

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA



Profesor Patrocinante:
Dr. Sergio Sobarzo Guzmán.

Informe de Memoria de Título
para optar al título de:
Ingeniero Civil en Telecomunicaciones

**Sistema de posicionamiento indoor mediante
seguimiento de direcciones MAC de equipos
móviles**

Universidad de Concepción
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica

Profesor Patrocinante:
Dr. Sergio Sobarzo Guzmán.

SISTEMA DE POSICIONAMIENTO INDOOR MEDIANTE SEGUIMIENTO DE DIRECCIONES MAC DE EQUIPOS MÓVILES

Aldo Nicolás Mellado Opazo

Informe de Memoria de Título
para optar al Título de

Ingeniero Civil en Telecomunicaciones

enero 2020

Índice General

Índice de Figuras	III
Índice de Tablas	v
1. Introducción	1
1.1. Antecedentes históricos	1
1.2. Definición del problema	1
1.3. Estado del arte	2
1.4. Hipótesis de trabajo	9
1.5. Objetivos	9
1.5.1. Objetivo general	9
1.5.2. Objetivos específicos	9
1.6. Alcances y limitaciones	9
1.7. Metodología	10
2. Marco Teórico	11
2.1. Introducción	11
2.2. Aplicaciones y aproximaciones en sistemas de posicionamiento indoor	11
2.2.1. Análisis de Escena	11
2.2.2. Detección de proximidad	12
2.2.3. Triangulación	12
2.2.3.1. Angulación	12
2.2.3.2. Lateración	15
3. Diseño de algoritmo	26
3.1. Introducción	26
3.2. Diseño general de trabajo	26
3.2.1. Offline	27
3.2.1.1. Medición	27
3.2.1.2. Extracción y procesamiento de datos	27
3.2.1.3. Diseño de Set de Entrenamiento	27
3.2.1.4. Entrenamiento y Optimización	27
3.2.2. Online	28

3.2.2.1. ANN para estimar la posición del objetivo.	28
4. Implementación	29
4.1. Introducción	29
4.2. Implementación general de trabajo	29
4.2.1. Fase Offline	30
4.2.1.1. Adquisición de RSSI	30
4.2.1.2. Set de entrenamiento	32
4.2.2. Red Neuronal	33
5. Resultados	35
5.1. Introducción	35
5.1.1. Etapa Offline	35
5.1.1.1. Medición	35
5.1.2. Laboratorio dividido en sectores.	37
5.2. Entrenamiento	37
5.2.1. Ajustes entrenamiento	37
5.3. Predicción	37
6. Conclusiones	38
6.1. Sumario	38
6.2. Conclusiones	38
6.3. Trabajo Futuro	39
7. Códigos	40
7.1. Derivación de fórmulas	40
7.2. Unidades especiales	40

Índice de Figuras

1.1. Imagen de la comparación de la cercanía según el nivel de RSS según un sistema basado en la interacción de beacons BLE	4
2.1. Algoritmo de Triangulación	13
2.3. Técnica de Ángulo de Llegada (AoA)	14
2.4. Técnica de Tiempo de Llegada (ToA)	15
2.5. Técnica de Diferencia de Tiempo de Llegada (TDoA)	16
2.6. Diagrama de medición de distancia usando Ruta de Viaje de Vuelo (del inglés <i>Round-Trip Time of Flight</i> , RToF)	17
2.7. Idea de algoritmo basado en RSSI	18
2.8. Distintos hiperplanos para el mismo set de datos.	19
2.9. Representación gráfica del funcionamiento de KNN según el valor de K	21
2.10. Comparación por similitud de una neurona con un perceptrón	22
2.11. Representación de un Perceptrón	22
2.12. Función Continua no lineal Sigmoide	23
3.1. Diagrama General de operación de Algoritmo de posicionamiento	26
4.1. Esquema de Implementación General de Trabajo	29
4.2. Salida de consola con datos de AP	30
4.3. Mapa Lab. de Redes de datos	31
4.4. Formato tipo de datos entregados	32

4.5. Diagrama de funcionamiento offline	33
5.1. Archivo antes y después de filtros sobre datos obtenidos	36

Índice de Tablas

Siglas

AoA Ángulo de llegada (del inglés *Angle of Arrival*)

AP Punto de Acceso (del inglés *Access Point*)

BLE Bluetooth de Baja Energía (del inglés *Bluetooth Low Energy*)

ESSID Identificador de Set de Servicio Extendido (del inglés *Extended Service Set Identifier*)

FCC Comisión Federal de Comunicaciones (del inglés *Federal Communications Commission*)

GPS Sistema de Posicionamiento Global (del inglés *Global Positioning System*)

IPS Sistema de Posicionamiento en Interiores (del inglés *Indoor Positioning System*)

KNN K-Vecinos más Cercanos (del inglés *K-Nearest Neighbour*)

MAC Control de Acceso al Medio (del inglés *Media Access Control*)

NAVSAT Sistema Satelital de Navegación Naval (del inglés *Navy Navigation Satellite System*)

NLOS No Línea Vista (del inglés *None Line of Sight*)

ANN Red Neuronal Artificial (del inglés *Artificial Neural Network*)

RNN Red Neuronal Recurrente (del inglés *Recurrent Neural Network*)

RSS Nivel de Intensidad de Señal Recibida (del inglés *Received Signal Strenght*)

RSSI Indicador de Intensidad de Señal Recibida (del inglés *Received Signal Strength Indicator*)

RToF Ruta de Viaje de Vuelo (del inglés *Round-Trip Time of Flight*)

RTT Tiempo-de-Viaje (del inglés *Round-Trip-Time*)

SNR Relación Señal a Ruido (del inglés *Signal Noise Rate*)

SVP Máquinas de vectores de soporte (del inglés *Support Vector Machine*)

TDoA Diferencia de Tiempo de Llegada (del inglés *Time Difference of Arrival*)

ToA Tiempo de llegada (del inglés *Time of Arrival*)

ToF Tiempo de Vuelo (del inglés *Time of Flight*)

UWB Banda Ultra Ancha (del inglés *Ultra Wide-Band*)

WLAN Red de área local inalámbrica (del inglés *Wide Local Area Network*)

1. Introducción

1.1. Antecedentes históricos

El ser humano, en la década del 60, ante la incipiente necesidad de saber su posición en el planeta desarrolló un sistema de posicionamiento llamado OMEGA y posteriormente otro llamado TRANSIT o Sistema Satelital de Navegación Naval (del inglés *Navy Navigation Satellite System*, NAVSAT), que fue resultado del trabajo conjunto de la NASA y el departamento de defensa de los Estados Unidos. Años después este acabó siendo reemplazado debido a la falta de precisión que este tenía, y que alcanzaba un error de hasta 250 metros.

Su sucesor apareció en la década del 70 bajo el nombre de Sistema de Posicionamiento Global (del inglés *Global Positioning System*, GPS) y su precisión permitía posicionar un objeto con un error de menos de 5 metros. Esto a través del cálculo del tiempo que tarda en llegar la señal al receptor, es decir, el efecto Doppler.

Funciona actualmente con un mínimo de 24 satélites en órbita sobre la tierra cuyas trayectorias sincronizadas le permiten mapear completamente el planeta y entregar posicionamiento casi exacto a dispositivos móviles y vehículos.

Existen desde ya hace décadas intentos por emular lo logrado con el GPS, pero para espacios interiores, estos intentos han recibido el nombre de Sistema de Posicionamiento en Interiores (del inglés *Indoor Positioning System*, IPS), que conforme a lo que se presentará en capítulos posteriores han logrado obtener resultados de posición con márgenes de error inferiores a los 3 m.

1.2. Definición del problema

Las condiciones bajo las cuales es posible para la señal propagarse no se cumplen en todos los ambientes. Existen lugares en que a diferencia de lo que sucede en el exterior, donde la señal se refleja haciendo posible la triangulación de la posición, esta se absorbe parcial o completamente y principalmente corresponden a espacios interiores, tales como una bodega, un centro comercial o

una oficina, haciendo que posicionarse dentro de estos espacios sea imposible a través del GPS, es por esto que a fin de brindar nuevas experiencias a usuarios a través del posicionamiento dentro de estos espacios se han desarrollado soluciones utilizando la banda de los 2.4 [GHz] que es la utilizada por, entre otros tecnologías, el Wi-Fi.

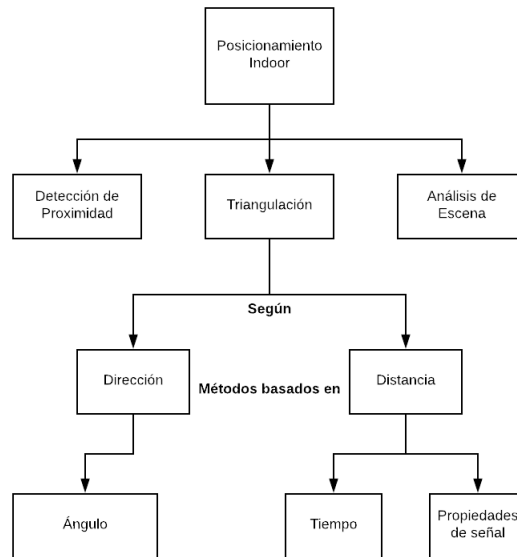
Esta ha sido ampliamente estudiada debido a la alta penetración comercial que ha alcanzado precisamente en estos espacios donde el GPS no da cobertura, sin embargo, dentro de los distintos enfoques en que se han abordado los estudios se tiene que la medición de los niveles de potencia radiada desde los Punto de Acceso (del inglés *Access Point*, AP), a través de los cuales se realiza la triangulación de la posición, se ven altamente afectados por la variación del escenario caracterizado. Este tipo de variaciones pueden ser inducidas por la presencia de personas u objetos que reflejen o absorban la señal.

Es por esto que el método a través del cual se de solución al problema del posicionamiento en interiores debe ser un IPS capaz de compensar la incidencia de estas variaciones en el indicador de intensidad de señal recibida (RSSI) dentro de la operación del algoritmo y así, estimar correctamente la posición de objetivo deseado identificable a través de su dirección de Control de Acceso al Medio (del inglés *Media Access Control*, MAC).

1.3. Estado del arte

En cuánto a los enfoques que se han considerado para evaluar la posición de un dispositivo móvil en un espacio interior se han podido desarrollar diversas soluciones que apuntan de una u otra manera a aproximaciones geométricas. A fin de ordenar según la técnica, la documentación consultada, se hará una clasificación y dentro de las categorías, explicará las investigaciones en el área.

