algoritmo converge. Para ejecutarlo, se debe proporcionar como entrada el conjunto de datos y un valor de “*k*”, correspondiente a los grupos que se desean. Las posiciones iniciales de los “*k*” centroides serán asignados de manera aleatoria a cualquier punto del conjunto de datos de entrada. El conjunto de datos serán las características o *features* de cada punto. Posteriormente, se iterará en dos pasos [18]:

1. Asignación de datos: En este paso, cada muestra del conjunto de datos se asigna al centroide más cercano basado en la distancia cuadrada euclidiana. Se utiliza la ecuación 4.1, donde $dist()$ es la distancia euclidiana estándar.

$$\arg \min_{c_k \in C} dist(c_i, x)^2 \quad (4.1)$$

donde c_i es el centroide de cada grupo y x la muestra de entrada.

2. Actualización del centroide: En este paso, los centroides de cada grupo se recalculan tomando la media de todos los puntos asignados en el paso anterior, siguiendo la ecuación 4.2.

$$c_i = \frac{1}{|S_i|} \sum_{x_i \in S_i} x_i \quad (4.2)$$

siendo S_i el número total de muestras del grupo, c_i el nuevo centro del grupo y x_i la muestra de entrada.

Cada centroide representa un conjunto de valores que define los grupos resultantes. Examinando el valor del peso de los centroides, se puede utilizar para interpretar cualitativamente qué tipo de grupo representa cada clúster. En este proyecto, idealmente existirán dos centroides, uno que representa la presencia humana y otro que representa la ausencia de esta. En algún caso, podría haber tres, añadiendo un caso en el que el algoritmo detecte un falso positivo o un caso indeterminado.

La biblioteca empleada para implementar el algoritmo de *k-means* es *scikit-learn*. Dentro de esta biblioteca, se puede aprovechar la implementación de *k-means* proporcionada por la clase *KMeans* del módulo *sklearn.cluster*. Esta clase ofrece una serie de parámetros que permite personalizar y ajustar el comportamiento del algoritmo para optimizarlo al máximo y adaptarlo a las necesidades específicas. Estos parámetros configurables han sido estudiados en el anexo C.

4.2.2. Hierarchical clustering

El agrupamiento jerárquico o *hierarchical clustering* es un algoritmo que construye una jerarquía de clústeres. Este algoritmo se basa en el concepto de fusionar o dividir los grupos de manera iterativa. Inicialmente, cada dato se asigna a su propio grupo individual y luego, el algoritmo procede a agruparlos en conjuntos más amplios en función de sus características o distancia entre las muestras. Este proceso se repite iterativamente hasta que se obtiene un único grupo que contiene todo el conjunto de muestras. El agrupamiento jerárquico se puede representar a través de un dendrograma, que es un gráfico en forma de árbol que muestra cómo se agruparon los datos, como se muestra en la figura 4.3.

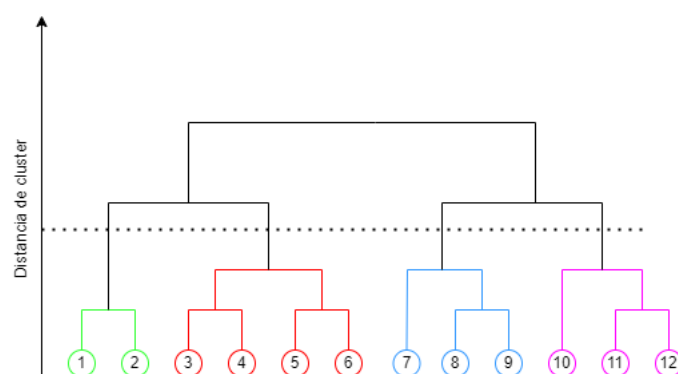


Figura 4.3: Ejemplo de *hierarchical clustering* (elaboración propia)

A diferencia del algoritmo *k-means*, el agrupamiento jerárquico no requiere especificar previamente el número de grupos. Una de las ventajas clave del agrupamiento jerárquico es su capacidad para detectar estructuras complejas y no convencionales en

los datos. Puede capturar grupos de diferentes formas y tamaños, lo que lo convierte en un algoritmo muy útil en conjuntos de datos donde los clústeres no son necesariamente esféricos o tienen una forma geométrica específica.

Gracias a una representación a través de un dendrograma, se puede elegir un nivel específico para obtener una división de los datos en clústeres. Esto proporciona flexibilidad en la interpretación y permite obtener agrupaciones a diferentes escalas de detalles. Además, tiene la ventaja de ser reproducible, es decir, que al ejecutar el algoritmo varias veces al mismo conjunto de datos, se obtendrá la misma estructura y los mismos resultados, lo que facilita la comparación y la repetición de resultados.

4.2.3. Comparación entre *k-means* y *hierarchical clustering*

Las diferencias entre los algoritmos de *k-means* y *hierarchical clustering* son [17]:

- A diferencia del agrupamiento *k-means*, el agrupamiento jerárquico no puede manejar grandes cantidades de datos. Esto se debe a que la complejidad temporal de *k-means* es lineal, es decir, $O(n)$, mientras que la del agrupamiento jerárquico es cuadrática, es decir, $O(n^2)$.
- En el agrupamiento *k-means* se comienza con una elección aleatoria de clústeres, y los resultados generados al ejecutar el algoritmo varias veces pueden ser distintos. En el agrupamiento jerárquico, los resultados son reproducibles, es decir se obtendrán siempre los mismos resultados.
- La agrupación *k-means* funciona bien cuando los grupos tienen una forma hipersférica, como un círculo en 2D o una esfera en 3D.
- *K-means* no permite datos con ruido, mientras que el agrupamiento jerárquico puede usar directamente el conjunto de datos con ruido para el agrupamiento.

Finalmente, se ha decidido emplear el algoritmo de *k-means* debido a que este algoritmo es más eficiente y compatible en comparación con el agrupamiento jerárquico, especialmente cuando se trabaja con grandes conjuntos de datos. Además, *k-means* produce una división clara de los datos en “k” grupos distintos, lo que facilita la comprensión e interpretación de los resultados obtenidos. Asimismo, en el algoritmo de *k-means*, el usuario es el que especifica el número de clústeres que desea, lo que es beneficioso si se tiene conocimiento previo sobre el conjunto de datos. Esto permite tener una predicción más precisa de las muestras a los diferentes grupos.

Una vez se ha elegido e implementado el algoritmo para el análisis de los datos y la obtención resultados, en el siguiente apartado se realizarán varias pruebas. Estas pruebas son fundamentales para evaluar la eficiencia y precisión del algoritmo en diversos escenarios y condiciones. Asimismo, permiten identificar posibles errores o mejoras para lograr un resultado óptimo. A través de este conjunto de pruebas se intentará garantizar una alta calidad y fiabilidad del algoritmo.

Capítulo 5

Pruebas y resultados

En este capítulo, se presentarán los detalles de las pruebas y los resultados obtenidos en el proyecto. Las pruebas se llevarán a cabo en una serie de escenarios para la detección de personas por cambios de radiofrecuencia. En los escenarios, se muestran diferentes situaciones que pueden suceder en la vida cotidiana. Cada uno tendrá un enfoque diferente y presentará una serie de desafíos específicos que deberá superar el algoritmo para proporcionar resultados precisos y fiables.

Por lo tanto, los escenarios que se llevarán a cabo son:

1. Un escenario donde los dispositivos tienen línea de visión directa y pasa una persona entre medias.
2. Un escenario donde los dispositivos tienen línea de visión directa y una persona se queda en medio del enlace.
3. Un escenario donde los dispositivos no tienen línea de visión directa y pasa una persona entre medias.
4. Un escenario donde los dispositivos no tienen línea de visión directa y una persona se queda en medio del enlace.
5. Por último, se realizará un análisis en profundidad de los falsos positivos y se determinarán las dimensiones mínimas de un objeto que el enlace puede detectar.

Todos estos escenarios serán simulados dentro de la Cátedra de Securitas Direct Verisure. Se ha intentado simular un escenario realista que represente una sala de una vivienda donde se encuentren instalados tanto la unidad central como el detector de acceso de Securitas Direct.

Para evaluar el algoritmo, se han tenido en cuenta las siguientes métricas [19]:

- La precisión o *accuracy* en inglés, es una métrica utilizada en aprendizaje automático y estadísticas para evaluar el rendimiento de un modelo. Se define como la proporción de predicciones correctas realizadas por el modelo, en relación con el total de predicciones realizadas, según la ecuación 5.1.

$$Accuracy[\%] = \frac{Correct\ predictions}{Total\ predictions} \cdot 100 \quad (5.1)$$

El resultado de la precisión será un valor entre 0 y 1, donde un valor más cercano a 1 indica un alto porcentaje de predicciones correctas y, por lo tanto, un mejor rendimiento del modelo.

- El *Adjusted Rand Score* (ARS) es una métrica utilizada para calcular una medida de similitud entre dos agrupaciones diferentes. El ARS tiene en cuenta todos los pares de muestras y cuenta los pares que se asignan a las mismas o a diferentes agrupaciones predichas en comparación con las agrupaciones verdaderas, ajustando la probabilidad aleatoria. Esta métrica se corresponde con la siguiente ecuación:

$$ARS = \frac{RI - ExpectedRI}{RI_{max} - ExpectedRI} \quad (5.2)$$

siendo *RI* (*Rand Index*) una medida de similitud que compara todas las muestras y cuántas se asignan a la misma o diferente grupo, dentro de las agrupaciones verdaderas y las agrupaciones predichas; *ExpectedRI* el valor esperado de *RI* y *RI_{max}* el valor máximo de *RI*.

El ARS asigna un valor cercano a 1 cuando las agrupaciones predichas son similares a las verdaderas, indicando un buen rendimiento del algoritmo. Un valor cercano a -1 indica grandes diferencias entre las agrupaciones predichas y las verdaderas, destacando un mal rendimiento del algoritmo.

- El índice de *Fowlkes Mallow Score* (FMS) es una puntuación que evalúa el grado en que las asignaciones de grupos son correctas. En particular, calcula la media geométrica entre precisión y recuperación o *recall*. Su fórmula es la siguiente:

$$FMI = \sqrt{accuracy \cdot recall} \quad (5.3)$$

El resultado será un valor acotado entre 0 y 1, donde valores más altos indican una mayor concordancia entre las asignaciones de grupos y, por lo tanto, un mejor rendimiento del algoritmo de agrupamiento. Un valor cercano a 0 indica que las asignaciones de grupos son prácticamente aleatorias.

- El *Silhouette Score* es una métrica que intenta describir qué tan similar es un punto de datos en relación con otros puntos dentro de su propio grupo y de fuera de este. Esta medida se utiliza para evaluar la calidad de las agrupaciones en términos de la coherencia y la separación entre los grupos. Esta métrica tiene la siguiente ecuación:

$$SilhouetteScore = \frac{1}{N} \sum_{i=1}^N \left(\frac{b_i - a_i}{\max(a_i, b_i)} \right) \quad (5.4)$$

siendo N el número total de muestras en el grupo, a_i la distancia media entre el punto i y todos los puntos de su grupo y b_i la distancia media entre el punto i y todos los puntos en el grupo más cercano que no contiene al punto i .

El *Silhouette Score* varía en un rango de -1 a 1, donde un valor cercano a 1 indica que los objetos están bien asignados a sus respectivos grupos y están separados de otros grupos, y un valor cercano a -1 indica que los objetos están asignados incorrectamente a los grupos y que están más cerca de objetos de otros grupos.

5.1. Escenario 1

En el primer escenario, los dos dispositivos tienen una línea de visión directa y se simula que el intruso pasa entre medias del enlace. Esta configuración permite evaluar el rendimiento del algoritmo en un entorno ideal, donde no hay obstáculos físicos que interfieran con la señal de radiofrecuencia, y se produce la detección más simple.

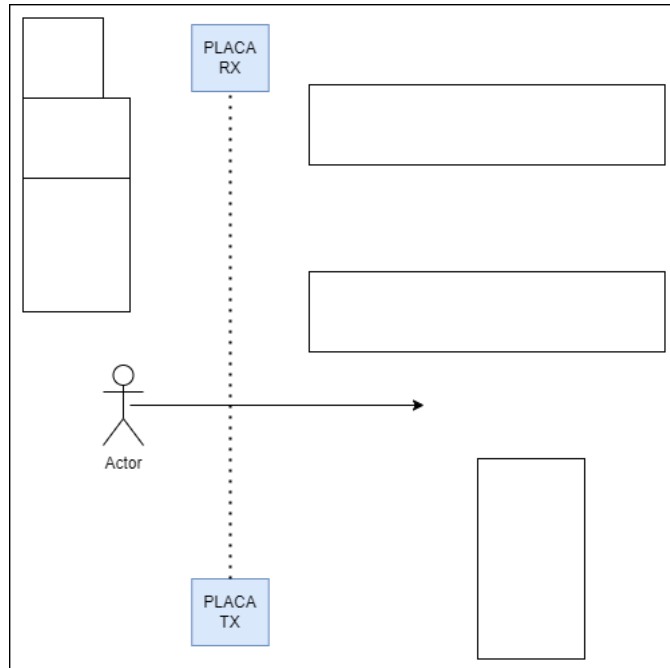


Figura 5.1: Esquema del escenario 1 (elaboración propia)

Se han realizado nueve pruebas para el entrenamiento del algoritmo, las cuales se han aleatorizado para evitar patrones temporales. En la figura 5.2 se muestra el resultado obtenido después del entrenamiento.