Se abordará siguiendo el esquema que está a continuación:



Así, comenzando por el tipo de algoritmos basados en *Detección de proximidad* se tiene que de acuerdo a lo visto en [6], algunas de las tecnologías que permiten esto, es aquella asociada al uso de beacons; pequeños dispositivos capaces de emitir señales de onda corta por medio de la tecnología Bluetooth, y que puede llegar hasta 50 metros de alcance. Es precisamente en el uso de estos dispositivos que en [20], los autores, con el fin de integrarlos a tecnologías del IoT, hacen todo un estudio respecto de los distintos tipos de beacons que existen en el mercado, haciendo énfasis en las características asociadas al consumo de energía, al uso del ancho de banda, y de la latencia asociada a la respuesta de los beacons del tipo Bluetooth de Baja Energía (del inglés *Bluetooth Low Energy*, BLE).

Presentan además como es que se han integrado estos dispositivos en el comercio, museos, hospitales y estaciones de trabajo, permitiendo así, el surgimiento potencial de una industria interesada en sus bondades. Explican allí que el uso de los beacons se traduce en que cada dispositivo posee una etiqueta única, que para efectos de este trabajo bien puede ser la dirección MAC, que al comunicarse con el dispositivo móvil permite a este, enviar una respuesta de cuáles dispositivos tiene a la vista y entonces, comparar el Nivel de Intensidad de Señal Recibida (del inglés *Received Signal Strength*, RSS) para así estimar respecto de cuál beacon está más cercano el usuario.

Sin embargo, pese a las ventajas asociadas el uso de estos dispositivos, se tiene que así como señalan en [1], la principal desventaja de algoritmos basados en tecnologías por proximidad, es

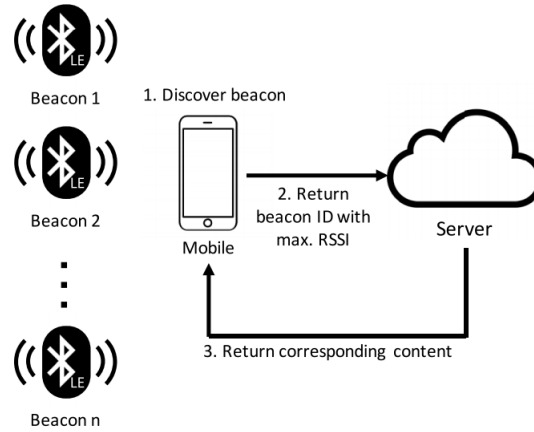


Fig. 1.1: Imagen de la comparación de la cercanía según el nivel de RSS según un sistema basado en la interacción de beacons BLE

que generalmente se requiere de hardware adicional, en lugar de los dispositivos que ya se usan e instalan para proveer de internet a usuarios. Además, hay que tener en cuenta que el hardware para este tipo de implementaciones hace uso de una parte no licenciada del espectro, por lo que es altamente susceptible a ser afectada por interferencias de equipos funcionando en la misma banda.

Luego, en el tipo de algoritmo basados en *Análisis de Escena* los autores en [1], presentan un sistema de posicionamiento indoor para navegación que puede funcionar indistintamente si se trata de un smartphone o de una tablet, en tanto se puedan usar los sensores que ya se incluyen en el funcionamiento interno de este. Para el reconocimiento de escena conciben el sistema como uno de dos partes [8]; una offline, que consiste en la extracción de características del entorno y un posterior almacenamiento en una base de datos que servirá para comparar las medidas que se obtengan en la etapa online. Durante esta etapa, la información recolectada es usada para obtener la posición estimada.

Las dificultades principales a las que se ven enfrentados estos métodos son los movimientos estructurales a las cuales los espacios se ven sometidos, ya que esto elimina puntos de referencia por simples movimientos de un lugar a otro, lo que conlleva una calibración y medición periódica.[20]

Finalmente, en lo que respecta al algoritmo de *Triangulación* se deben hacer algunas subclasificaciones a fin de dejar más claro cómo se ha abordado el problema, de esta manera se tiene que para los métodos basados en:

1. Triangulación:

- **Ángulo de llegada (del inglés *Angle of Arrival*, AoA):** Referido, como su nombre lo dice, se refiere al ángulo de llegada de la señal del dispositivo móvil, proveniente de una ubicación desconocida, la cual es recibida en múltiples estaciones base. Este algoritmo es evaluado por diversos autores; en [1] los autores lo trabajan como un complemento al uso de los sensores propios del dispositivo, logrando una performance destacable en cuanto a la media del error, sin embargo, no presentan resultados numéricos que permitan conocer el área de cobertura, así como la disponibilidad.

En [20], señalan que una de las desventajas de este método radica en que debido a que no todos los dispositivos están en línea vista con el equipo receptor, la precisión de las mediciones se degrada, haciendo hincapié en que la disposición del arreglo de antenas emisoras juega un rol fundamental en estimar la posición.

En [21] comienzan señalando que si bien existen otros algoritmos tales como ToF o TDoF, mucho menos complejos que el de AoA, si es posible implementar un método que alcance una buena precisión pese a las imprecisiones en las mediciones de ángulo y el bajo número de beacons dispuestos. No obstante, en [5] los autores hablan de aproximaciones híbridos que trabajan con lo mejor de los algoritmos basados en tiempo, y en ángulo para reducir la complejidad que introduce el ambiente indoor. Sin embargo, no logran reducir ni el número de antenas ni el ancho de banda utilizado.

2. Trilateración: Este algoritmo puede ser abordado a través de distintas técnicas, algunas de ellas son:

- **Tiempo de Vuelo (del inglés *Time of Flight*, ToF):** este concepto está referido a la técnica a través de la cual se realizan cálculos de posición a través del tiempo de llegada de la señal respecto de diversos puntos de referencia.[5]

Algunas de las dificultades que exponen los autores en [22] es que la medición de la señal basada en el tiempo de propagación es que requiere señales de Banda Ultra Ancha (del inglés *Ultra Wide-Band*, UWB), lo que lleva a un consumo de energía considerable, que no ofrece mayor precisión en largas distancias. Muy por el contrario, la precisión de la medición es directamente proporcional a la distancia, pudiendo alcanzar detalle de centímetros cuando la distancia es corta, pero fallas considerables si es que la señal ultrasónica debe viajar sufriendo las atenuaciones del aire así como del efecto multitrayectoria.

No obstante, en el sistema *TWINS* de [2] los autores lograron montar un sistema

basado en este algoritmo donde utilizan un rango de dos vías, aprovechando el intercambio de tráfico de DATA/ACK, que luego de sortear problemas asociados al protocolo de comunicación, esto es: ajustar la frecuencia de trabajo, sincronizar los equipos para que funcionen en un mismo tiempo, les permite posicionar usuarios que pueden moverse libremente por el espacio delimitado por el sistema.

- **Indicador de Intensidad de Señal Recibida (del inglés *Received Signal Strength Indicator*, RSSI):** A través de esta técnica los autores en [3] y [4] relacionan los niveles de RSS desde los AP existentes con la distancia que hay a un AP particular. A través de los modelos de propagación pueden hacer una regresión que, cuando se ha calculado la distancia que hay a los, como mínimo 3, AP se puede entonces crear un sistema de ecuaciones cuyas soluciones son las coordenadas espaciales del objetivo.

Otra técnica, algo menos acabada en comparación a lo visto anteriormente, es aquella vista en [13] que hace uso de los niveles de potencia recibida de cada uno de los AP para caracterizar un espacio interior y así, hacer un sistema de coordenadas que permita estimar la posición del usuario. Sin embargo, conocidas las propiedades de los espacios indoor, es sabido que estos son muy dinámicos, que cualquier superficie puede reflejar y absorber potencia, haciendo así variar estos niveles e inducir un margen de error que posicione erróneamente al usuario.

Un algoritmo basado en RSSI, pero que no usa propiedades geométricas para hacer el posicionamiento es aquel llamado *Fingerprinting Based Indoor Localization*, en [6] los autores lo presentan como una técnica que consiste en dos etapas; la primera, conocida como *offline*, es la de hacer una colección de los niveles de potencia recibidos en las distintas áreas de interés y almacenarlas en una base de datos que sirva de entrenamiento para las predicciones en las cuales se basará la etapa siguiente; luego, en la etapa *online*, lo que se busca es medir los niveles de potencia y compararlos con el vector de RSS que más se le parezca.

Para esto, la etapa *offline* puede verse apoyado por herramientas estadísticas como las de los trabajos [8], [10],[11],[12] y [14], o bien, por algoritmos basados técnicas de reconocimiento de patrones probabilísticas y determinísticas tales como KNN, ANN, SVM u otros.

- **K-Vecinos más Cercanos (del inglés *K-Nearest Neighbour*, KNN):** Este algoritmo introducido teóricamente en [7] aparece como un método que autores en [23] usaron KNN para eliminar aquellos puntos individuales en el dataset de muestreo, que bien podría corresponder a aquellos vectores que surgieron de

capturar más AP. Luego, usaron SVM para posicionar al usuario. En este trabajo los resultados muestran que el algoritmo híbrido mejora considerablemente la precisión y la velocidad de este.

A diferencia de el trabajo anterior, en [24] lo usan para directamente obtener la posición del dispositivo, a la vez que en [25] señalan que el comportamiento de KNN superó en términos de precisión al más sofisticado algoritmo basado en un modelo probabilístico.

- **Red Neuronal Recurrente (del inglés *Recurrent Neural Network*, RNN):**

En [15] señalan que las RNN han sido usadas como algoritmos de clasificación en diversas aplicaciones, debido a que puede manejar varias series de datos de tiempo, pero que actualmente no han habido investigaciones que asocien sistemas de posicionamiento indoor al uso de este algoritmo.

Por este motivo, ellos desarrollaron un sistema que se basa en RNN y que permite clasificar pares ordenados (x,y) basados en el RSSI de cada sensor. Específicamente para implementar un tipo de regresión que permite predecir las coordenadas al considerar cada AP como nodo. Algo parecido a lo que se deseaba conseguir en el trabajo presentado en [28].

- **Red Neuronal Artificial (del inglés *Artificial Neural Network*, ANN):**

En se [26] aborde el problema de posicionamiento indoor a través del mismo enfoque visto anteriormente en 2, es decir, se busca transformar los RSS a distancias y así determinar las coordenadas del objetivo. La diferencia radica en que acá no se usan AP, sino que otra tecnología llamada ZigBee que también actúa como un sensor de red, señalando que el uso de este, en conjunto con el modelo, trabajan sustancialmente mejor que la técnica de triangulación que por si sola es inestable.