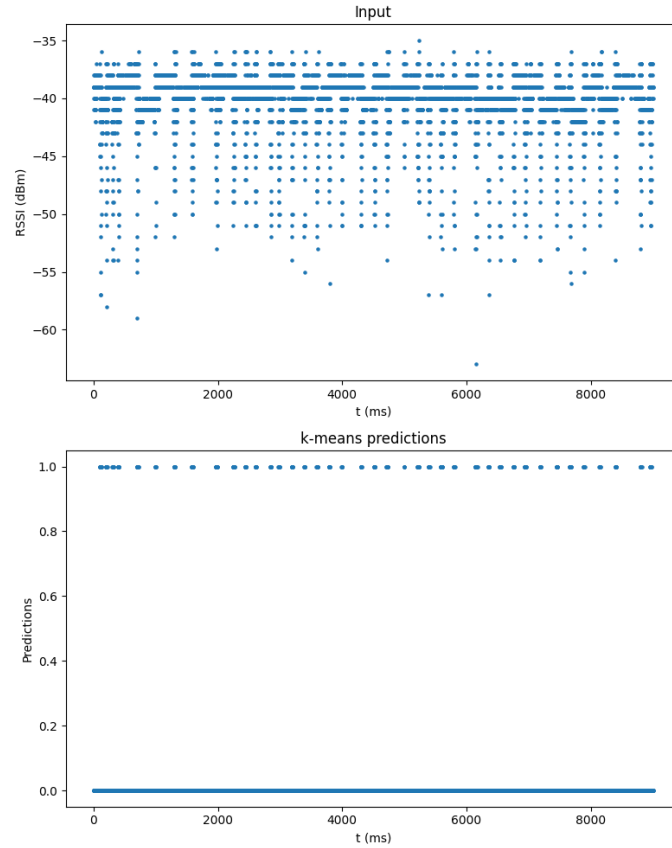
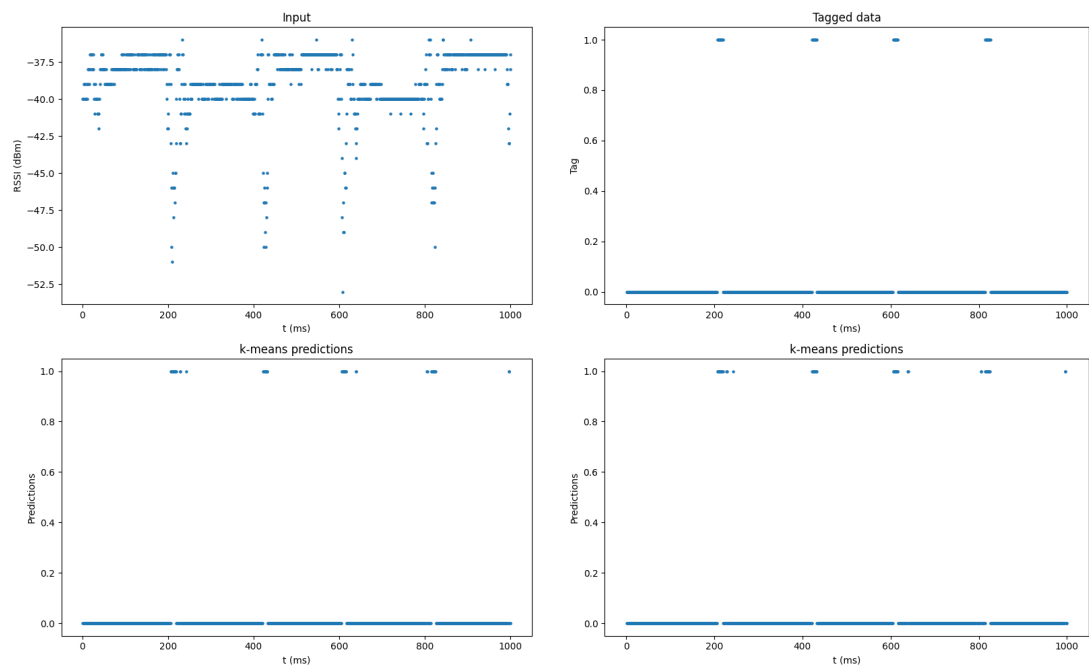


Figura 5.2: Entrenamiento en el escenario 1

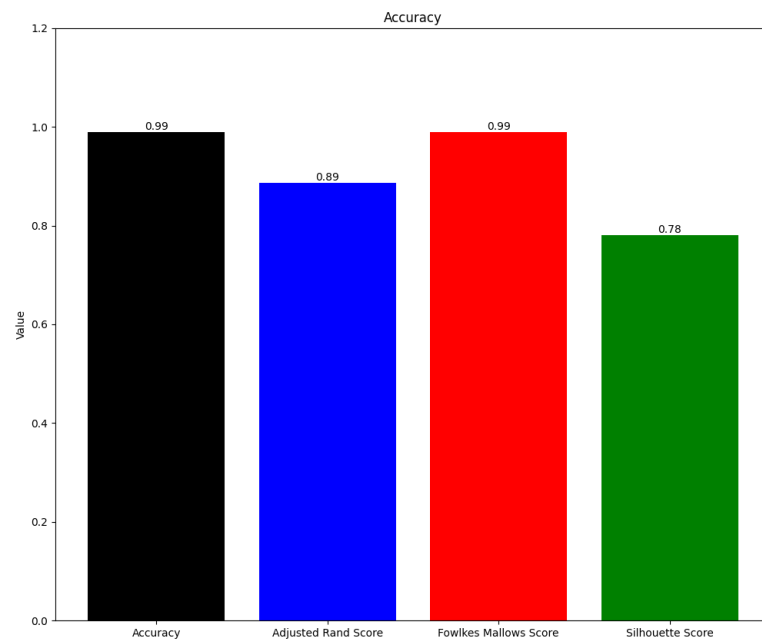
Por último, se ha realizado una décima prueba, la cual fue etiquetada para evaluar el rendimiento del algoritmo. En esta prueba, se ha intentado pasar entre medias del enlace cada 20 segundos, que equivale a cada 200 muestras. En la figura 5.3a se muestran las predicciones de esta última prueba, en la figura 5.3b la comparación entre las etiquetas y las predicciones del algoritmo, y por último, en la figura 5.3c se muestran los parámetros de evaluación del algoritmo *k-means*.

Analizando los resultados, se revela que en un escenario ideal, el algoritmo logra un nivel de precisión alto, con un valor de 99 % que indica que la mayoría de las predicciones fueron correctas. El ARS de 0,89 nos indica un rendimiento sólido en la detección de presencia y movimiento humano. Asimismo, se obtuvo un FMS de 0,99 lo que muestra que el algoritmo logró clasificar con precisión los datos en los grupos correspondientes. Por otro lado, el *Silhouette Score* de 0,78 indica que la separación entre los grupos encontrados es adecuada y que los datos se agrupan de manera coherente. En resumen, estos resultados indican que el algoritmo detecta de manera muy fiable la presencia y el movimiento humano.



(a) Predicciones del algoritmo

(b) Comparación de etiquetas y predicciones



(c) Evaluación del algoritmo

Figura 5.3: Algoritmo *k-means* en el escenario 1

5.2. Escenario 2

En el segundo escenario, los dos dispositivos tienen una línea de visión directa y simularemos que el intruso pasa entre medias del enlace, pero esta vez se quedará parado en medio. Esta configuración nos permite evaluar el rendimiento del algoritmo en un entorno más realista, donde buscamos que el algoritmo siga detectando presencia, aunque el intruso se encuentre quieto en medio del enlace.

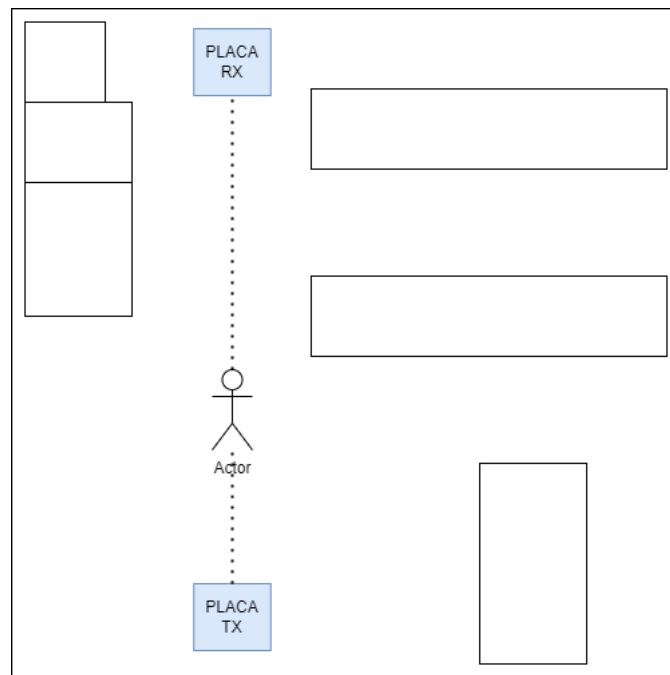


Figura 5.4: Esquema del escenario 2 (elaboración propia)

Se han realizado nueve pruebas para el entrenamiento del algoritmo, las cuales se han aleatorizado para evitar patrones temporales. En la figura 5.5 se muestra el resultado obtenido después del entrenamiento.

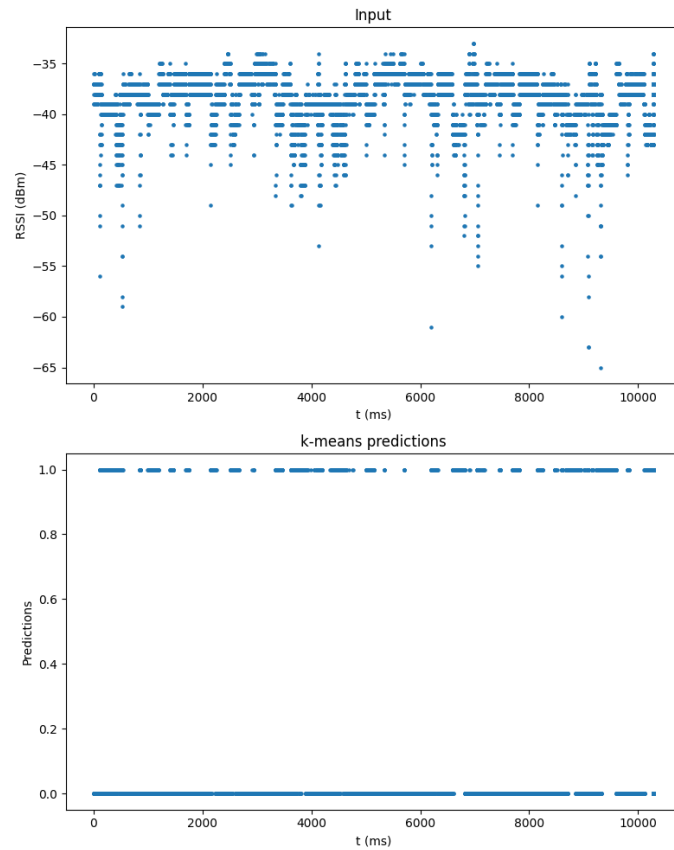
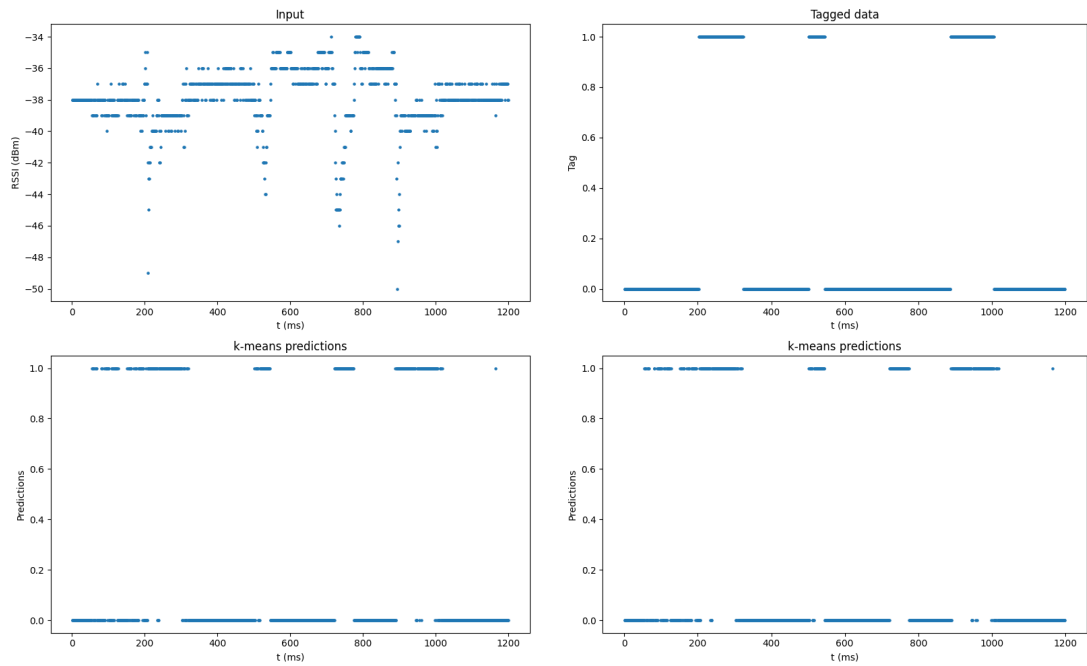


Figura 5.5: Entrenamiento en el escenario 2

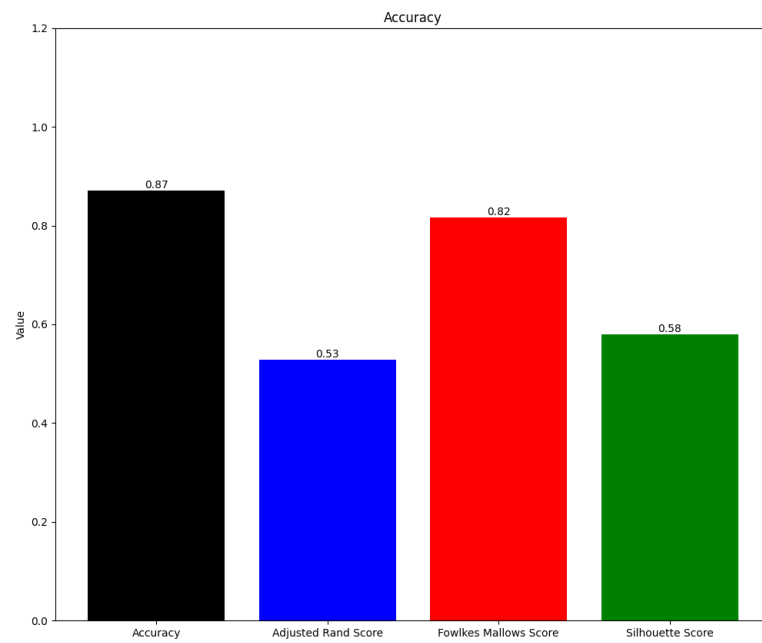
Por último, se ha realizado una décima prueba, la cual fue etiquetada para evaluar el rendimiento del algoritmo. En esta prueba, se ha intentado pasar entre medias del enlace en el segundo 50, correspondiente a la muestra 500. Además, el intruso se ha quedado en medio del enlace durante aproximadamente 10 segundos, entre los segundos 20 y 30, que equivale a las muestras 200-300, y entre los segundos 90 y 100, que corresponden a las muestras 900-1000. En la figura 5.6a se muestran las predicciones de esta última prueba, en la 5.6b la comparación entre las etiquetas y las predicciones del algoritmo, y por último, en la 5.6c se muestran los parámetros de evaluación del algoritmo *k-means*.