Por otro lado en [27] las personas de la investigación comentan que en investigaciones anteriores o en general, en los trabajos que buscan determinar la posición basándose en una ANN utilizan características del ambiente de la Red de área local inalámbrica (del inglés *Wide Local Area Network*, WLAN) para caracterizarlo, estas pueden ser: RSS, que es lo más recurrente a la hora de abordar estos problemas, o bien, Relación Señal a Ruido (del inglés *Signal Noise Rate*, SNR) de los múltiples AP a fin de mapear el espacio en la etapa offline. En contraste a lo anterior señalan que su trabajo es capaz de alcanzar una precisión del 92.86 % solo con el uso de una red de apenas 3 capas que permiten posicionar con una exactitud de tres metros sin incurrir en un margen de error importan-

te, recalcando que si se realiza una división por área, entonces la performance del sistema mejora considerablemente y cuyos resultados fueron contrastados a través de KNN.

Finalmente en [28], al igual que en [27] utilizaron una ANN que se compuso de 4 capas, y estas a su vez, se desglosaron en 1 capa de entradas, 3 capas ocultas, con 28 nodos cada capa, y 1 de salida, con dos nodos. Interesante señalar que el resultado de agregar 2 capas se tradujo en alcanzar una precisión de 2 metros, que lleva a la conclusión de que la precisión guarda cierta dependencia con el número de nodos y capas ocultas.

1.4. Hipótesis de trabajo

"Es posible posicionar un dispositivo, identificable a través de su MAC, en un espacio interior utilizando una Red Neuronal Artificial para compensar las variaciones de RSS de los AP usados en la trilateración"

1.5. Objetivos

A continuación se señalan los objetivos que apuntan a resolver el problema presentado y a probar la hipótesis de trabajo.

1.5.1. Objetivo general

Posicionar un dispositivo móvil, reconocible a través de su MAC dentro de un espacio interior, a través de algoritmos de trilateración y ANN.

1.5.2. Objetivos específicos

- Evaluar ventajas y desventajas los modelos utilizados para posicionamiento Indoor
- Realizar mediciones de RSS
- Desarrollar una ANN para hacer el posicionamiento

1.6. Alcances y limitaciones

El alcance de este proyecto estará limitado a encontrar la posición de un dispositivo móvil dentro del pasillo del segundo piso de Edificio Tecnológico Mecánico de la Universidad de Concepción. Se ejecutará el algoritmo desarrollado en una laptop personal, utilizando Raspberry Pi para la extracción de datos y el resultado del posicionamiento se desplegará en una interfaz gráfica desarrollada en python.

1.7. Metodología

Para lograr posicionar un dispositivo móvil dentro del espacio interior, se buscará en primera instancia extraer los datos de RSSI obtenidos desde las tarjetas de Red de las placas. Posterior a esto se caracterizará el espacio de trabajo con sus respectivos niveles de RSS. Estas serán usadas para entrenar un modelo que, a través de RNN, pueda detectar la posición en que se halla un usuario, respecto de los niveles de potencia y la certidumbre que arroje el entrenamiento realizado. Finalmente, el resultado será desplegado en una interfaz gráfica.

2. Marco Teórico

2.1. Introducción

En este capítulo se sentarán las bases teóricas que soportan y respaldan el trabajar con el algoritmo escogido.

2.2. Aplicaciones y aproximaciones en sistemas de posicionamiento indoor

En cuánto a los sistemas de posicionamiento indoor, diversas documentaciones revisadas señalan la importancia de desarrollar esta tecnología a fin de resolver necesidades ya reconocidas dentro de espacios interiores tales como hospitales, museos, parques, etc. Sin embargo, todos concuerdan en que la tarea no es sencilla y por este motivo se ha intentado abordar este problema de distintas maneras, teniendo cada una sus respectivas ventajas y limitaciones.

La clasificación de estos métodos pasa por si es que en ellos se utilizan propiedades físicas o bien, reconocimiento de imágenes. De este modo, un posible ordenamiento es el que sigue:

- Análisis de escena
- Detección de proximidad
- Triangulación

2.2.1. Análisis de Escena

El algoritmo de análisis de escena se refiere al tipo de algoritmo que primero recolecta características de una sala, habitación o espacio y luego estima la posición del objetivo según coincidan los elementos del espacio.

Este algoritmo tiene la ventaja de entregar precisión, a la vez de no requerir la localización de los AP WiFi, y por ende, juega un rol sumamente relevante dentro del algoritmo de posicionamiento indoor. Sin embargo, las desventajas de estos radica en que, computacionalmente hablando, son altamente costosos y que en realidad, no están orientados a entregar posición, sino a reconocer patrones u objetivos, lo que lleva a no conocer coordenadas y geolocalización.

2.2.2. Detección de proximidad

Este algoritmo es relativamente simple, pero no preciso. Es, en general, utilizado para apoyar el posicionamiento outdoor, ya que este ayuda a estimar la posición de un objetivo utilizando la relación de proximidad entre un objetivo y los Puntos de Acceso cercanos. Cuando el dispositivo móvil objetivo recibe diferentes señales de los distintos AP, permite asociar la intensidad de señal más fuerte a una ubicación particular.

La precisión de este algoritmo viene dada por la distribución de densidad y el rango de señal de los AP, de modo que si utilizamos antenas adicionales, podemos mejorar la confiabilidad de los resultados, no obstante, esto podría encarecer la implementación.

2.2.3. Triangulación

Este método utiliza propiedades geométricas de triángulos para determinar la ubicación. Tiene dos derivaciones, angulación y lateración. De la primera, se desprende la técnica de angulación que se basa en la fase de la señal (AoA), mientras que de la segunda, aquellas basadas en la medida de la propagación por tiempo (ToA, TDOA y RTOF) y en los niveles de intensidad de señal recibida (RSSI).

2.2.3.1. Angulación

1. **AoA:** es una técnica que determina el ángulo de llegada de la señal, respecto de la ubicación conocida de estaciones base a través de envío de múltiples mensajes desde el Dispositivo móvil hacia el AP. Dado que la precisión de la medición del ángulo de llegada es la que permite estimar la posición, se tiene que estas deben ser altamente direccionales o bien, tratarse de un arreglo de antenas, lo que encarece el costo de sistemas que implementen este tipo de solución. Además, se tiene que esta técnica ve alteradas sus mediciones debido

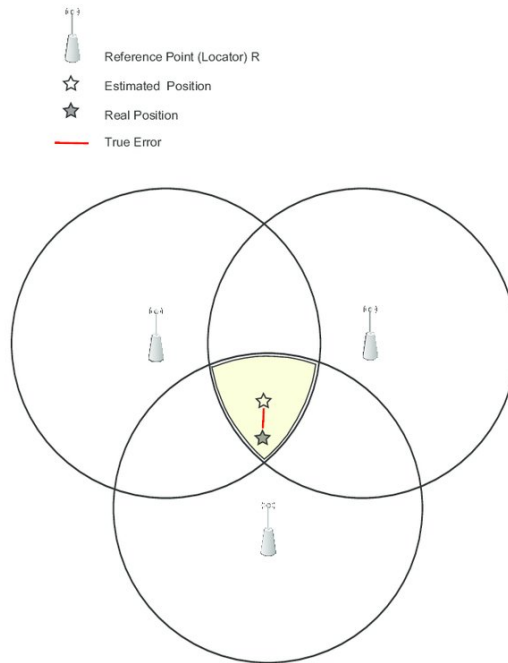
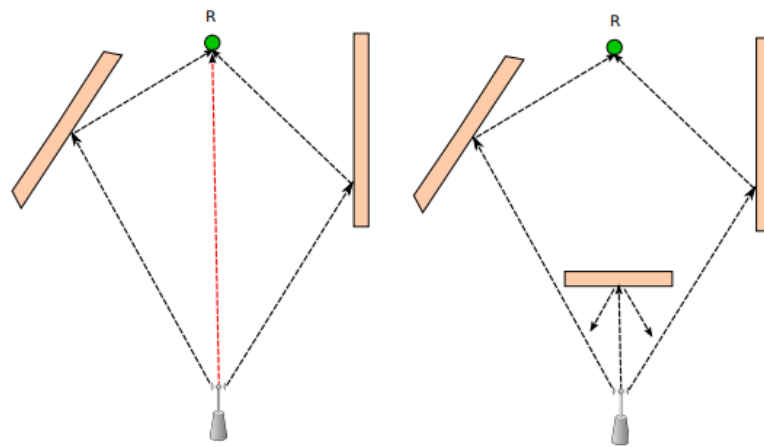


Fig. 2.1: Algoritmo de Triangulación

a efectos de la propagación multicamino de la señal cuando se trata de dispositivos que no se hallan en No Línea Vista (del inglés *None Line of Sight*, NLOS).



(a) Señal enviada en LOS

(b) Señal enviada en NLOS

Por esta razón, el sistema, aunque preciso en espacios exteriores, no es del todo idóneo para espacios interiores, ya que las reflexiones que se producen en paredes y otros objetos que cambian significativamente la dirección o ángulo de llegada de la señal y por tanto, degradan la precisión de la estimación de posición.

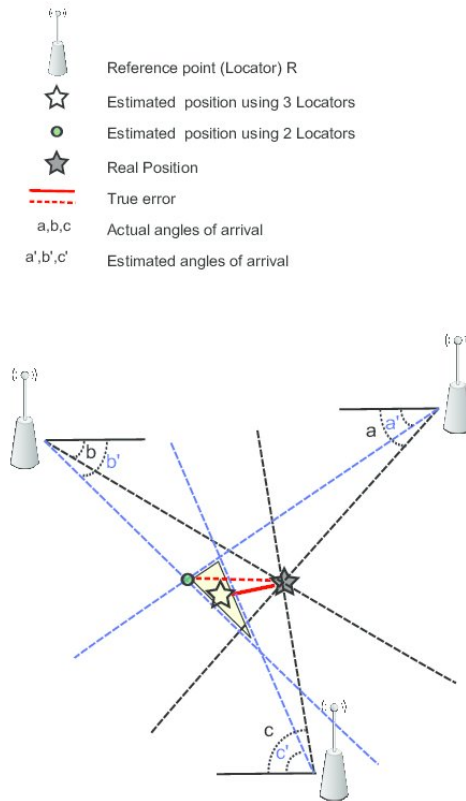


Fig. 2.3: Técnica de Ángulo de Llegada (AoA)

2.2.3.2. Lateración

■ Técnicas basadas en la propagación por tiempo

1. **Tiempo de llegada (del inglés *Time of Arrival*, ToA):** El tiempo de llegada aparece como el tiempo de viaje del transmisor hasta el receptor y permite, si es que se le multiplica por la velocidad de la luz, obtener la distancia que separa ambos dispositivos. Para medir el tiempo de llegada en el aire, este método requiere sincronización previa de los dispositivos, que para hacer posicionamiento en el espacio, requiere de al menos 4 transmisores.