Analizando los resultados, revelan que en este escenario, el algoritmo logra un nivel de precisión alto, con un valor de 87% que indica que la mayoría de las predicciones fueron correctas. El ARS de 0,53 nos indica un nivel razonable de similitud entre las dos agrupaciones. Asimismo, se obtuvo un FMS de 0,87 lo que muestra que el algoritmo logró clasificar con una buena precisión los datos en los grupos correspondientes. Por otro lado, el *Silhouette Score* de 0,58 indica que existe un pequeño solapamiento a la hora de asignar la muestra en los grupos. En resumen, estos resultados indican que el algoritmo detecta de manera fiable la presencia y el movimiento humano.



(a) Predicciones del algoritmo

(b) Comparación de etiquetas y predicciones



(c) Evaluación del algoritmo

Figura 5.6: Algoritmo *k-means* en el escenario 2

5.3. Escenario 3

En el tercer escenario, es una situación similar al primero, pero con la particularidad de que los dos dispositivos no tienen una línea de visión directa. Esta configuración nos permite evaluar el rendimiento del sistema en un entorno más realista, donde se introducen obstáculos físicos que pueden afectar a la propagación de la señal de radiofrecuencia y a la detección.

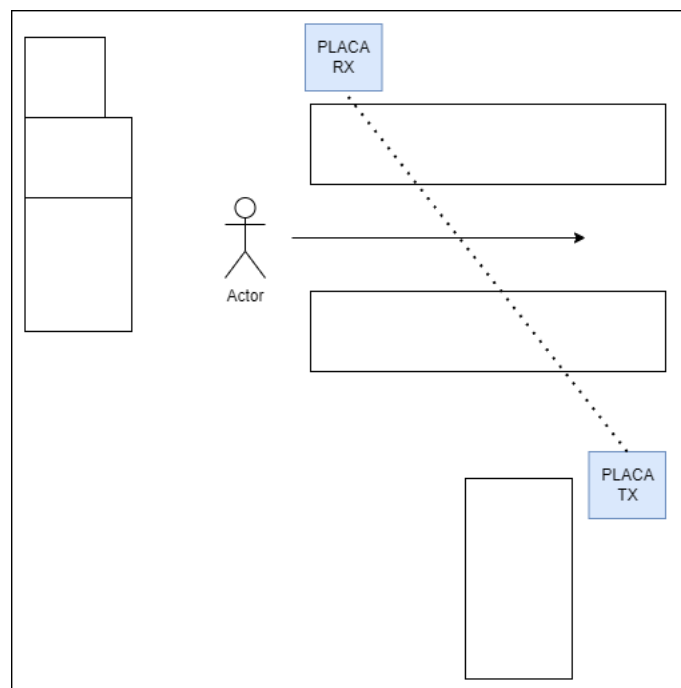


Figura 5.7: Esquema del escenario 3 (elaboración propia)

Se han realizado nueve pruebas para el entrenamiento del algoritmo, las cuales se han aleatorizado para evitar patrones temporales. En la figura 5.8 se muestra el resultado obtenido después del entrenamiento.

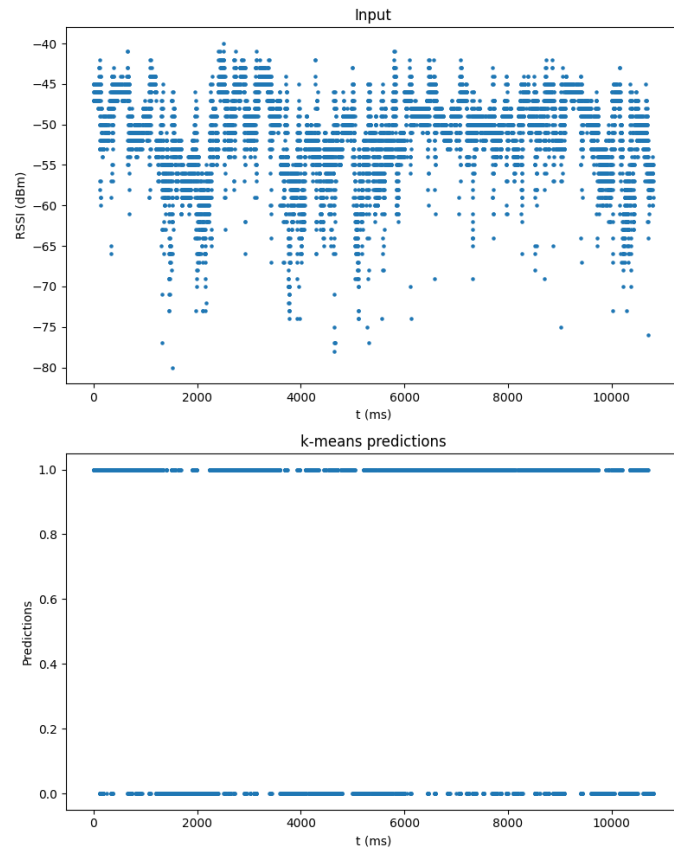
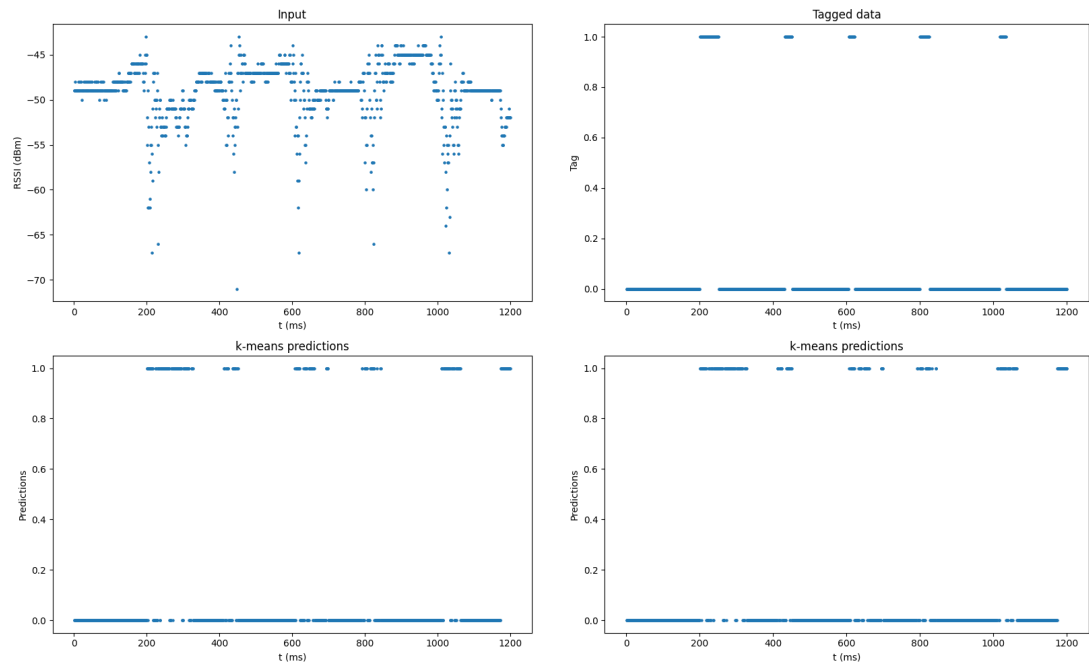


Figura 5.8: Entrenamiento en el escenario 3

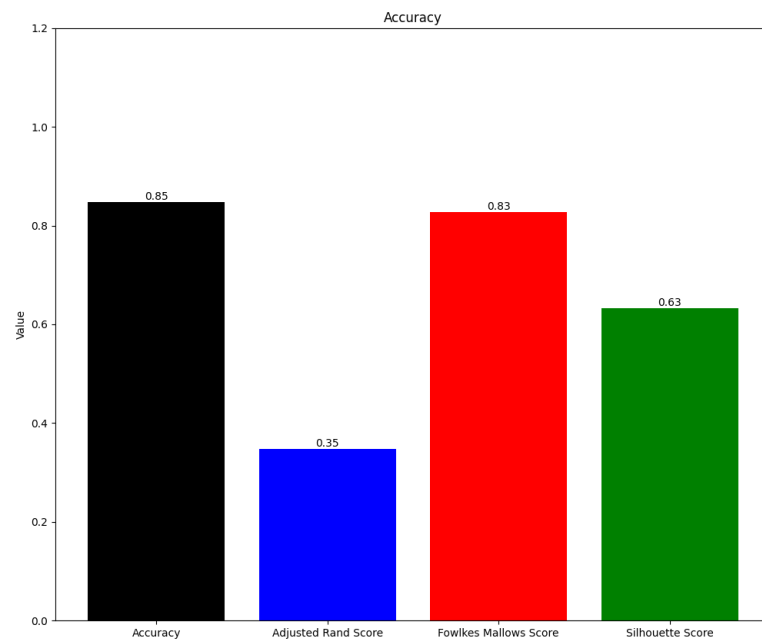
Por último, se ha realizado una décima prueba, la cual fue etiquetada para evaluar el rendimiento del algoritmo. En esta prueba, se ha intentado pasar entre medias del enlace cada 20 segundos, que equivale a cada 200 muestras. En la figura 5.9a se muestran las predicciones de esta última prueba, en la 5.9b la comparación entre las etiquetas y las predicciones del algoritmo, y por último, en la 5.9c se muestran los parámetros de evaluación del algoritmo *k-means*.

Analizando los resultados, revelan que en este escenario donde no existe línea de visión directa, el algoritmo logra un buen nivel de precisión, con un valor de 85 % que indica que la gran parte de las predicciones fueron correctas. El ARS de 0,35 nos indica que el algoritmo tiene una capacidad limitada para discriminar entre presencia y no presencia, ya que al haber obstáculos, la potencia es menor. Sin embargo, el valor de FMS es de 0,83 lo que muestra que el algoritmo logró clasificar con precisión los datos en los grupos correspondientes. Por otro lado, el *Silhouette Score* de 0,63 indica que la separación entre los grupos encontrados es adecuada y que los datos se agrupan de manera coherente. En resumen, estos resultados indican que, a pesar de la falta de línea de visión directa, el algoritmo aún logra realizar una detección efectiva de presencia humana.



(a) Predicciones del algoritmo

(b) Comparación de etiquetas y predicciones



(c) Evaluación del algoritmo

Figura 5.9: Algoritmo *k-means* en el escenario 3

5.4. Escenario 4

En el cuarto escenario, es una situación similar al segundo, pero con la particularidad de que los dos dispositivos no tienen una línea de visión directa. Además, nos enfocamos en buscar que el algoritmo siga detectando presencia, aunque el intruso se encuentre quieto en medio del enlace. Esta configuración nos permite evaluar el rendimiento del sistema en un entorno más realista, donde se introducen obstáculos físicos que pueden afectar a la propagación de la señal de radiofrecuencia y a la detección.