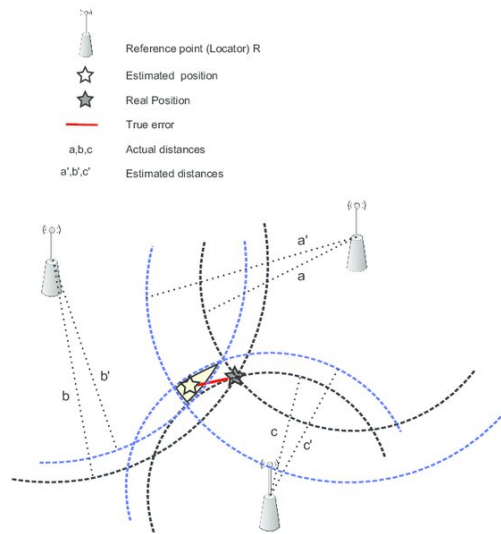


Fig. 2.4: Técnica de Tiempo de Llegada (ToA)

En ésta técnica, dado que los errores de estimación de distancia se derivan de mediciones de tiempo de salida y llegada del mensaje, se debe procurar que los tiempos sean bien medidos, esto es, que todos los dispositivos transmisor y receptor estén sincronizados en el mismo tiempo. La correcta implementación de un sistema de medición basado en ToA permite filtrar los efectos de la multitrayectoria.

2. **Diferencia de Tiempo de Llegada (del inglés *Time Difference of Arrival*, TDoA):** Esta técnica, acorde a lo mencionado en [17], consiste en medir la diferencia de tiempo de llegada de la señal que se envía desde un objeto y que es recibida por otros tres dispositivos.

Con TDoA, se puede iniciar la transmisión de la señal en un tiempo desconocido en tanto sea conocido por los otros nodos receptores. Lo único que se requiere para que esto funcione, es que entre los receptores, exista una sincronización de tiempo. Esta técnica, también conocida como Multilateración tiene como obstáculo el hecho de requerir una cantidad significativa de ancho de banda, así como también, el alto costo del equipamiento utilizado para aplicarlo.

Y es que la banda usada para la implementación de esta tecnología es la UWB, definido por la Comisión Federal de Comunicaciones (del inglés *Federal Communications Commission*, FCC) como la porción del espectro electromagnético con frecuencias que van desde los 3.1 hasta los 10.6 GHz.

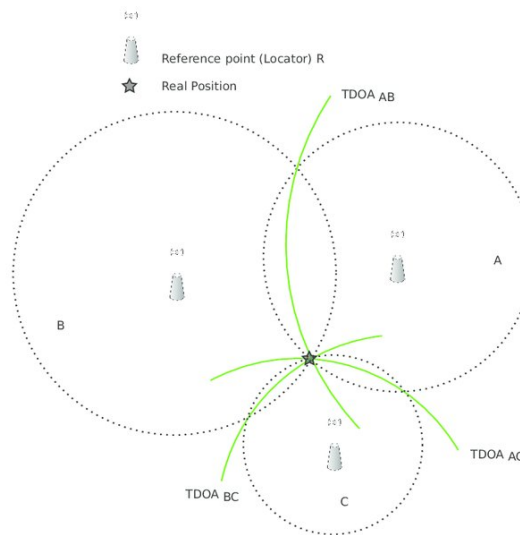


Fig. 2.5: Técnica de Diferencia de Tiempo de Llegada (TDoA)

3. **RToF**: en este método basado en 1 que consiste en el tiempo total que toma la conexión entre los dos dispositivos que participan en la comunicación. A diferencia de 2.5, acá se considera el tiempo de ida y regreso de la señal, no sólo el tiempo que tarda desde el transmisor hasta el receptor. Este método de medida no requiere ninguna sincronización entre los dispositivos, de modo que se puede implementar necesitan únicamente enviar el resultado de la medición del Tiempo-de-Viaje (del inglés *Round-Trip-Time*, RTT) para hacer el cálculo bajo la siguiente fórmula:

$$d = \frac{c \cdot (T_{Total} - T_{proc})}{2} \quad (2.1)$$

Donde d , representa la distancia entre los dos dispositivos; c , la velocidad de la luz; T_{proc} , el tiempo que le toma a la estación realizar el calculo. Siendo este último, un parámetro complejo de obtener con precisión, ya que depende de la latencia del sistema. Esto se escala como la principal desventaja de este sistema ya que calcular la posición para varios dispositivos se traduce en lidiar con latencias que impidan obtener la posición de dispositivos que se mueven rápidamente.

Fig. 2.6: Diagrama de medición de distancia usando RToF

■ Técnicas basadas en los niveles de intensidad de señal recibida

1. **RSSI**: Respecto de esta técnica se tiene que algunos autores la introducen como una técnica propia de algoritmos de Detección de Proximidad (2.2.2) debido a que los niveles de los indicadores de intensidad de señal recibida, al guardar relación con la distancia, tal y como señalan los modelos de propagación para interiores, establecen que a menor distancia del AP, mayor debe ser la intensidad de señal recibida.

Los mismos autores que la clasifican dentro de la sección mencionada, utilizan esta técnica para trazar un sistema de coordenadas basadas en las intensidades recibidas respecto de cada AP, esto les permite mapear la zona en una etapa que llaman offline y ubicar al dispositivo en la etapa posterior, llamada online.

En esta etapa, se asigna a través de un sistema de coordenadas, los niveles de intensidad de señal, tal que es posible posicionar el objetivo según los RSS de los AP, como se ve en la imagen.

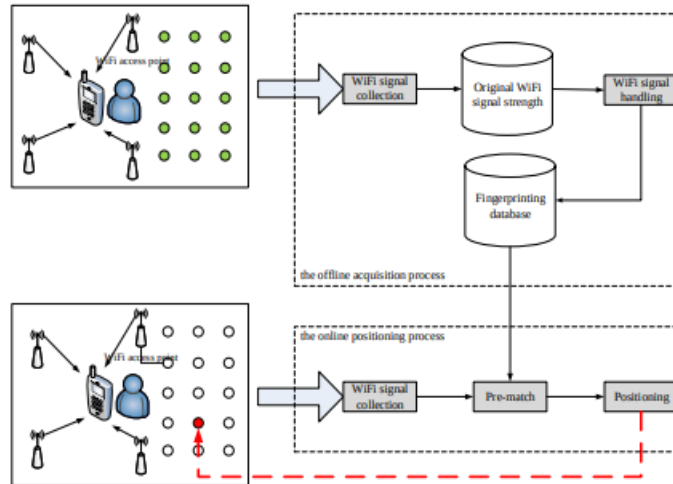


Fig. 2.7: Idea de algoritmo basado en RSSI

2. Algoritmos de posicionamiento para método por características de posición

- a) **Método Probabilístico:** En este método, se considera el estimar la posición del objetivo, como un problema de clasificación, ya que desde un vector de RSS capturadas desde los APs, se puede estimar la probabilidad de que un valor S , pertenezca a un AP particular, y por ende, a un punto de coordenadas aproximado.

Se puede hacer el cálculo de la posición a través de modelos de inferencia y asociarlo también a un tipo de Cadenas de Markov, así como también, apoyar el cálculo introduciendo medidas de dispersión tales como la desviación estándar y la varianza, o de una medida de tendencia central como es la media.

- b) **Método Máquinas de vectores de soporte (del inglés *Support Vector Machine*, SVP)**: En la literatura consultada [7] los autores señalan a esta técnica como "Una nueva u prometedora técnica para clasificación de datos y *Machine Learning*.", que ha sido usada con resultados satisfactorios en el mapeo de intensidad de señales inalámbricas.

Su implementación consiste igualmente en dos fases; una de entrenamiento, que es aquella en que los sets con los datos de intensidad de señal de los AP son ingresados, y con los cuales se buscará diseñar un hiperplano que clasifique los vectores en dos clases, procurando maximizar la distancia entre estos.

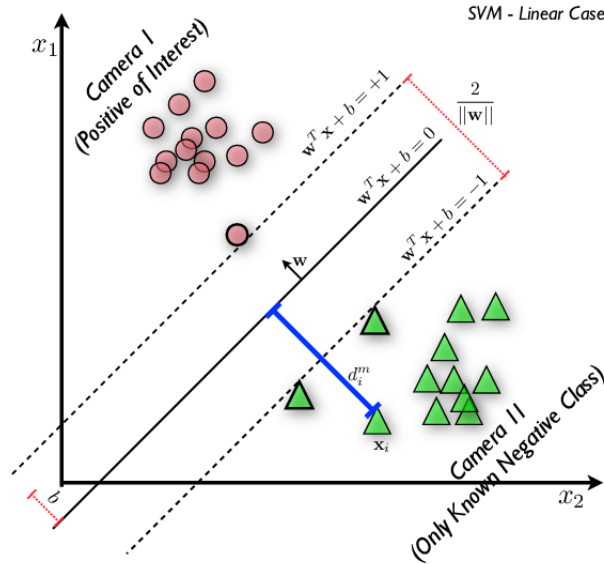


Fig. 2.8: Distintos hiperplanos para el mismo set de datos.

Sin embargo, dado que en un espacio indoor, la cantidad de AP es , por lo general, superior a 2, la posibilidad de que en un set de datos podamos diferenciar entre un AP u otro mediante el procedimiento de encontrar el hiperplano que separa maximizando la distancia entre los valores, es nula, e implementar un algoritmo recursivo que vaya evaluando para cada AP, el hiperplano que clasifica la medición del objetivo en alguno de ellos, es costosa computacionalmente además de compleja.

- c) **Método KNN**: El método de vecinos más cercano es uno de los métodos de clasificación más simple dentro de la librería de Machine Learning ya que consiste, en una clasificación por similitud respecto de otros puntos en un dataset de entrenamiento.

El objetivo de este método es el de agrupar vectores de características, que no son más que representaciones matemáticas para la información. Ha de tenerse en cuenta que el vector de características puede requerir un pre procesamiento si es que estas no son necesariamente numéricas, ya que de este modo se podrá agrupar los datos como si se tratara de un vector de R^N .

Ahora, a diferencia de otros métodos de clasificación, como es el caso de $2b$, en KNN la fase de entrenamiento no existe, de manera que sólo se trata de una fase. En cambio, se tiene que cualquier caso por abstraer o generalizar la información que allí está, es un hecho en la etapa de clasificación.

Esto significa que se puede trabajar inmediatamente sobre los datos, en tanto estos, tal y como se mencionó previamente, no requieran un preprocesamiento. Dado que lo que hace este método es agrupar valores, el resultado óptimo será cuando exista un reducido número de data-sets y estos no tengan valores no numéricos.

Así, se tiene que los datos que se clasificarán pueden también ser representados como una matriz de $M \times N$ datos, de donde ; M , es el número de datos; y N , son las características de los datos. Dicho esto, el proceso de agrupamiento de los datos se hace según las siguientes consideraciones:

- Se calcularán las distancias entre los valores a clasificar, y aquellos contenidos en el data-set. En el cálculo de la distancia suele usarse la distancia euclidiana, que es básicamente, la magnitud del vector que se obtiene al restar un punto de muestra del data-set de entrenamiento, con el punto a clasificar

$$E(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (2.2)$$

- Se escogerá un valor K , que por lo general, no es ni múltiplo de la cantidad

de clases, ni par, y que representa la cantidad de valores K con las distancias más bajas. Este puede ser escogido arbitrariamente o bien, en caso que sea necesario, obtenido por validación cruzada.¹

- Finalmente, se agruparán los datos a una clase, a través de la comparación de las distancias obtenidas. Si este, posee la mayoría de los K de puntos cuya distancia los asocia a una clase A, entonces este será clasificado en dicha clase. Si por el contrario, el número de puntos cuya distancia se acerca a los datos objetivo no son mayoría, entonces este pertenecerá a B.

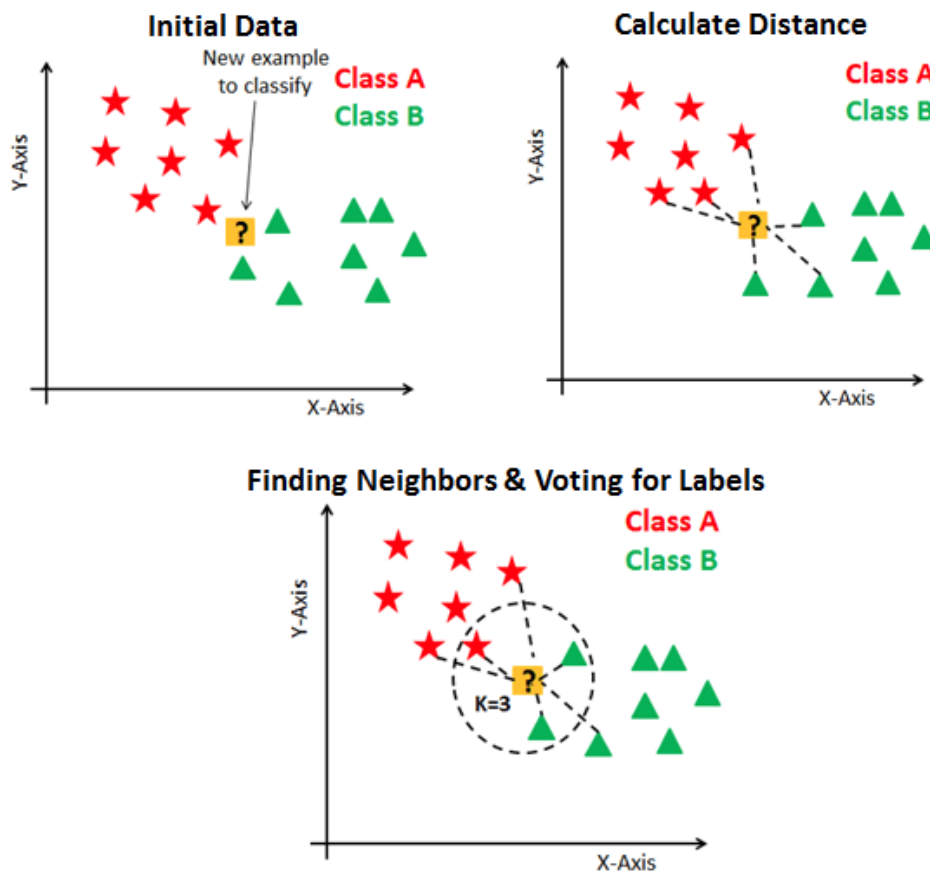


Fig. 2.9: Representación gráfica del funcionamiento de KNN según el valor de K

¹La validación cruzada o cross-validation es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba.

d) **Método ANN:** Este método se vale de una neurona artificial llamada **perceptrón**, algo que fuera desarrollado entre los años 1950s y 1960s por el científico Frank Rosenblatt, inspirado por el trabajo de Warren McCulloch y Walter Pitts «A Logical Calculus of the Ideas Immanent in Nervous Activity» [19] donde señalan la posibilidad de rescatar la idea de un sistema complejo de redes y conexiones, de impulsos y señales, como es nuestro sistema nervioso, y extrapolarlo a uno computarizado, donde en lugar de estímulos, se tuvieran señales.

Se mencionó en [29] que uno de los objetivos del trabajo de McCulloch y Pitts fue el de introducir la noción de neurona artificial, por esto es que el perceptrón le debe su estructura, presentada a continuación:

Fig. 2.10: Comparación por similitud de una neurona con un perceptrón

En el cerebro humano, existen distintos tipos de neuronas las cuales están destinadas para tareas específicas, por esta razón, puede darse que una neurona reciba señales de temperatura, presión, luz, entre otras, antes de transmitir un impulso. Este último viene en algunos casos a ser la suma ponderada de los distintos estímulos nerviosos que provienen de las otras terminaciones sinápticas, tal que es en el soma donde se evalúa si esta excede el potencial de acción o no. Si lo hace, la señal será enviada a través del soma y las terminaciones sinápticas hasta un órgano diana u otra neurona, pudiendo ser esta señal inhibitoria, excitadora o moduladora.

Tal como se puede ver en 2.11, se tiene que están las conexiones de la capa de entrada, representadas por impulsos o señales de entrada x_j , que luego se multiplican por los pesos w_j y pasan a evaluarse en la función de activación.

Las neuronas son representadas como unidades individuales de procesamiento, que al agruparse forman una red neuronal cuya función de activación, y luego, las capas se resumen en tres:

- **Capa de entrada:** las unidades que componen esta capa son los campos de entrada, por donde ingresan los primeros estímulos
- **Conexiones ponderadas:** Las operaciones que se realizan en las capas, se

Fig. 2.11: Representación de un Perceptrón

logran a través de las fuerzas de conexión variables, o ponderaciones, que resultan del producto entre los pesos w_j , que representan la fuerza que tiene un nodo en particular, y los estímulos x_j . Los datos de entrada se presentan en la primera capa y estos se propagan a través de las capas ocultas, en forma de *función de propagación*, tal cual hiciera el impulso eléctrico en las neuronas, hasta llegar a la capa de salida que es donde el dato de salida, nos permitirá tomar decisiones o evaluar resultados.

- **Capas Ocultas:** Son aquellas unidades de procesamiento que transforman el impulso de entrada en otro diferente pero que no son realmente ni una entrada ni una salida.
- **Función de activación:** Es quizás la característica principal o definitoria de las neuronas, la que mejor define el comportamiento de la misma. Se usan diferentes tipos de funciones, desde simples funciones simples de umbral a funciones no lineales. Se encarga de calcular el nivel o estado de activación de la neurona en función de la entrada total.

La función de activación, en este caso Sigmoide, corresponde una función continua no lineal cuyo rango va entre 0 y 1 que puede ser representada gráficamente a través de 2.12 y matemáticamente a través de 2.3:

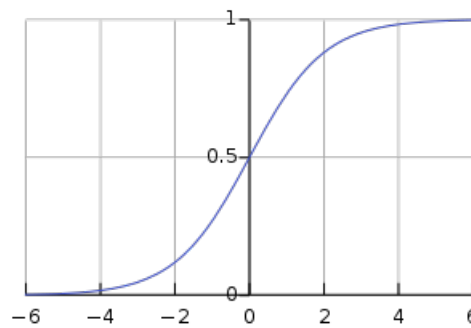


Fig. 2.12: Función Continua no lineal Sigmoide

$$S(u_j) = \frac{1}{1 + e^{-u_j}} \quad (2.3)$$

- **Capa de salida:** Compuesta por una o varias unidades que representan los campos de salida y que entregan el resultado que finalmente es modelado por

la siguiente expresión:

$$O_j = S(w_j x_j + w_0) \quad (2.4)$$

Donde w_0 , es el Bias, una constante que ayuda al modelo de una manera que puede ajustarse mejor a los datos dados y O_j , el resultado particular de la evaluación de un dato en la función Sigmoide.

Al igual que hace el cerebro por acción de las neuronas, una red es capaz de aprender examinando los registros individuales, generando una predicción para cada uno de estos y ajustando, a través de los pesos, las ponderaciones en caso que el resultado no haya sido el óptimo. Este proceso es iterativo y puede tener lugar cientos de veces antes que la red sea capaz de mejorar sus predicciones antes de alcanzar resultados que satisfagan los criterios de parada.

En un comienzo, las ponderaciones son aleatorias, y las respuestas suelen ser muy dispares, sin embargo, y conforme va avanzando el entrenamiento, esta va mejorando, haciéndose más precisa en aprender a partir de los resultados conocidos y estimar la posición.

No obstante, obtener una buena precisión, por si sola no es un buen indicador de que nuestra red está realizando una buena estimación de la posición, por esto generalmente se extraen otros indicadores tales como:

- **Bias:** El sesgo estadístico, es una característica de una técnica estadística o de sus resultados, donde el valor esperado difiere del verdadero resultado.

Dicho de otro modo, es la incapacidad de un modelo de Aprendizaje automático o *machine learning*, en este caso, de una Red Neuronal Artificial, de extraer la verdadera relación entre los valores reales y el predicho.

Existe una relación subyacente entre el valor del bias y qué tan exacta es la predicción hecha por el modelo, esta es tal que si existe un:

- **Bajo Bias:** Los valores predichos son cercanos a los valores reales y el modelo no necesita extraer demasiadas características para lograr este re-

sultado.