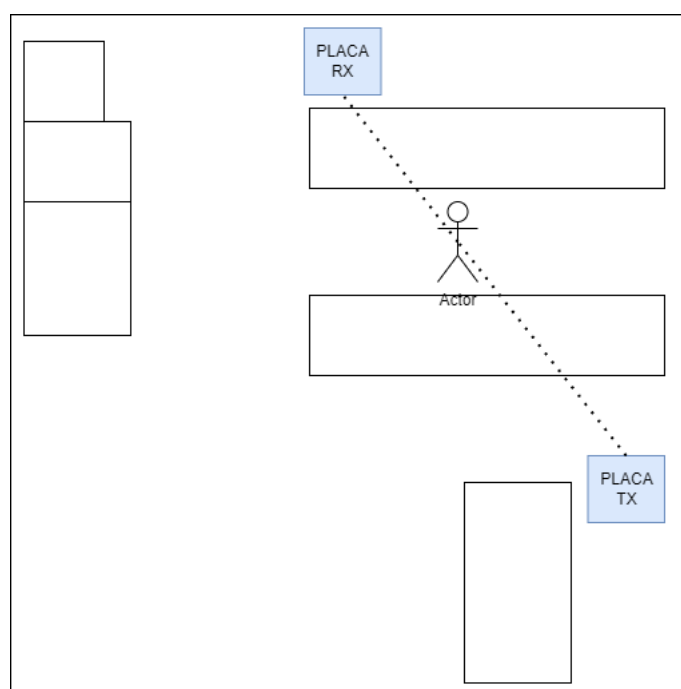


Figura 5.10: Esquema del escenario 4 (elaboración propia)

Se han realizado nueve pruebas para el entrenamiento del algoritmo, las cuales se han aleatorizado para evitar patrones temporales. En la figura 5.11 se muestra el resultado obtenido después del entrenamiento.

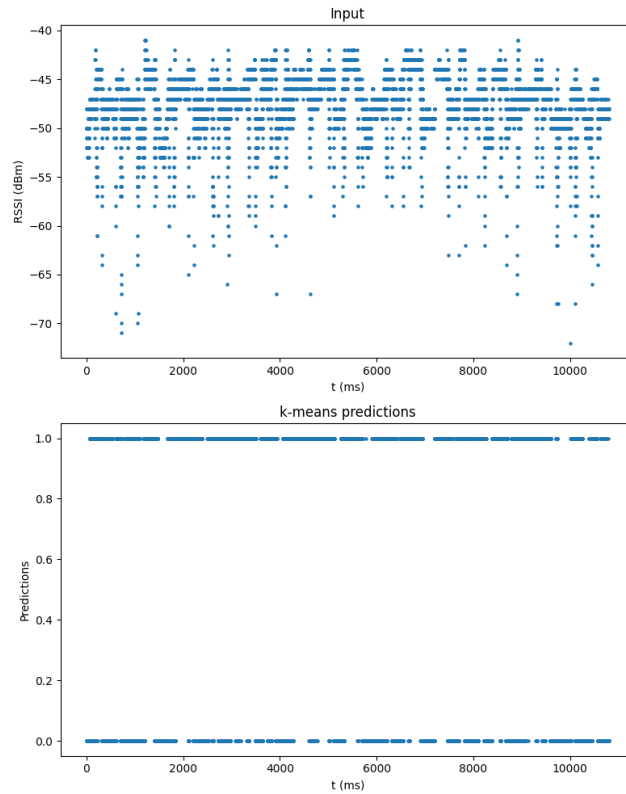
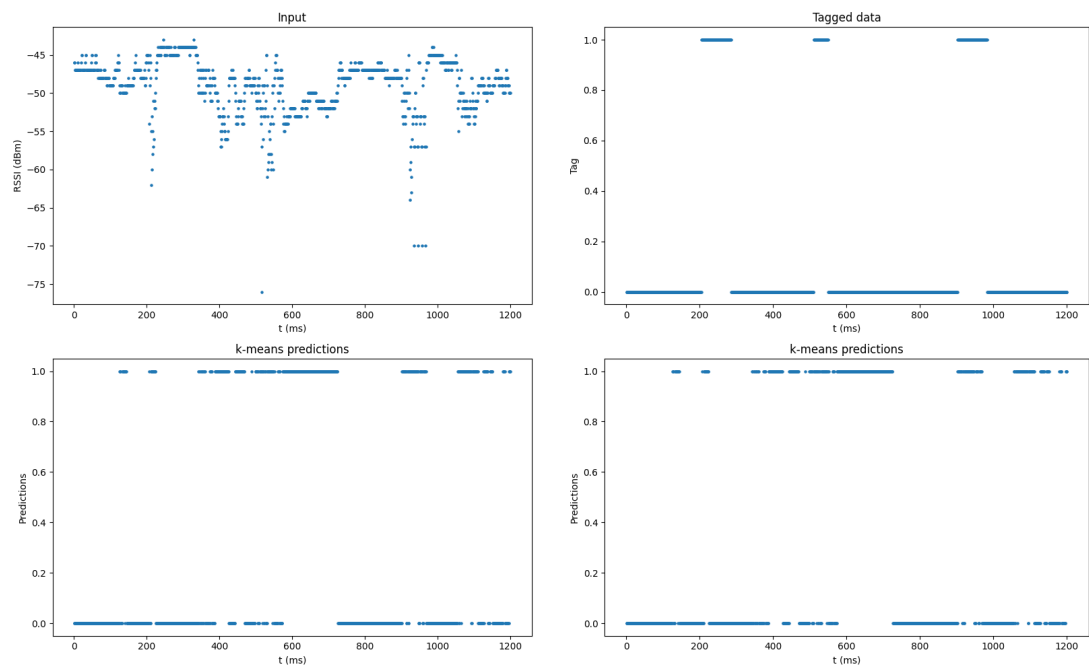


Figura 5.11: Entrenamiento en el escenario 4

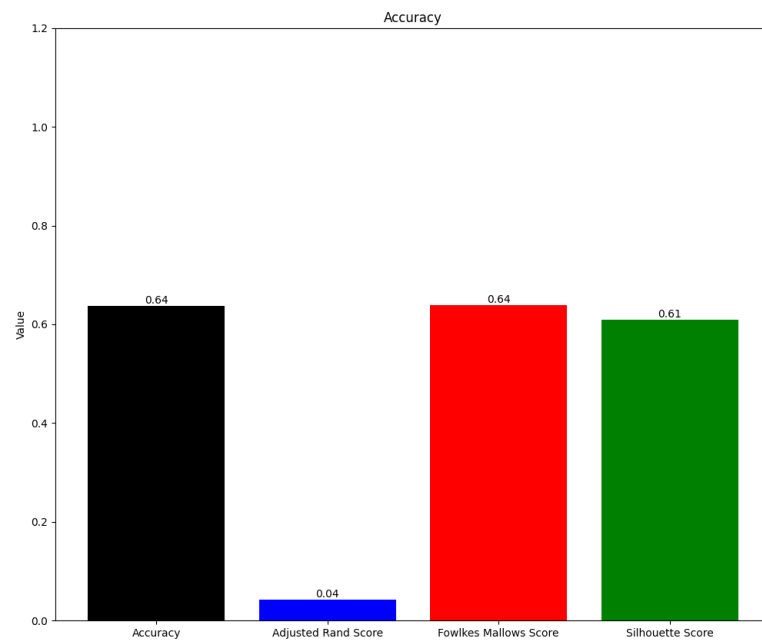
Por último, se ha realizado una décima prueba, la cual fue etiquetada para evaluar el rendimiento del algoritmo. En esta prueba, se ha intentado pasar entre medias del enlace en el segundo 50, correspondiente a la muestra 500. Asimismo, el intruso se ha quedado en medio del enlace durante aproximadamente 10 segundos, entre los segundos 20 y 30, que equivale a las muestras 200-300, y entre los segundos 90 y 100, que corresponden a las muestras 900-1000. En la figura 5.12a se muestran las predicciones de esta última prueba, en la 5.12b la comparación entre las etiquetas y las predicciones del algoritmo, y por último, en la 5.12c se muestran los parámetros de evaluación del algoritmo *k-means*.

Analizando los resultados se revela que en este escenario donde no existe línea de visión directa, el algoritmo obtiene una precisión del 64 %, esto indica que el algoritmo tiene una capacidad razonable de detección. El ARS de 0,04 sugiere que existen dificultades considerables para discriminar entre la presencia y la ausencia de una persona. Sin embargo, el FMS tiene un valor aceptable de 0,64, lo que muestra que el algoritmo logró clasificar con cierta precisión los datos en los grupos correspondientes. Asimismo, el *Silhouette Score* de 0,61 indica que la separación entre los grupos encontrados es adecuada y que los datos se agrupan de manera coherente. En resumen, estos resultados indican que, a pesar de la falta de línea de visión directa y habiendo una persona quieta en el enlace, el algoritmo aún logra realizar una detección de presencia humana con una precisión razonable.



(a) Predicciones del algoritmo

(b) Comparación de etiquetas y predicciones



(c) Evaluación del algoritmo

Figura 5.12: Algoritmo *k-means* en el escenario 4

5.5. Análisis de los falsos positivos

La detección de los falsos positivos es un aspecto crítico en el análisis de datos de detección de presencia y movimiento humano. Después de evaluar los resultados en los cuatro escenarios estudiados en los apartados anteriores, es importante analizar y comprender los casos en los que se producen falsos positivos. Un falso positivo se produce cuando el sistema detecta erróneamente la presencia de un objeto en el enlace, y lo predice como una persona. Esta evaluación permitirá identificar las limitaciones del algoritmo y determinar el tamaño mínimo de un objeto que puede ser detectado. Mediante un análisis detallado de los falsos positivos, se busca mejorar la precisión y la efectividad del sistema de detección de presencia y movimiento humano, lo que resulta fundamental en aplicaciones de seguridad y control de interiores.

Para llevar a cabo este análisis, se utilizaron diferentes objetos con el objetivo de simular un escenario realista, como por ejemplo que se caiga un objeto de una estantería o el paso de un animal doméstico. Los objetos empleados son:

- Una caja de dimensiones 7,5x5x2 cm, que representa un objeto de tamaño pequeño.
- Una caja de dimensiones 20x18x4 cm, que representa un objeto de tamaño mediano.
- Una caja de dimensiones 30x20x7,5 cm, que representa un objeto de tamaño grande.
- Una silla de oficina, que representa un objeto con una forma y tamaño variado.

Estos objetos fueron empleados en el escenario ideal, representado en la figura 5.13, donde los dispositivos tienen línea de visión directa.

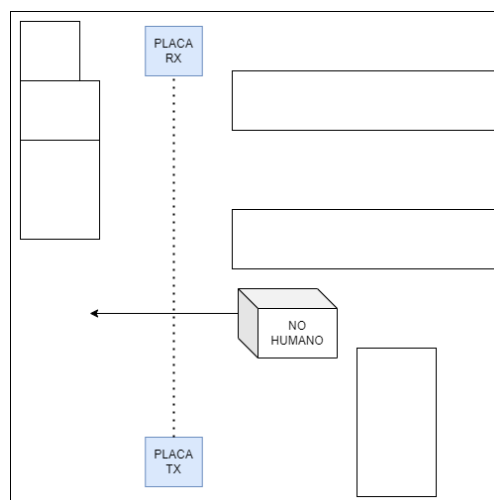


Figura 5.13: Esquema de escenario de detección de falsos positivos (elaboración propia)

Durante las pruebas, se combinarán los elementos mencionados previamente, junto con el paso de una persona entre el enlace, con el fin de evaluar el comportamiento del algoritmo. Cada 20 segundos se pasará el objeto por el enlace, correspondiendo a las muestras 200, 800 y 1400; y cada 40 segundos, pasará la persona por el enlace, correspondiendo a las muestras 400, 1000 y 1600.

Para la caja pequeña se obtienen los resultados representados en la figura 5.14. Se puede observar que el enlace no logra detectar la presencia de la caja pequeña. Esto se atribuye a que el RSSI utilizado en la simulación no se ve afectado por la presencia de un elemento de pequeñas dimensiones. En cambio, detecta perfectamente a la persona humana en las muestras 400, 1000 y 1600.

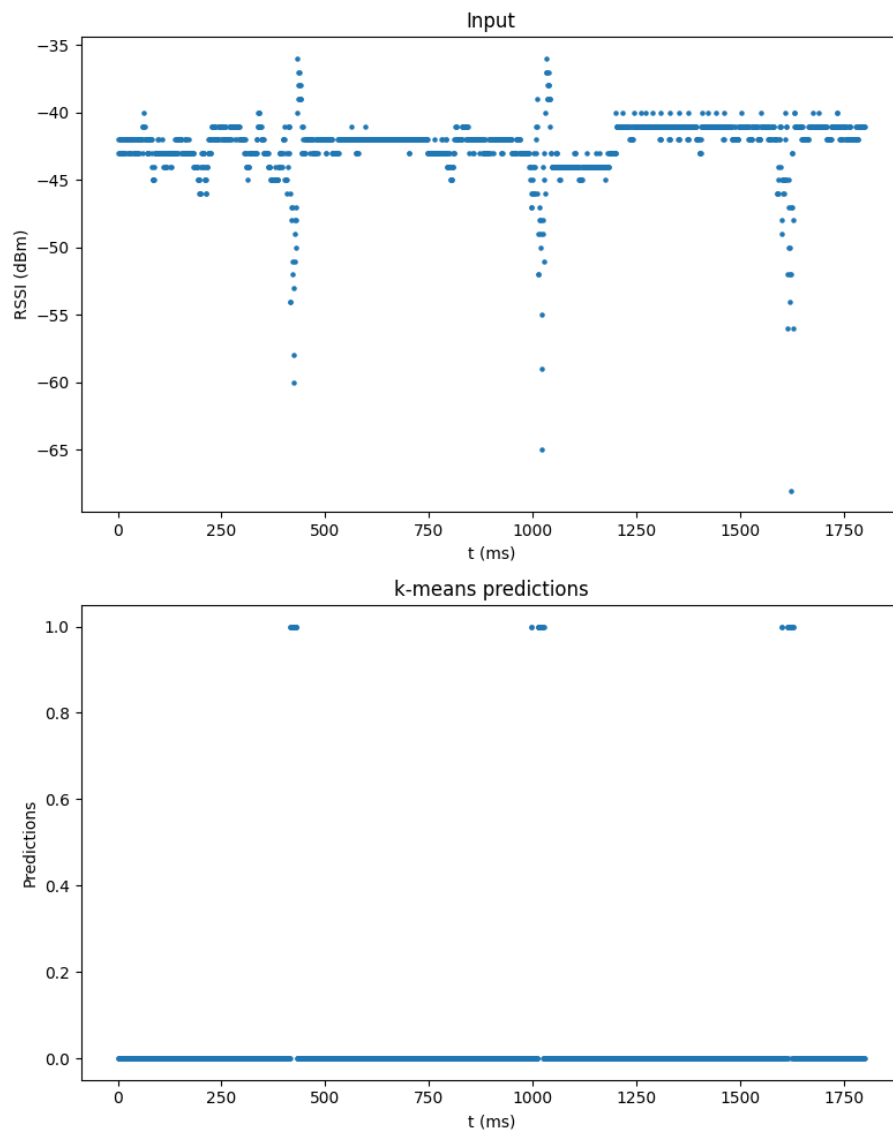


Figura 5.14: Predicciones para una caja pequeña

Para la caja mediana se obtienen los resultados representados en la figura 5.15. Se puede observar que el enlace tampoco logra detectar la presencia de la caja mediana. Esto se atribuye a que el RSSI utilizado en la simulación no se ve afectado por la presencia de un elemento de esas dimensiones. En cambio, detecta perfectamente a la persona humana en las muestras 400, 1000 y 1600.