- **Alto Bias:** Los valores predichos son cercanos a los valores reales y la red debe hacer un mayor trabajo con los datos a fin de extraer las características que este ignoró y que le permitirán llegar al resultado.

Lo anterior tiene por consecuencia que aun cuando el modelo intenta reducir el sesgo o bias, este tiende a sobre ajustar los datos, produciendo lo que llamamos, *Overfitting*.

- **Varianza:** Definida como la medida de cuán alejados, o esparcidos, están unos de otros los datos, respecto de su valor medio, la varianza permite entender cómo, están repartidas las predicciones hechas por el modelo. Así, los distintos valores de este indicador señalarán características interesantes de la red tales como que si la varianza es baja, los valores se encuentran cerca unos de otros y los ajustes en caso que el bias sea alto, son menores. En cambio, si la varianza es alta, aun cuando se tenga un bias bajo, el funcionamiento de la red no será el óptimo pues cambiará significativamente según los datos de entrenamiento.

Es entonces que a partir de combinaciones de estos parámetros se obtiene una guía para comprender cuándo los valores estimados por la red serán los correctos

- Verdaderos positivos
- Falsos positivos
- Verdaderos Negativos
- Falsos Negativos

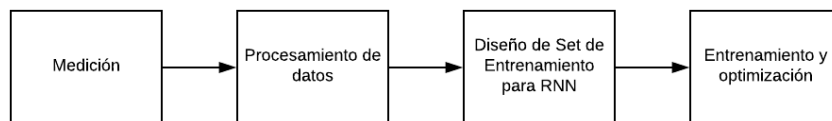
3. Diseño de algoritmo

3.1. Introducción

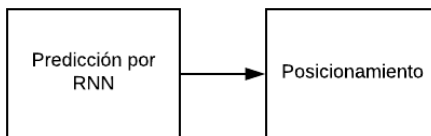
En este capítulo se pretende abordar el diseño para llevar a cabo el objetivo principal y algunos de los específicos del proyecto.

3.2. Diseño general de trabajo

En la figura 3.1 se observan las partes que componen el trabajo del sistema. En este sentido, es necesario hacer una diferenciación entre la parte *offline* y la parte *online* ya que cada parte por separado conlleva requerimientos específicos.



(a) Etapa Online



(b) Etapa Offline

Fig. 3.1: Diagrama General de operación de Algoritmo de posicionamiento

3.2.1. Offline

3.2.1.1. Medición

En éste módulo se extraerán los datos usando como hardware, un ordenador de placa reducida como es Raspberry Pi 2 y un Adaptador USB Nano Inalámbrico N de 150Mbps, modelo TL-WN725N. Los cuales tendrán por función medir los niveles de intensidad de potencia de cada uno de los Access Points que se hallen a su alrededor.

3.2.1.2. Extracción y procesamiento de datos

Para este módulo se trabajará con un algoritmo diseñado cuya función es hacer el pre-procesamiento de los datos. Esto se traduce en que el algoritmo escribe toda la información recopilada en un archivo que posteriormente es leído para filtrar de este sólo los datos que requerimos, esto es; la MAC del AP, el Identificador de Set de Servicio Extendido (del inglés *Extended Service Set Identifier*, ESSID), así como el RSSI. Luego, el archivo obtenido corresponde al dataset de entrenamiento que será el usado entrenar la red.

3.2.1.3. Diseño de Set de Entrenamiento

Con el objetivo de relacionar los Niveles de Intensidad de Señal Recibida, con una posición en particular, es menester asociar estos a coordenadas en el espacio, para ello en este punto se leerá el archivo obtenido en puntos anteriores y coordenadas presentadas como pares ordenados, serán incorporados como un elemento más al dataset de entrenamiento.

3.2.1.4. Entrenamiento y Optimización

Dado que en principio, la red es incapaz de percibir la relación existente entre la posición y la densidad de potencia radiada, es preciso relacionar los datos con que se alimenta la red. Por esta razón es necesario hacer un ajuste adicional sobre estos.

Para que el sistema aprenda, es preciso que el dataset sea tan robusto como se pueda. Sin embargo, en lo que respecta a redes neuronales conceptos como sobreajuste (*Overfitting*) y subajuste (*Underfitting*) llevan a poner especial cuidado en que el sistema, con un número

óptimo de iteraciones, datos y neuronas, funciones de activación y capas, alcance efectivamente el resultado deseado.

Por ello se implementarán técnicas de optimización de redes neuronales que cuidarán que los parámetros antes mencionados no lleven a resultados que sean distintos de los óptimos.

3.2.2. Online

3.2.2.1. ANN para estimar la posición del objetivo.

En este punto, se tomarán nuevas muestras de intensidad de señal a fin de entregarlas valores de prueba para predecir, según lo que aprendió la red en su entrenamiento, cuál es la posición en que se haya el objetivo.

Como salida de la red se tendrá la posición y la precisión con que esta fue estimada.

4. Implementación

4.1. Introducción

En este capítulo se describirá la implementación del proyecto, dando énfasis en las tecnologías y dispositivos usados para el posicionamiento.

4.2. Implementación general de trabajo

A continuación, se observa el esquema general del algoritmo a implementar y cómo confluyen ambas partes para llevar a cabo el posicionamiento.

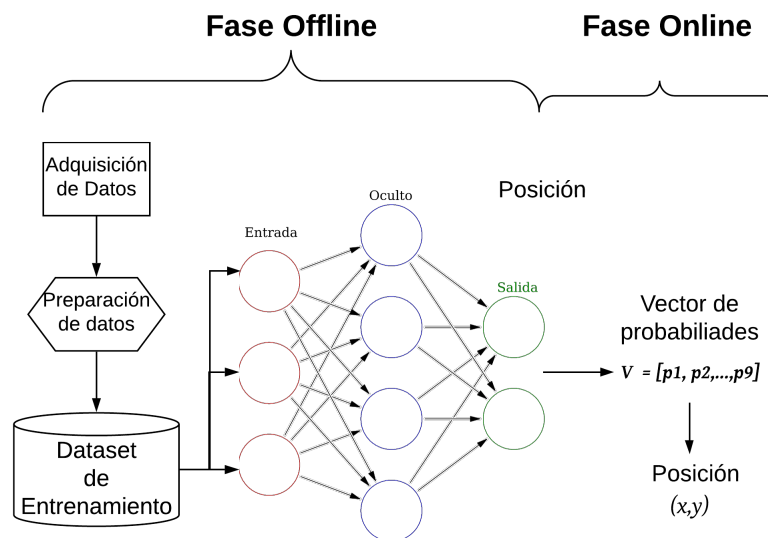


Fig. 4.1: Esquema de Implementación General de Trabajo

4.2.1. Fase Offline

4.2.1.1. Adquisición de RSSI

La adquisición de estos datos se hará con la librería de linux llamada `iwlist`. Una breve descripción de esta, a continuación:

- **Descripción:** `Iwlist` es usada para mostrar información adicional de dispositivos inalámbricos de una red. El argumento principal es usado para seleccionar una categoría de información, `iwlist` muestra información en detalle para toda la información relacionada a esta categoría, incluyendo la información que se muestra en `iwconfig`.
- **Scan:** Entrega una lista completa de los Access Points y celdas Ad-Hoc en el rango, y opcionalmente, información adicional sobre ellos. Algunos de estos datos adicionales son ESSID, Quality, Frequency, Mode, etc que permiten entre otras cosas, reconocer la **dirección MAC** de los dispositivos, el **RSSI**, y el **nombre de la red**.

```
Cell 01 - Address: A8:7D:12:D3:45:04
          Quality=67/70  Signal level=-43 dBm
          ESSID:"HUAWEI P20 lite"
Cell 02 - Address: C8:D7:19:DC:7D:C7
          Quality=33/70  Signal level=-77 dBm
          ESSID:"Sistemas Paralelos"
Cell 03 - Address: EC:22:80:F3:EB:EA
          Quality=28/70  Signal level=-82 dBm
          ESSID:"Baxter"
Cell 04 - Address: CC:2D:E0:7A:79:AB
          Quality=27/70  Signal level=-83 dBm
          ESSID:"Dinamica de suelos"
Cell 05 - Address: 00:02:6F:E5:B6:5A
          Quality=50/70  Signal level=-60 dBm
          ESSID:"DIE"
Cell 06 - Address: 10:FE:ED:C0:6D:C0
          Quality=29/70  Signal level=-81 dBm
          ESSID:"OpenWrt"
Cell 07 - Address: 74:DA:DA:FE:21:D4
          Quality=25/70  Signal level=-85 dBm
          ESSID:"Laboratorio de Accionamiento"
Cell 08 - Address: 94:10:3E:86:13:20
          Quality=27/70  Signal level=-83 dBm
          ESSID:"LabSerc 2.4"
Cell 09 - Address: E8:DE:27:6D:4A:E7
          Quality=39/70  Signal level=-71 dBm
          ESSID:"VLSI_2.4GHz"
Cell 10 - Address: FA:DA:0C:83:7D:17
          Quality=39/70  Signal level=-71 dBm
          ESSID:"DIRECT-17-HP M130f LaserJet"
```

Fig. 4.2: Salida de consola con datos de AP

En esta parte, es necesario recalcar que las características a desplegar dependerán de la tarjeta de red de manera que para efectos del proyecto, se limitarán al adaptador de red inalámbrico conectado a la raspberry pi

Con lo mencionado anteriormente podemos hacer un mapeo del laboratorio tal que para una posición específica sea, por ejemplo $(x, y) = (0, 0)$ el punto de origen de las mediciones, puedan obtenerse los niveles de intensidad de señal. De este modo, extendiendo esta metodología a las demás posiciones es que se caracterizará el espacio en función del parámetro RSSI en 3.1.

Las mediciones se harán en espacio accesible y de tráfico relativamente bajo como es el laboratorio de redes de datos. Este para efectos prácticos fue dividido en dos partes como se presenta en la siguiente imagen:

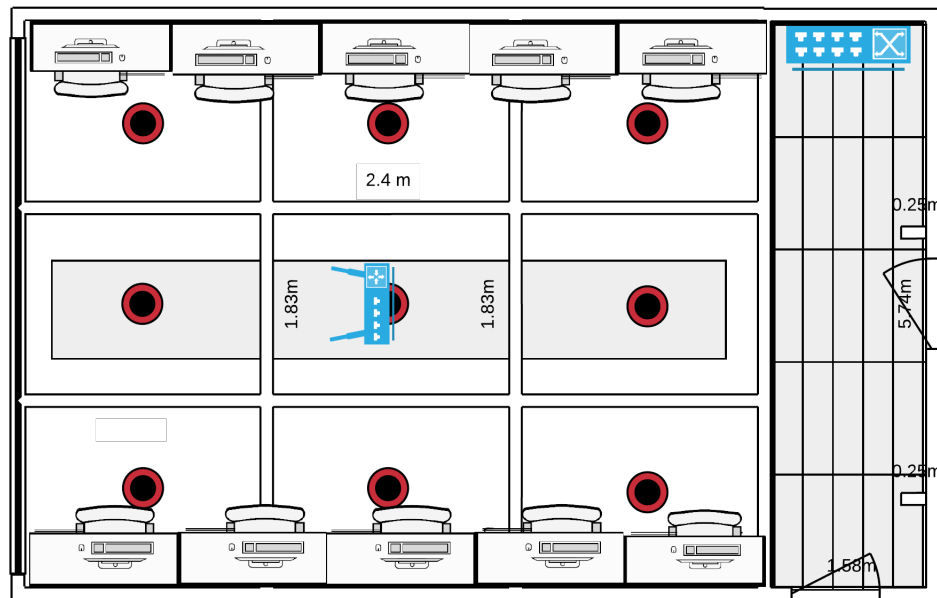


Fig. 4.3: Mapa Lab. de Redes de datos

- **Escritorios y computadores:** corresponden a la parte que no está ennegrecida y que fue dividida en 9 sectores de 2.4 x 1.83 m en los cuales se dispusieron los equipos -representados a través de los puntos rojos con negro, a una altura de 0.8 m para así medir. Estos fueron ubicados equidistantes uno de otro.
- **Franja no utilizable:** este sector corresponde a un espacio que dada su relación asimétrica de profundidad y densidad de materiales no fue considerado en el estudio.

4.2.1.2. Set de entrenamiento

En esta parte, los datos son filtrados, separados y agrupados en un archivo .csv para la posterior lectura y procesamiento de la Red Neuronal.

El formato de los datos de entrada proveídos a la red es el siguiente.

```

["Convertidores Estaticos"
20:AA:4B:AE:FD:0A", ""UdeC""
84:D4:7E:29:3E:80", ""FabLab""
AC:84:C6:3E:EF:51", ""SMA1930055648""
B0:72:BF:05:3C:3D", ""Lab ERNC""
00:90:4C:91:00:03", ""Eduescobl""
68:FF:7B:1D:44:01", ""eduroam""
84:D4:7E:29:3E:82", ""LCDA""
B0:48:7A:DE:14:8C", ""Sistemas Electronicos""
64:70:02:40:6C:DE", ""AP_1612927617""
BC:54:F9:FC:7F:60", ""Laboratorio de Redes""
48:F8:B3:22:4C:FA", ""Microondas""
48:F8:B3:22:4D:4B", ""Lab Energia Renovable 2""
98:FC:11:54:32:53", ""CE5CFB""
98:FC:11:E2:A6:A5", ""Sala Dise\xC3\xB1o""
0A:18:D6:C8:01:A4", ""Ing. Civil Telecomunicaciones""
00:1E:E5:7B:E0:BE", ""Lab Energias Renovables""
F4:EC:38:D2:09:78", ""Vibraciones""
50:C7:BF:DD:C3:B3", ""LabSerc 2.4""
94:10:3E:86:13:20", ""Die231""
EC:08:6B:42:B5:4A", ""Controlistas""
54:E6:FC:E7:FA:D2", ""Ingenieria""
4A:D9:E7:21:01:36", ""Lab 212""
00:23:CD:18:65:9C"
-65,-65,-72,-79,-72,-79,-76,-79,-79,-77,-54,-68,-72,-72,-71,-66,-71,-72,-77,-79
,-89,-95,-95,-87
-65,-65,-74,-79,-71,-79,-76,-79,-77,-77,-54,-66,-72,-72,-72,-66,-72,-72,-77,-79
,-89,-95,-92,-87,-95,-75
-68,-65,-75,-79,-72,-79,-76,-79,-77,-77,-54,-66,-72,-72,-73,-66,-73,-72,-77,-79
,-89,-95,-92,-87,-95,-75
-68,-65,-74,-79,-72,-79,-76,-79,-76,-77,-54,-66,-72,-72,-73,-66,-73,-72,-77,-79
,-89,-96,-91,-86,-95,-75
-68,-65,-74,-79,-72,-79,-76,-79,-75,-77,-54,-65,-74,-72,-73,-66,-73,-72,-77,-79
,-89,-95,-91,-86,-95,-75,-65,-79
-67,-65,-75,-79,-73,-79,-77,-79,-74,-78,-53,-65,-74,-72,-73,-64,-73,-73,-77,-79
,-86,-95,-91,-86,-95,-75,-65,-79
-67,-75,-79,-72,-79,-77,-79,-74,-78,-53,-65,-74,-72,-72,-65,-73,-73,-76,-79,-86
,-94,-91,-86,-95,-74,-65,-79,-64
-67,-74,-79,-72,-79,-76,-79,-74,-78,-52,-65,-74,-72,-72,-64,-72,-73,-76,-79,-86
,-94,-92,-86,-95,-74,-65,-79,-64
-67,-73,-79,-72,-79,-76,-79,-73,-78,-52,-65,-73,-72,-72,-65,-72,-73,-76,-79,-85
,-94,-93,-84,-95,-75,-65,-79,-65,-97
-67,-73,-79,-72,-79,-76,-79,-73,-78,-52,-65,-73,-72,-72,-65,-72,-73,-76,-79,-85
,-94,-93,-84,-75,-65,-79,-65,-97

```

Fig. 4.4: Formato tipo de datos entregados

Del listado anterior cabe recalcar que lo único relevante es la potencia recibida, medida en dBm, desde los AP cercanos. La información de las etiquetas de las columnas resulta prescindible por cuanto no es objeto de interés de este trabajo, el establecer una correlación entre los RSSI de AP específicos y la posición de un objetivo, sino el hacerlo indistintamente de cuál sea el AP.

4.2.2. Red Neuronal

La red neuronal se compone de las siguientes capas:

- **Capa de entrada:** Esta capa recibirá tantos valores de entrada como mediciones de RSSI haya para el total de posiciones que se desee estimar. Esto se traduce en la práctica en obtener el número de filas del archivo `csv` mencionado previamente.
- **Capas ocultas o *Hidden layers*:** Corresponde al espacio intermedio existente entre la capa de entrada y la capa de salida, que bien puede tratarse de una o varias capas que propagan los valores de entrada hacia la salida ajustando según sea necesario, los pesos que llevan, en este caso, los valores de RSSI a una predicción correcta.
- **Capa de salida:** Corresponde a la capa final del modelo y representa la salida de la red. En este caso, sus valores de salida son la probabilidad con que se estima que pueda ser cada posición. De este modo, para un valor de entrada, es decir, un vector de potencias, la salida luego del paso de la red corresponde a una lista con las probabilidades de que sea una u otra posición. Cabe mencionar que posterior a la salida, debe ejecutarse una interpretación de los resultados, esto es, traducir desde lo que son las probabilidades, a lo que es una posición dada.

```
Model: "sequential_6"
```

Layer (type)	Output Shape	Param #
dense_31 (Dense)	(None, 18001)	252014
dropout_26 (Dropout)	(None, 18001)	0
dense_32 (Dense)	(None, 18)	324036
dropout_27 (Dropout)	(None, 18)	0
dense_33 (Dense)	(None, 24)	456
dropout_28 (Dropout)	(None, 24)	0
dense_34 (Dense)	(None, 16)	400
dropout_29 (Dropout)	(None, 16)	0
dense_35 (Dense)	(None, 12)	204
dropout_30 (Dropout)	(None, 12)	0
dense_36 (Dense)	(None, 9)	117
Total params: 577,227		
Trainable params: 577,227		
Non-trainable params: 0		

Fig. 4.5: Diagrama de funcionamiento offline

Es sobre estas capas en las que se aplicarán las técnicas de Cross-validation y Grid-Search para entregar información respecto de; cuál es el valor de las métricas que reflejan el com-

portamiento real de la red y cuáles son los parámetros (n° de capas y epochs, optimizador y batch-size) y con ello, optimizar el funcionamiento de la red. Cabe mencionar, que dentro de los indicadores que se obtendrán se encuentra la Matriz de Confusión, con esta se apoyará la exactitud y precisión de la red.

5. Resultados

5.1. Introducción

Los resultados obtenidos en esta sección representan la aplicación metódica de cada una de las etapas presentadas como objetivos, en la 4.1, de modo que, habiendo seguido la estructura de trabajo antes mencionada, lo presentado en esta parte se desglosará en dos partes: *resultados de la etapa online* y *resultados de la etapa offline*.

5.1.1. Etapa Offline

El presente trabajo se desarrolló en el laboratorio de redes de la carrera Ingeniería civil en Telecomunicaciones, ubicado en el segundo piso del Departamento de Ingeniería Eléctrica de la Facultad de Ingeniería.

Se escogió este espacio dadas sus características estructurales y uso, pues esto entrega una muestra fidedigna en cuanto a si es o no factible implementar la solución propuesta en este trabajo hacia espacios de mayor superficie y flujo de personas.

5.1.1.1. Medición

La medición como se mencionó anteriormente, se llevó a cabo usando un adaptador de red Wifi conectado a una raspberry pi 2, sobre la cual se ejecutó un comando de la librería `iwtools`, que permite obtener mediciones de intensidad de señal, paquetes recibidos y enviados, así como también el ESSID y la MAC de los APs que se encuentran alrededor de este equipo.

Estos datos (4.2) son obtenidos para luego ser escritos en un archivo de texto plano (5.1a) sobre el cual se aplican distintos filtros de búsqueda que permitan extraer específicamente los datos de la intensidad de señal obteniéndose lo siguiente:

A fin de estimar cuál era la cantidad idónea de muestras a tomar por ubicación, se optó por realizar un ensayo. Para esto, dividiendo el laboratorio en 4 sectores, se posicionó una raspberry

5.1.2. Laboratorio dividido en sectores.

Tal y como se mencionó en 4.3, el laboratorio fue dividido en 9 sectores en los que se realizaron mediciones de potencia con una raspberry en cada ubicación.