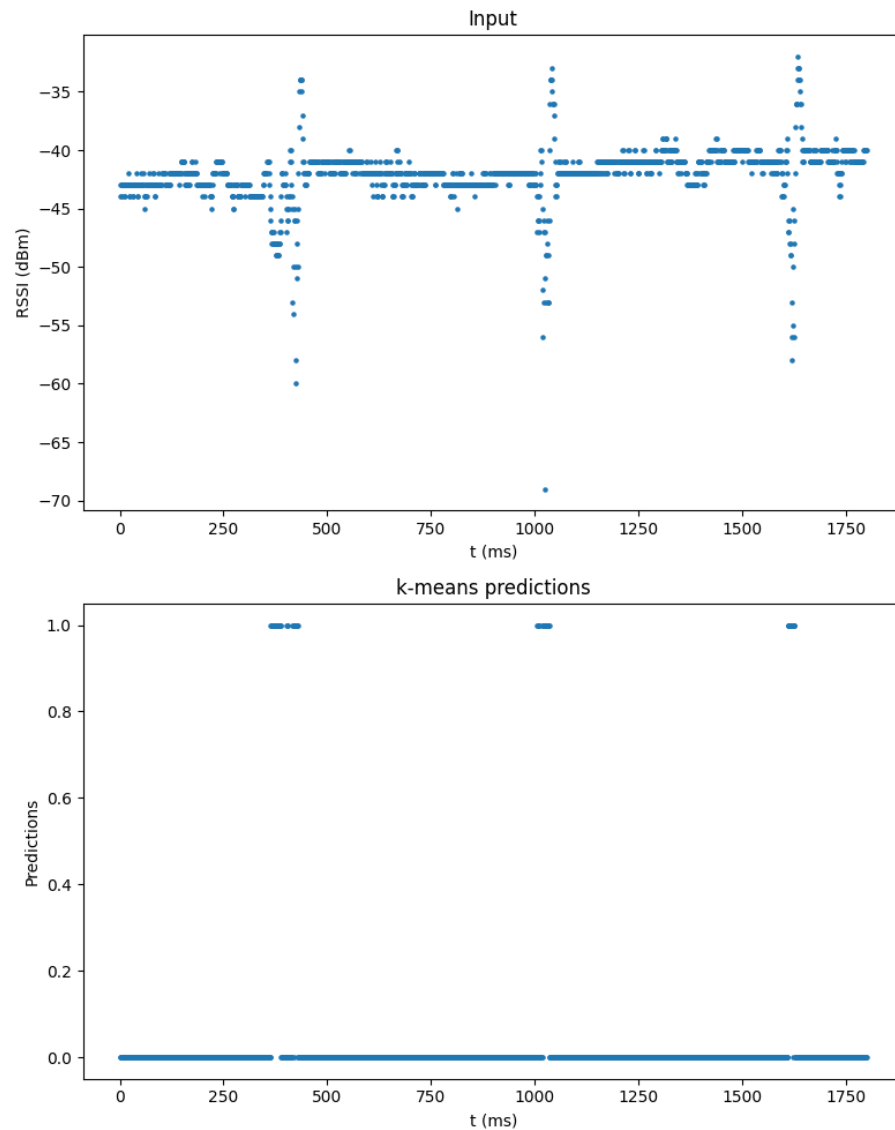


Figura 5.15: Predicciones para una caja mediana

Para la caja grande se obtienen los resultados representados en la figura 5.16. Se puede observar que, esta vez, el enlace si logra detectar la presencia de la caja. Se puede apreciar alrededor de las muestras 200, 800 y 1400 una disminución de RSSI provocado por el objeto. Sin embargo, el algoritmo no lo predice como presencia ya que la potencia recibida cuando pasa una persona, en las muestras 400, 1000 y 1600, es mucho menor.

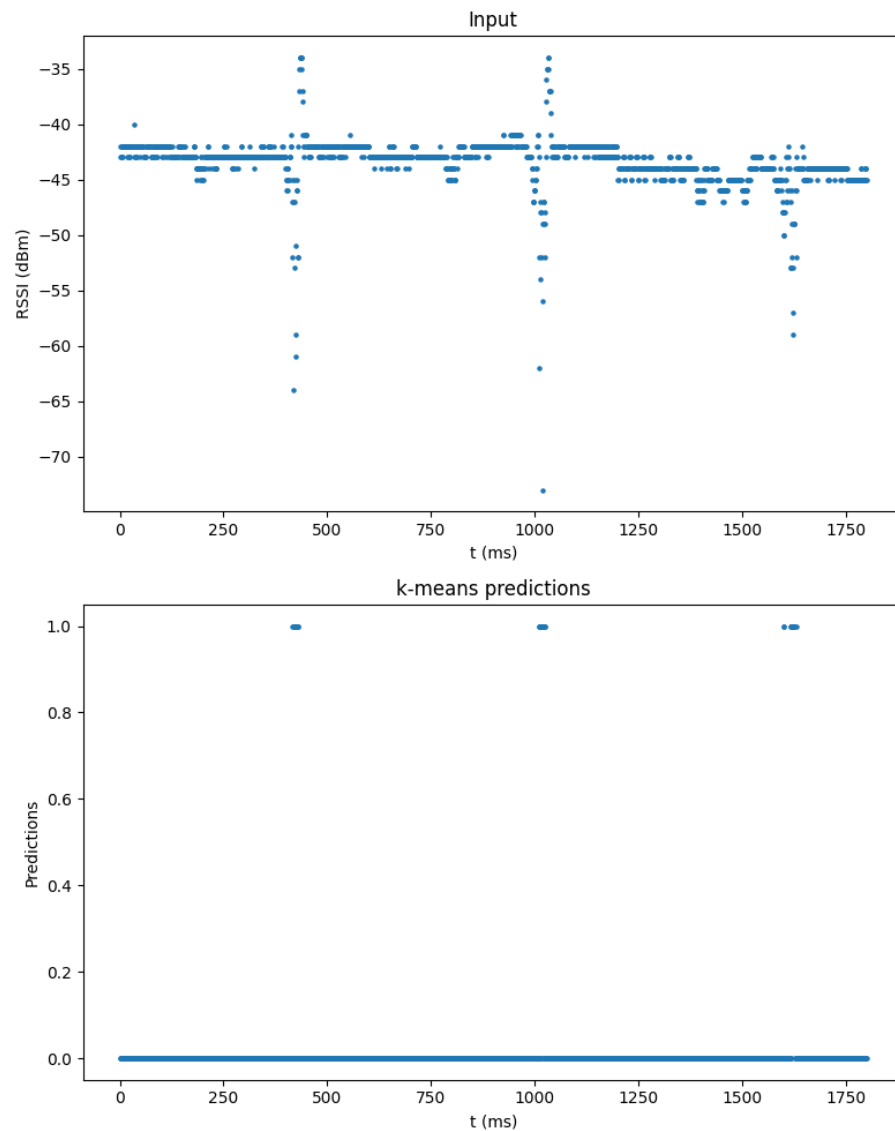


Figura 5.16: Predicciones para una caja grande

Para la silla de oficina se obtienen los resultados representados en la figura 5.17. Se puede observar que, como en el caso anterior, el enlace sí logra detectar la presencia de la silla. Se puede apreciar alrededor de las muestras 200, 800 y 1400 una disminución de RSSI provocado por el objeto. Esta vez, el algoritmo si que predice como presencia cuando la silla pasa por el enlace, ya que la potencia recibida es similar a cuando pasa una persona, en las muestras 400, 1000 y 1600.

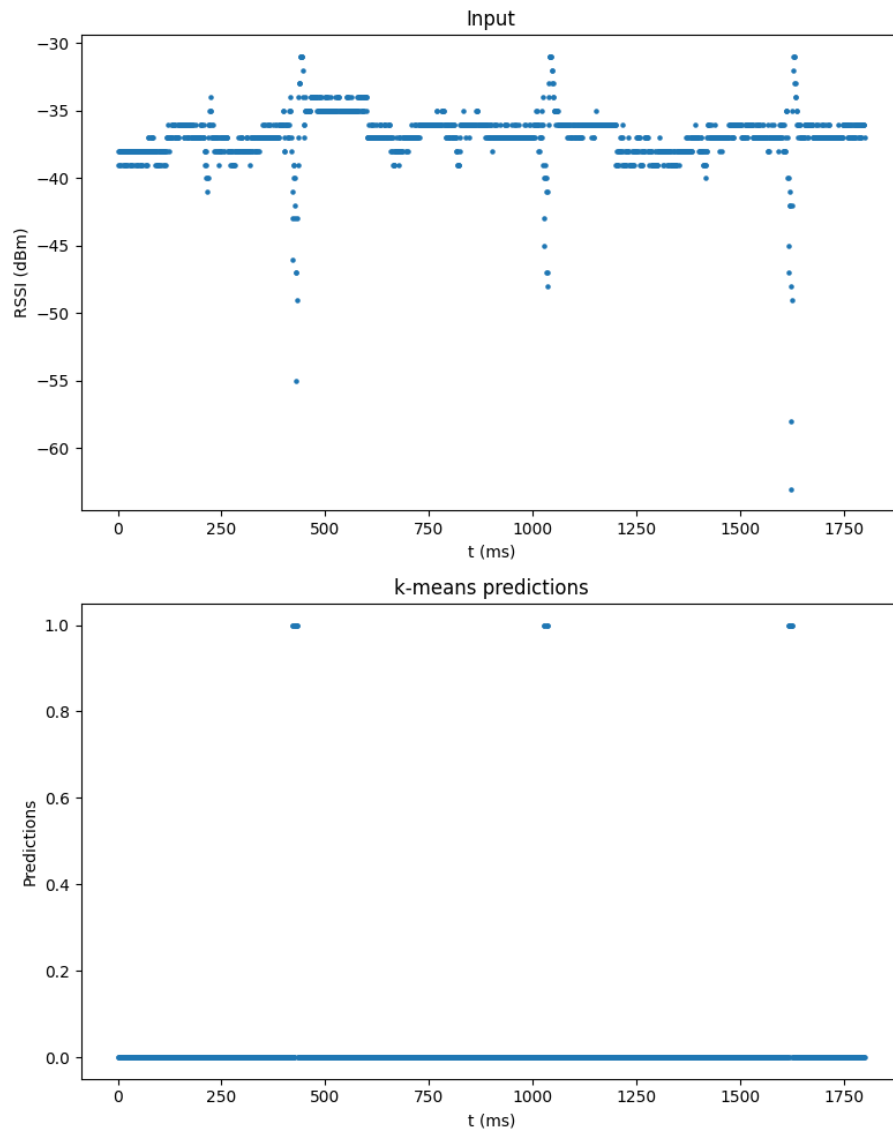


Figura 5.17: Predicciones para una silla de oficina con dos clústeres

Una solución para abordar esta situación es aumentar el número de clústeres del algoritmo a 3. Al incrementar el número de grupos, se pueden distinguir y clasificar de manera más precisa los siguientes escenarios: la presencia humana, la ausencia de esta y, por último, los falsos positivos u objetos que afectan significativamente la potencia del RSSI, pero no tanto como una presencia humana. Realizando este cambio

obtenemos los resultados de la figura 5.18.

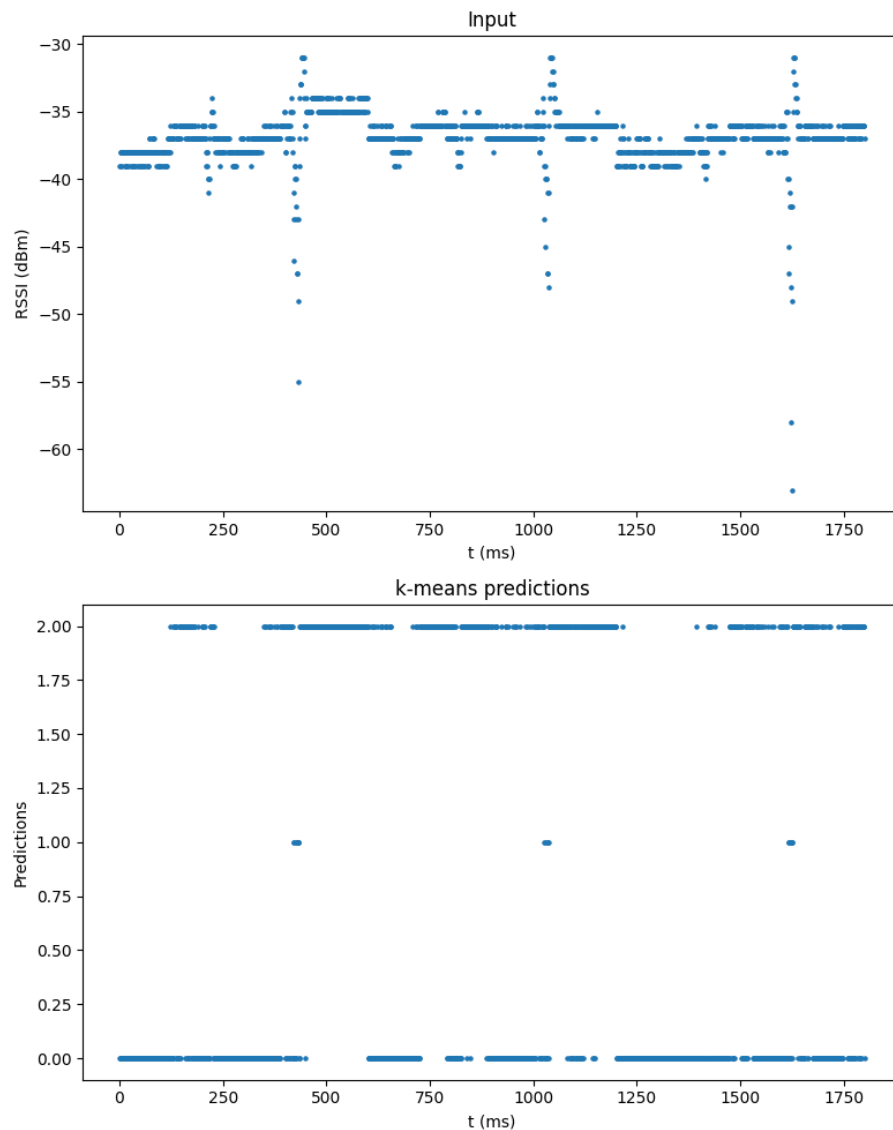


Figura 5.18: Predicciones para una silla de oficina con tres clústeres

Al configurar el algoritmo *k-means* con tres clústeres, se obtiene una segmentación más detallada de los datos y se asigna cada muestra a uno de los tres grupos mencionados. En este caso, podemos ver como se diferencian claramente con un 0 la ausencia de presencia, con un 1 la presencia humana, y con un 2 los falsos positivos y las pérdidas de potencia de RSSI causadas por interferencias, un falso positivo o por el entorno.

5.6. Conclusiones de las pruebas

En este capítulo se presentaron los detalles de las pruebas y resultados obtenidos en el proyecto de detección de personas mediante los cambios de la señal de radiofrecuencia. Se llevaron a cabo diferentes escenarios que simulaban las situaciones de un hogar u oficina para evaluar el rendimiento y la eficacia del algoritmo.

- En el primer escenario se simulaba un enlace donde los dispositivos tenían una línea de visión directa y una persona pasaba entre medias del enlace. El algoritmo demostró un alto nivel de precisión y una detección efectiva de la presencia.
- En el segundo escenario se simulaba un enlace donde los dispositivos tenían una línea de visión directa y una persona se quedaba en medio del enlace. En este caso, el algoritmo también logró detectar con precisión la presencia incluso cuando el intruso se quedaba quieto en el enlace.
- En el tercer escenario se simulaba un enlace donde los dispositivos no tenían una línea de visión directa y una persona pasaba entre medias del enlace. El algoritmo mostró un buen nivel de precisión, a pesar de los obstáculos físicos que afectaban a la potencia de la señal.
- En el cuarto escenario se simulaba un enlace donde los dispositivos no tenían una línea de visión directa y una persona se quedaba en medio del enlace. En este caso, el algoritmo logró realizar una detección de presencia con una precisión razonable, aunque existieron algunas dificultades para discriminar la presencia de la persona.

Por último, se realizó un análisis de los falsos positivos para identificar las limitaciones del algoritmo, y determinar el tamaño mínimo de los objetos detectables. Para estas pruebas se emplearon 4 objetos: una caja pequeña, otra mediana, otra grande y por último, una silla de oficina. Se analizó que el enlace no logra detectar las dos primeras cajas. Asimismo, la tercera caja si se puede detectar, pero el algoritmo no lo predice como una presencia humana. En cambio, el algoritmo predice una presencia humana con el paso de la silla de oficina. Una solución para esto es emplear tres clústeres para poder agrupar la presencia en uno, en el segundo la ausencia de esta y en el tercero juntar los falsos positivos o las pérdidas provocadas por el entorno.

En resumen, el algoritmo de detección de presencia y movimiento humano mostró un gran rendimiento en escenarios de línea de visión directa, con altos valores de precisión, ARS, FMS y *Silhouette Score*. Sin embargo, en escenarios sin línea de visión directa, se observaron resultados más variados, con precisiones y métricas menores. Aun así, en estos escenarios, el algoritmo logra realizar unas predicciones razonables. Finalmente, podemos concluir que ni el rendimiento ni las predicciones del sistema se ven afectados por objetos de dimensiones pequeñas, y empleando un número de clústeres mayor consigue diferenciar objetos de mayores dimensiones, sin que estos afecten a la detección de presencia y movimiento humano.

Capítulo 6

Conclusiones y líneas futuras

En este capítulo se presentan las conclusiones del Trabajo Fin de Grado. Se analizan los resultados obtenidos, así como las limitaciones y desafíos encontrados durante su desarrollo. Además, se proponen líneas futuras de investigación y desarrollo para mejorar y expandir este proyecto.

6.1. Conclusiones

El objetivo de este Trabajo Fin de Grado es realizar un análisis de las distintas posibilidades existentes para el cálculo de presencia y movimiento basado en cambios en la señal de radiofrecuencia. Específicamente, el enfoque del proyecto se ha centrado en el análisis de señales a una frecuencia de 868 MHz.

Este proyecto se ha realizado en la Cátedra de Securitas Direct Verisure de la Escuela Técnica Superior de Ingenieros de Telecomunicación. Dentro de este entorno académico, se ha contado con la supervisión y orientación de profesionales y docentes cualificados. Esto, ha proporcionado una oportunidad para unir los conocimientos teóricos y prácticos aprendidos durante el grado.

Las conclusiones obtenidas al realizar este proyecto son:

1. Primeramente se investigaron las diversas técnicas de detección de señales inalámbricas. Esto ha permitido analizar y evaluar las características, fortalezas y limitaciones de cada una de las técnicas. Las cuatro técnicas más importantes son RSSI, CSI, FMCW y *Doppler*. Finalmente, se optó por emplear la técnica RSSI o intensidad de la señal recibida en el proyecto, la cual permite evaluar el nivel de potencia de las señales recibidas por un dispositivo.
2. Posteriormente, se realizó una simulación de un enlace entre dos puntos para analizar los cambios en la señal de radiofrecuencia. Para ello, se emplearon dos placas *EFR32FG* de Silicon Labs junto con un *Flex Gecko* capaz de trabajar a una frecuencia de 868 MHz (ver figura 3.2).

3. A continuación, se realizó un análisis en profundidad de diferentes algoritmos de decisión no supervisados, con especial énfasis en los algoritmos de *clustering*. Dentro de los algoritmos de agrupamiento, destacan el algoritmo de *k-means* y el de *hierarchical clustering*. Analizando sus fortalezas y limitaciones, se optó por desarrollar el algoritmo de *k-means* debido a su eficiencia y su capacidad de trabajar con grandes conjuntos de datos.
4. Por último, se llevaron a cabo diferentes pruebas para evaluar el rendimiento del algoritmo de detección en escenarios realistas. Se emplearon métricas de precisión, ARS, FMS y *Silhouette Score*, para medir la eficacia de las detecciones. Se pudo observar que el algoritmo presentó un gran rendimiento en escenarios de línea de visión directa, en cambio en escenarios sin línea de visión directa se obtuvieron resultados más dispares, pero logrando realizar unas detecciones razonablemente acertadas. Además, ni el rendimiento ni las predicciones del sistema se ven afectados por objetos de dimensiones pequeñas evitando problemas de falsos positivos.

Finalmente, es importante tener en cuenta que los resultados obtenidos no deben considerarse como una verdad absoluta, ya que su eficacia puede variar según el entorno y la situación en la que se utilice. No es recomendable que el sistema diseñado funcione de manera independiente, es mejor que sea utilizado como un apoyo junto a otros dispositivos. En este caso, complementando este proyecto con otros dispositivos de seguridad de Securitas Direct, se logran unos resultados más fiables y completos, ya que se podrían aprovechar las fortalezas y capacidades de cada uno de ellos.

6.2. Líneas futuras

En esta sección se van a valorar y comentar futuras líneas de trabajo en este proyecto. Algunas de ellas podrían ser:

- Explorar el ahorro de energía. El ahorro de energía es un aspecto muy relevante en el diseño de sistemas electrónicos. Es crucial encontrar un balance entre cuantos paquetes se pueden enviar por segundo para detectar presencia y no consumir la batería del dispositivo en poco tiempo.
- Mejorar el rendimiento en escenarios sin línea de visión directa. Se puede investigar y desarrollar otras técnicas que no se basen solamente en RSSI para lograr un mejor rendimiento del sistema en los entornos cotidianos.
- Desarrollar más algoritmos que permitan mejorar la detección de presencia y movimiento humano. Cuantas más predicciones se realicen de las muestras, se logrará un mayor rendimiento.
- Explorar la detección de objetos de diferentes tamaños y formas. Ampliar el alcance del sistema de detección, para que pueda identificar y distinguir objetos de diversas dimensiones y formas, permitirá una detección más precisa en entornos reales.

- Implementar técnicas de seguimiento y reconocimiento de patrones. Esta investigación permitiría mejorar la identificación del movimiento, así como reconocer las acciones específicas realizadas por el individuo en cada momento.
- Implementar un sistema de detección de presencia y movimiento humano que permita cubrir un área más amplia en lugar de depender únicamente de un enlace punto a punto. Esta estrategia de triangulación podría mejorar la precisión y la cobertura del sistema, brindando una detección más completa y confiable.

Referencias

- [1] Mariana Molina Matute. Estudio de banda de frecuencias sub-ghz para redes de sensores inalámbricas e implementación en plataforma modular. Marzo 2011.
- [2] Lingchao Guo, Xiangming Wen, Zhaoming Lu, Xinbin Shen, and Zijun Han. Wiroi: Spatial region of interest human sensing with commodity wifi. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2019.
- [3] Jian Liu, Hongbo Liu, Yingying Chen, Yan Wang, and Chen Wang. Wireless sensing for human activity: A survey. *IEEE Communications Surveys Tutorials*, 22(3):1629–1645, 2020.
- [4] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. Non-invasive detection of moving and stationary human with wifi. *IEEE Journal on Selected Areas in Communications*, 33(11):2329–2342, 2015.
- [5] Yaoxuan Yuan, Jizhong Zhao, Chen Qiu, and Wei Xi. Estimating crowd density in an rf-based dynamic environment. *IEEE Sensors Journal*, 13(10):3837–3845, 2013.
- [6] Fei Wang, Jianwei Feng, Yinliang Zhao, Xiaobin Zhang, Shiyuan Zhang, and Jinsong Han. Joint activity recognition and indoor localization with wifi fingerprints. *IEEE Access*, 7:80058–80068, 2019.
- [7] Shinya Fukushima, Hiroyoshi Yamada, Hirokazu Kobayashi, and Yoshio Yamaguchi. Human motion detection and extraction by using fm-cw range-doppler radar. In *2014 International Symposium on Antennas and Propagation Conference Proceedings*, pages 173–174, 2014.
- [8] Krohne. Medida de nivel de radar 80 ghz.
- [9] Securitas Direct. Nuestra tecnología de seguridad, securitas direct verisure.
- [10] DigiKey. Placas de desarrollo, kits, programadores kits de desarrollo y evaluación de rf, placas.
- [11] Silicon Labs. Developing proprietary connect applications.
- [12] Mabuthi. Setting up range test application on efr32. *Silicon Labs*, 2016.
- [13] Silicon Labs. Flex sdk v3.x range test demouser’s guide.
- [14] Wikipedia. Algoritmos.
- [15] Juan Francisco Vallalta Rueda. Aprendizaje supervisado y no supervisado.
- [16] Inna Kaler Fan Zhang. So you have some clusters, now what?
- [17] Vihar Kurama. Unsupervised learning with python: A beginner’s guide. *Built in*, 2022.

-
- [18] Na8. K-means en python paso a paso.
 - [19] Callum Ballard. How to evaluate unsupervised learning models. *Towards sata science*, 2020.
 - [20] Ana Muñoz Vita. El ingeniero de telecomunicación, un perfil que escasea. *Cinco Días; El Pais*, 2017.
 - [21] Glassdoor. Sueldos para el puesto de ingeniero de telecomunicaciones en españa.
 - [22] Sujeeva Kumaratunga. Kmeans hyper-parameters explained with examples. *Towards Dara Science*, 2020.

Apéndice A

Aspectos éticos, económicos, sociales y ambientales

Este apéndice tiene como objetivo estudiar los aspectos éticos, económicos, sociales y ambientales de este Trabajo Fin de Grado. Al analizar estos aspectos, se busca conseguir una gestión responsable y sostenible del proyecto, generando un impacto positivo en el entorno y en la sociedad.

Cabe recordar que el objetivo de este proyecto es analizar las distintas posibilidades para la detección de presencia y movimiento basado en cambios en la señal de radio, haciendo uso de diferentes algoritmos de *machine learning* para poder inferir resultados.

A.1. Impacto ético

El proyecto plantea importantes cuestiones éticas relacionadas con la privacidad y la protección de datos de las personas. La recolección y análisis de datos e información puede implicar la recopilación de datos personales, privados o sensibles. Garantizar la confidencialidad y el consentimiento de las personas humanas involucradas es uno de los desafíos más importantes y fundamentales de este proyecto.

A.2. Impacto económico

En cuanto a los impactos económicos, se puede pensar que el desarrollo de algoritmos de *machine learning* y la obtención de los equipos y herramientas necesarias pueden significar grandes gastos. Sin embargo, este proyecto está pensado para implementarse en equipos sencillos, lo cual reduce significativamente los costes. Además, al reemplazar a los sensores de movimiento, sensores basados en visión, sensores acústicos y sensores infrarrojos piroeléctricos (PIR), que tienen un coste elevado, se logra una solución más rentable.

A.3. Impacto social

Uno de los mayores impactos de este Trabajo Fin de Grado es el social. Este proyecto tiene el potencial de mejorar la seguridad y la eficiencia en hogares y oficinas. La detección de presencia y movimientos de intrusos contribuye a la prevención de robos y monitorización de interiores. Esto mejora la seguridad de las personas, aumentando así su calidad de vida.

A.4. Impacto ambiental

Por último, en términos del impacto ambiental, este proyecto puede tener efectos positivos al reemplazar a numerosos equipos y dispositivos de detección. Al detectar una presencia de un intruso, se activará el dispositivo, lo que conlleva grandes ahorros energéticos y una reducción de la huella de carbono. Sin embargo, también hay que recalcar el ciclo de vida de los dispositivos y su impacto en términos de consumo y generación de residuos electrónicos.

A.5. Conclusiones

En resumen, este Trabajo Fin de Grado de detección de presencia, presenta impactos éticos en la protección de datos y privacidad. Desde el punto de vista económico, se destaca su implementación sencilla y rentable, y su capacidad de reemplazar a los sensores convencionales. A nivel social, mejora la seguridad y calidad de vida de los usuarios. Por último, a nivel ambiental cabe destacar los ahorros energéticos al activarse solo en momentos específicos.

Apéndice B

Presupuesto económico

Este apéndice del Trabajo Fin de Grado tiene como objetivo estudiar el aspecto económico del proyecto. El presupuesto económico es una herramienta para el cálculo, planificación y ejecución de un proyecto, que permite estimar y controlar los recursos necesarios para llevarlo a cabo de manera eficiente y efectiva.

- **Personal:** En la tabla B.1 se representa la asignación de los costes relacionados con el personal que se ha visto involucrado en el proyecto. Para determinarlo, se ha considerado el salario medio bruto de los becarios que estudian ingeniería de telecomunicación, el cual es de 12.000 euros anuales, equivalente a 6 euros por hora [20]. Asimismo, el salario medio bruto de los ingenieros de telecomunicación es de 30.000 euros anuales, equivalente a 15 euros por hora [21].

	Coste horario (€)	Horas	Total (€)
Director del trabajo	15	30	450
Estudiante de ingeniería	6	324	1944
TOTAL			2394

Tabla B.1: Costes de personal

- **Costes de recursos materiales:** En la tabla B.2, se presenta la asignación de costos relacionados con los recursos materiales utilizados en el proyecto. Para determinarlos, se ha realizado un análisis de los precios de cada material utilizado, teniendo en cuenta el tiempo de vida, unidades, coste, amortización y su uso.

	Vida útil (años)	Uds.	Coste (€)	Amortización (€/mes)	Uso (meses)	Total (€)
Ordenador	4	1	1300	27,08	9	243,72
Placas <i>EFR32FG</i>		2	126,24	0	0	252,48
Pilas de botón		1	6,99	0	0	6,99
TOTAL						503,19

Tabla B.2: Costes de recursos materiales

Para finalizar, el coste del proyecto incluyendo costes de personal y material se ve reflejado en la tabla B.3. Se estima un coste indirecto del 10 % de los costes directos, y un beneficio industrial del 5 % sobre ambos costes.

	Coste
Costes de personal	2394 €
Costes de material	503,19 €
Costes directos	2897,19 €
Costes indirectos	289,72 €
Costes directos + indirectos	3186,91 €
Beneficio industrial	159,35 €
Subtotal	3346,26 €
IVA (21 %)	702,72 €
Total	4048,98 €

Tabla B.3: Costes totales

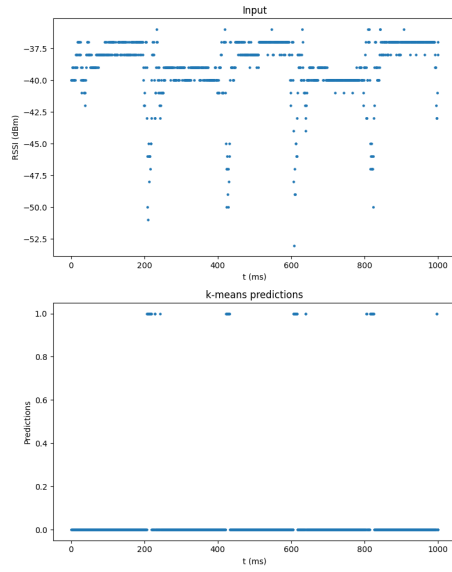
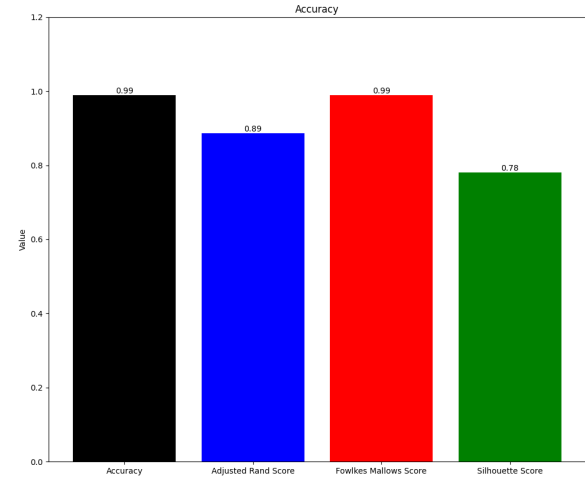
Apéndice C

Análisis de parámetros del algoritmo *k-means*

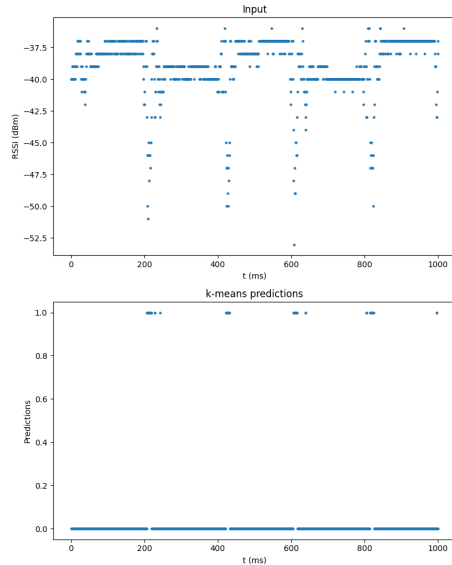
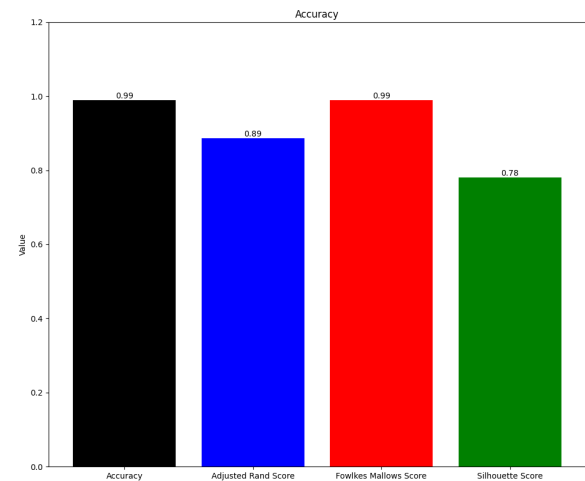
Este anexo tiene como objetivo analizar los parámetros configurables del algoritmo *k-means* estudiado en el capítulo 4.

La biblioteca empleada para implementar el algoritmo de *k-means* es *scikit-learn*. Dentro de esta biblioteca, se puede aprovechar la implementación de *k-means* proporcionada por la clase *KMeans* del módulo *sklearn.cluster*. Esta clase ofrece una serie de parámetros que permite personalizar y ajustar el comportamiento del algoritmo para optimizarlo al máximo y adaptarlo a las necesidades específicas. Los parámetros configurables del algoritmo *k-means* son [22]:

- *n_clusters*: este parámetro define el número de grupos que se van a emplear. Este parámetro se tiene que inicializar obligatoriamente con el número de grupos que queremos. Como se ha mencionado previamente, para este proyecto idealmente el valor será de 2, uno que representa la presencia humana y otro que representa la ausencia de esta.
- *init*: este parámetro especifica el método de inicialización de los centroides de los clústeres. Existen 2 opciones posibles:
 - *k-means++*: es un método para seleccionar los centroides iniciales de forma más eficiente. Se utiliza como valor por defecto. Este enfoque tiende a producir mejores resultados y una convergencia más rápida en comparación con la inicialización aleatoria.

(a) Predicción del algoritmo *k-means*(b) Precisión del algoritmo *k-means*Figura C.1: Configuración de parámetro $init = k - means ++$

- *random*: inicializa los centroides de manera aleatoria. En este caso, los resultados pueden variar en cada ejecución del algoritmo, ya que la inicialización aleatoria puede llevar a diferentes configuraciones iniciales de los centroides.

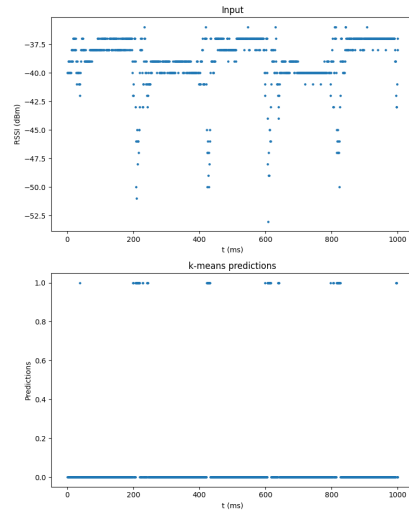
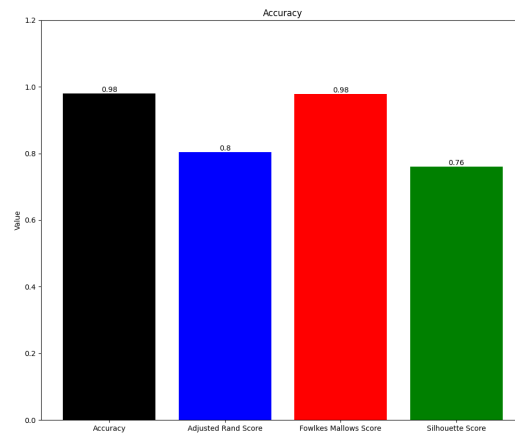
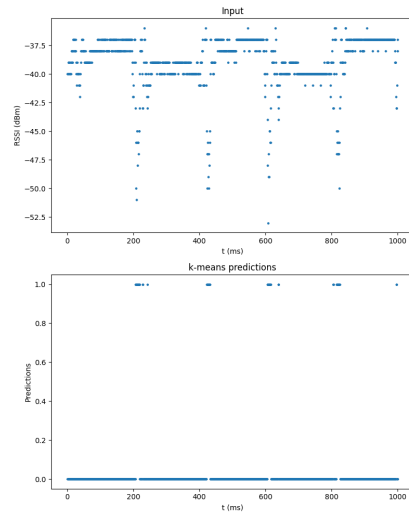
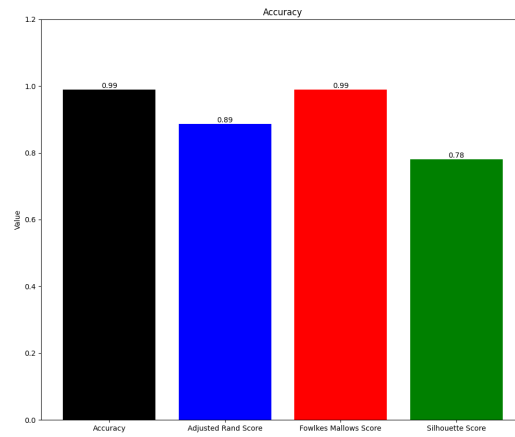
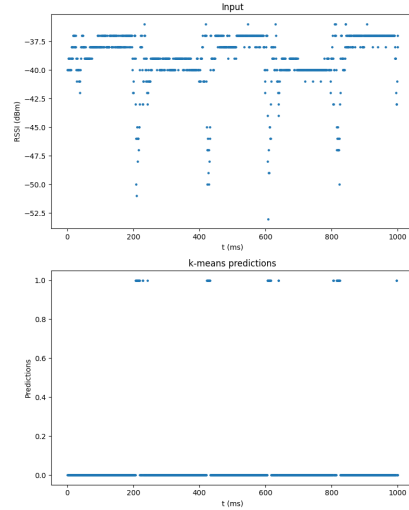
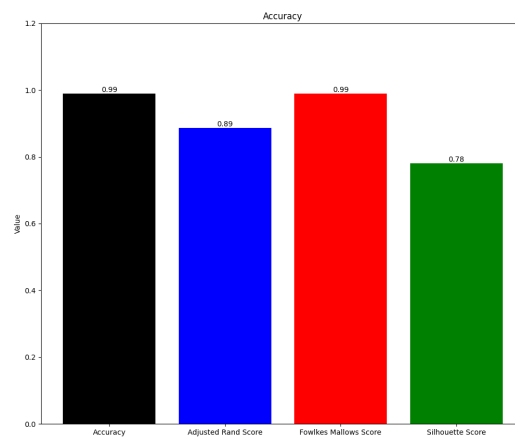
(a) Predicción del algoritmo *k-means*(b) Precisión del algoritmo *k-means*Figura C.2: Configuración de parámetro $init = random$

Como se ve en las figuras C.1 y C.2, la variación de este parámetro no tiene gran

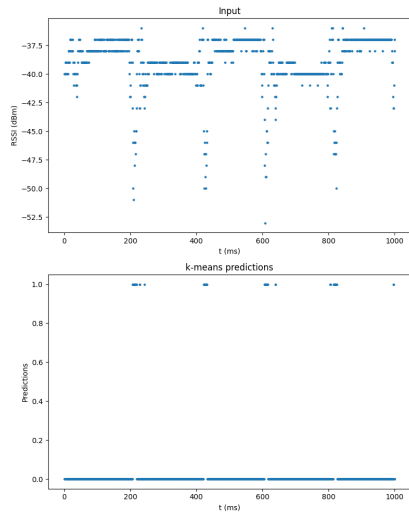
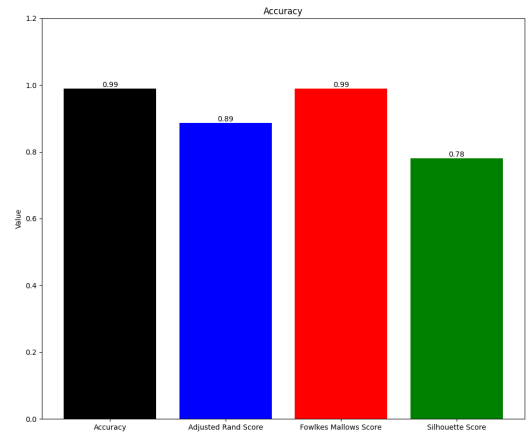
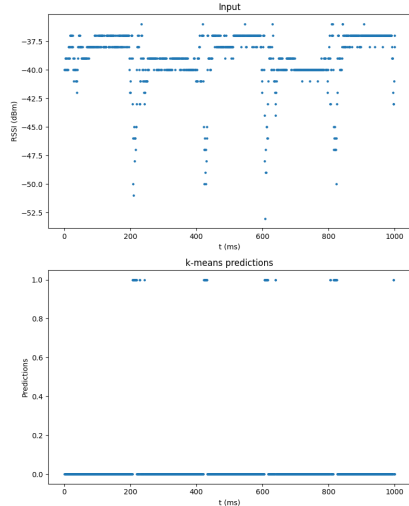
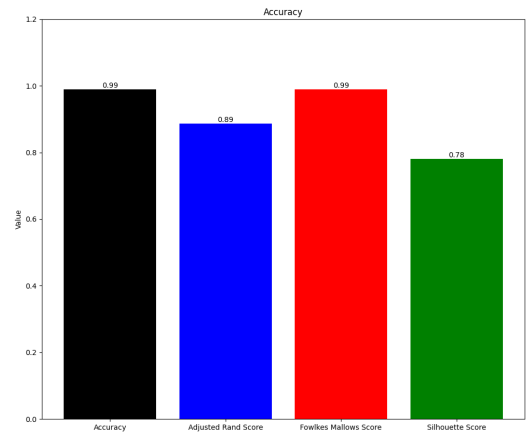
impacto si empleamos solo 2 o 3 centroides, por lo tanto, se dejará este valor por defecto, es decir *k-means++*.

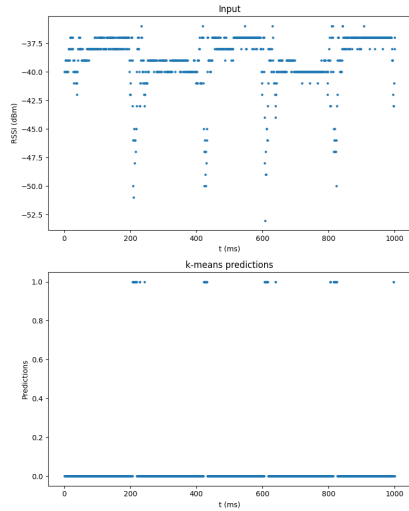
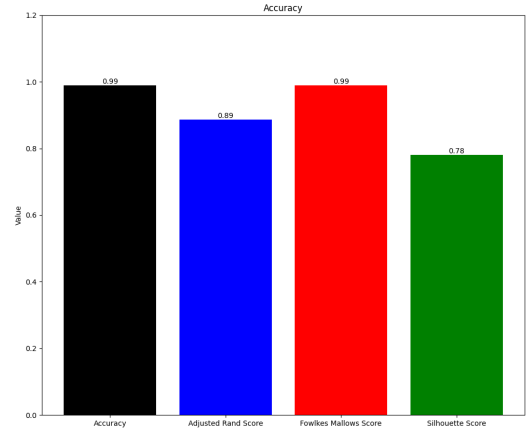
- *n_init*: este parámetro indica el número de veces que se ejecutará el algoritmo con diferentes centroides iniciales. Al finalizar, se devuelve el algoritmo con mejores resultados según la inercia.
 - Con un valor de *n_init* = 1, solo se realizará una única inicialización aleatoria. Es útil si se quiere una ejecución rápida del algoritmo o el conjunto de muestras es pequeño.
 - Con un valor de *n_init* = 10, que es el valor por defecto, se realizarán 10 inicializaciones aleatorias y seleccionará la mejor solución según sus características.
 - Con un valor de *n_init* = 20, existen más posibilidades de obtener un resultado óptimo al aumentar el valor de *n_init*, aunque el tiempo de ejecución será mayor.

Como se ve en las figuras C.3, la variación de este parámetro no tiene ningún impacto; por lo tanto, se dejará este valor por defecto, es decir 10 ejecuciones con diferentes centroides iniciales.

(a) Predicción con $n_init = 1$ (b) Precisión con $n_init = 1$ (c) Predicción con $n_init = 10$ (d) Precisión con $n_init = 10$ (e) Predicción con $n_init = 20$ (f) Precisión con $n_init = 20$ Figura C.3: Configuración de parámetro n_init

- *max_iter*: este parámetro especifica el número máximo de iteraciones que se realizarán en cada ejecución del algoritmo.
 - Con un valor de *max_iter* = 100, se tiene un menor número de iteraciones y evaluar rápidamente el comportamiento del algoritmo gracias a una convergencia más rápida.
 - Con un valor de *max_iter* = 300, que es el valor por defecto, se tiene una cantidad aceptable de iteraciones para evaluar el comportamiento del algoritmo.
 - Con un valor de *max_iter* = 1000, garantiza explorar todo el espacio de características, pero se reduce el rendimiento.

(a) Predicción con *max_iter* = 100(b) Precisión con *max_iter* = 100(c) Predicción con *max_iter* = 300(d) Precisión con *max_iter* = 300

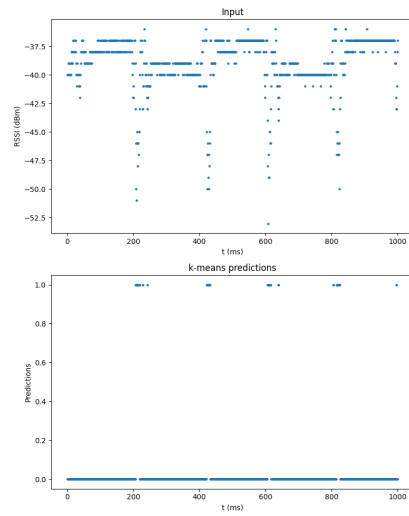
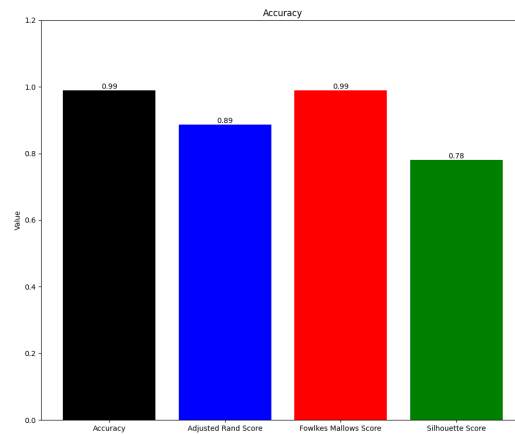
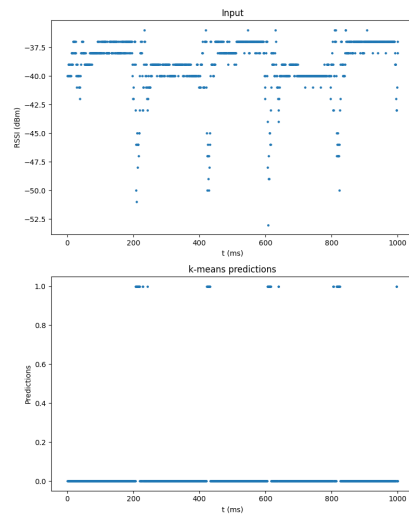
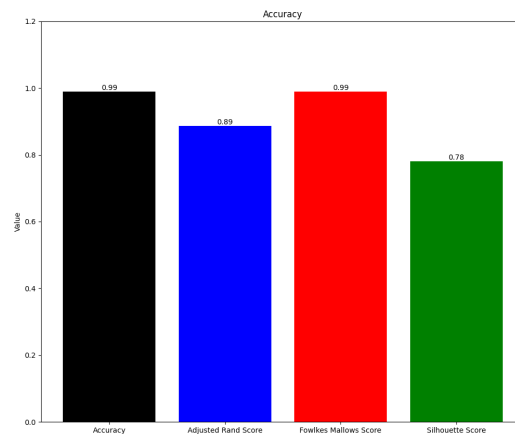
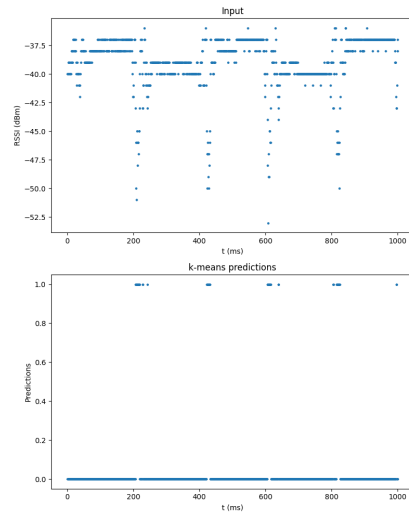
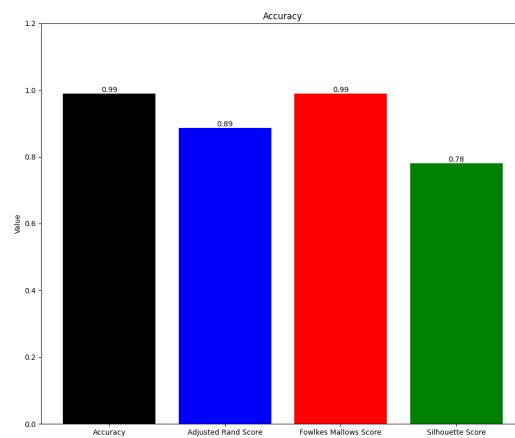
(e) Predicción con $max_iter = 1000$ (f) Precisión con $max_iter = 1000$ Figura C.4: Configuración de parámetro max_iter

En este proyecto, se busca encontrar un equilibrio entre el número de iteraciones necesarias para evaluar adecuadamente el algoritmo y obtener una respuesta rápida. Para lograrlo, se ha seleccionado un valor de $max_iter = 300$. Como se puede observar en la figura C.4, con este valor se obtienen los resultados deseados.

- tol : este parámetro hace referencia a la tolerancia relativa respecto a la inercia. El algoritmo convergerá cuando la diferencia entre las inercias de dos iteraciones consecutivas sea menor o igual al valor de tol .
 - Con un valor de $tol = 10^{-2}$, se permite una mayor inercia lo que provoca una convergencia más rápida y menor tiempo de ejecución, pero se pierde precisión en los resultados.
 - Con un valor de $tol = 10^{-4}$, que es el valor predeterminado, se logra tener una inercia aceptable para evaluar el comportamiento del algoritmo, manteniendo la precisión de los resultados.
 - Con un valor de $tol = 10^{-6}$, se obtiene una convergencia más precisa pero un valor de ejecución elevado.

Como se ve en la figura C.5, aumentar este valor significa que se tolera un cambio mayor en la inercia antes de converger. Es mejor mantenerlo en un valor bajo para que el algoritmo realice todas las iteraciones hasta converger. Para este proyecto, se mantendrá en 10^{-4} .

- Los parámetros restantes son: $precompute_distances$, $verbose$, $random_state$, $copy_x$, n_jobs y $algorithm$. Estos parámetros afectan a la eficiencia de computación, y como se va a tener un conjunto grande de muestras, lo mejor es mantener estos parámetros con los valores predeterminados.

(a) Predicción con $tol = 10^{-2}$ (b) Precisión con $tol = 10^{-2}$ (c) Predicción con $tol = 10^{-4}$ (d) Precisión con $tol = 10^{-4}$ (e) Predicción con $tol = 10^{-6}$ (f) Precisión con $tol = 10^{-6}$ Figura C.5: Configuración de parámetro tol