Puntos a tratar en esta sección:

1. Cantidad apropiada de datos
2. Tiempo que tarda la medición
3. Tiempo que tarda el algoritmo de recolección de datos

5.2. Entrenamiento

5.2.1. Ajustes entrenamiento

1. Cantidad apropiada de neuronas
2. Cantidad de epochs
3. Tiempo que tarda el algoritmo de búsqueda de los mejores parámetros
4. Tiempo que tarda el algoritmo que entrega la media del valor de precisión de la red
5. Tiempo que tarda la red en entrenarse con esos parámetros

5.3. Predicción

6. Conclusiones

6.1. Sumario

Dentro del desarrollo del estudio, se contemplaron diversos enfoques, se ponderó la posibilidad de implementar en la memoria de título un sistema de posicionamiento indoor basándose en algoritmos de estimación de posición geométricos y otros, en parámetros más complejos, no obstante, la poca practicidad y adaptabilidad de estos modelos al ambiente donde se podría implementar, supone una desventaja innegable.

6.2. Conclusiones

Por todo lo revisado, estudiado y expuesto, se plantea la posibilidad de que, como alternativa a los algoritmos geométricos o basados en potencia, se consideren alternativas más efectivas, de menor coste computacional y que no dependan de parámetros relacionados con el ambiente, dado que en ambientes indoor o en sectores muy concurridos hay métricas que varían considerablemente, al punto de obligar a realizar mediciones correctivas.

Además, se tiene que la efectividad y precisión del método que recoge las solicitudes IRDP, es considerablemente superior al no depender de otros parámetros salvo la llegada de los paquetes al router.

Se dio cumplimiento a los objetivos planteados inicialmente, mediante la investigación de las alternativas, que fueron abordadas a fondo a través de literatura. Se pudo llegar a una comprensión justificada de las ventajas y desventajas de los métodos, pudiendo así, escoger con base en distintos estudios teóricos y empíricos, cuál es el método que convendría desarrollar en detalle en la memoria de título.

En suma, se destaca la importancia de llegar a desarrollar un sistema de posicionamiento indoor que cuente con una precisión que no sea directamente proporcional a la capacidad de cómputo, tanto si se desea resolver necesidades humanas tales como, hallar la mejor ruta de acceso para una persona con capacidades diferentes, a través de una aplicación móvil que desde

el seguimiento de su interfaz vaya dándole instrucciones. Como si se desea implementar un sistema que recoja las solicitudes de los clientes que desean conectarse y con ello, hacer un seguimiento de las posiciones, a fin de que este, permita generar reportes e informes estadísticos que relacionen patrones de conducta y también de consumo, que favorezcan a un mejoramiento en las estrategias de venta y en la experiencia de compra del cliente.

6.3. Trabajo Futuro

Se listan las posibles líneas de investigación que se deducen directamente de la presente obra.

1. Implementación de un sistema que realice el seguimiento de las solicitudes IRDP donde pueda visualizarse el movimiento de las interfaces de red
2. Realizar un estudio que contraste la precisión y eficacia con la que los métodos geométricos y el de las solicitudes IRDP realizan la estimación de posición.
3. Estudiar algoritmos de posicionamiento basado en redes neuronales.
4. Estudiar la precisión de las mediciones obtenidas por el uso del sistema presentado en [8]

Se quiere que al menos existan unas 3 a 5 ideas en las que se pueda seguir investigando.

7. Códigos

7.1. Derivación de fórmulas

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

7.2. Unidades especiales

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Bibliografía

- [1] VAN VINH NGUYEN y WEON LEE JONG, «Self-Positioning System for Indoor Navigation on Mobile Phones» *IEEE International Conference on Consumer Electronics (ICCE)*, 2012
- [2] REA MAURICIO, CORDOBÉS HÉCTOR, GUSTINIANO DOMENICO, «Twins: Time-of-flight based Wireless Indoor Navigation System»
- [3] BIANCA BOBESCU, MARIAN ALEXANDRU «Mobile indoor positioning using Wifi localization» Review of the Air Force Academy. Transilvania University, Brasov, Romania
- [4] MAYUR TAWARI, ONKAR PATHAK, RAJESH PALASKAR & RAJESH PALKAR, «Wi-Fi Indoor Positioning System Based on RSSI Measurements from Wi-Fi Access Points –A Trilateration Approach», *International Journal of Scientific & Engineering Research*. V5, 2014.
- [5] C. YANG AND H. R. SHAO, «WiFi-based indoor positioning,» in *IEEE Communications Magazine*, vol. 53, no. 3, pp. 150-157, March 2015. doi: 10.1109/MCOM.2015.7060497
- [6] ZAHID FARID, ROSDIADEE NORDIN & MAHAMOD ISMAIL, «Recent Advances in Wireless Indoor Localization Techniques and System», *Journal of Computer Networks and Communications*, 2013.
- [7] ZHAO KAI, LI BINGHAO y ANDREW DEMPSTER, «A Comparison of algorithms adopted in fingerprinting indoor positioning systems», *IGNSS Symposium*, 2013.
- [8] MA, R., GUO, Q., HU, C., & XUE, J. (2015). «An Improved WiFi Indoor Positioning Algorithm by Weighted Fusion». *Sensors*.
- [9] HABIBI LASHKARI, ARASH & PARHIZKAR, BEHRANG & NG AH NGAN, MIKE. (2010). «WIFI-based indoor positioning system». *Computer and Network Technology, International Conference on*. 76-78. 10.1109/ICCNT.2010.33.
- [10] ZOURMAND A., SHENG N., LAI KUN A., ET AL, «Human Counting and Indoor Positioning System Using WiFi Technology» , *IEEE, International Conference on Automatic Control and Intelligent Systems*. October 20, 2018, Shah Alam, Malaysia.
- [11] DAVIES K., JONES I. y SHAPIRO J., «A Bayesian Approach to Dealing with Device Heterogeneity in an Indoor Positioning System», *International Conference on Indoor Positioning and Indoor Navigation*. September 2018, Nantes, France.

- [12] LE DORTZ N, ZETTERBERG P. AND GAIN F., «Wifi Fingerpirnt Indoor Positioning System Using Probability Distribution Comparison», *IEEE ICASSP 2012*
- [13] BEOM-JU SHIN, KWANG-WON LEE, SUN-HO CHOI, ET AL,«Indoor WiFi positining System fior Android Based Smartphone» , *ICTC 2010*
- [14] LU Y., ZHANG D. CHENG Y., ET AL,«An Improved Method and Implementation of Indoor Positioning Fingerprint Matching Localization Based on WLAN», *Springer Nature*,Switzerland AG 2019
- [15] H. HSIEH, S. W. PRAKOSA AND J. LEU, «Towards the Implementation of Recurrent Neural Network Schemes for WiFi Fingerprint-Based Indoor Positioning,» 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 2018, pp. 1-5.doi: 10.1109/VTCFall.2018.8690989
- [16] Y LU ET AL,«Implementation of Fingerprint Matching Method for Indoor Position Based on Virtual Reality Technology», *An Improved Method and Implementation*
- [17] ALARIFI, ABDULRAHMAN & AL-SALMAN, ABDULMALIK & ALSALEH, MANSOUR & AL-NAFESSAH, AHMAD & ALHADHRAMI, SUHEER & A. AL-AMMAR, MAI & AL-KHALIFA, HEND. (2016). «Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances». *Sensors*. 16. 1-36. 10.3390/s16050707.
- [18] LUKASZ ZWIRELLO, TOM SCHIPPER, MARLENE HARTER, AND THOMAS ZWICK, «UWB Localization System for Indoor Applications: Concept, Realization and Analysis», *Journal of Electrical and Computer Engineering*, vol. 2012, Article ID 849638, 11 pages, 2012. <https://doi.org/10.1155/2012/849638>.
- [19] McCULLOCH, W.S. & PITTS, W. «A logical calculus of the ideas immanent in nervous activity» *Bulletin of Mathematical Biophysics* (1943) 5: 115.
- [20] K. E. JEON, J. SHE, P. SOONSAWAD AND P. C. NG, «BLE Beacons for Internet of Things Applications: Survey, Challenges, and Opportunities» in *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 811-828, April 2018. doi: 10.1109/JIOT.2017.2788449 <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8242361&isnumber=8334665>
- [21] R. P AND M. L. SICHITIU, «Angle of Arrival Localization for Wireless Sensor Networks,» 2006 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks, Reston, VA, 2006, pp. 374-382. doi: 10.1109/SAHCN.2006.288442

- [22] LUO, JUAN & YIN, XIXI & ZHENG, YANLIU & WANG, CHUNHUA. (2018). «Secure Indoor Localization Based on Extracting Trusted Fingerprint». *Sensors*. 18. 469. 10.3390/s18020469.
- [23] ZHOU, J & LI, W & JIN, L & CHEN, X. (2015). «Indoor positioning system based on KNN-SVM algorithm.» 43. 517-520. 10.13245/j.hust.15S1123.
- [24] DARI, YOHANES & SUYOTO, SUYOTO & PRANOWO, PRANOWO. (2018). CAPTURE: «A Mobile Based Indoor Positioning System using Wireless Indoor Positioning System.» *International Journal of Interactive Mobile Technologies (iJIM)*. 12. 61. 10.3991/ijim.v12i1.7632.
- [25] REN, JIN & YUNAN, WANG & CHANGLIU, NIU & WEI, SONG. (2019). «A Novel High Precision and Low Consumption Indoor Positioning Algorithm for Internet of Things.» *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2019.2924992.
- [26] CHEN, CHIH-YUNG & YIN, LI-PENG & CHEN, YU-JU & HWANG, REY-CHUE. (2012). «A modified probability neural network indoor positioning technique». *Proceedings - 3rd International Conference on Information Security and Intelligent Control, ISIC 2012*. 317-320. 10.1109/ISIC.2012.6449770.
- [27] MU, ZHOU & YUBIN, XU & LIN, MA. (2009). «ANN indoor position determination based on area correlation in WLAN environment.» *2009 International Conference on Wireless Communications and Signal Processing, WCSP 2009*. 10.1109/WCSP.2009.5371722.
- [28] F MALIK, R & GUSTIFA, R & FARISSI, A & STIAWAN, DERIS & UBAYA, H & AHMAD, MOHD RIDUAN & AL-KHALEEFA, AHMED. (2019). «The Indoor Positioning System Using Fingerprint Method Based Deep Neural Network.» *IOP Conference Series: Earth and Environmental Science*. 248. 012077. 10.1088/1755-1315/248/1/012077.
- [29] ANTONIO GULLI & SUJIT PAL «Deep Learning with Keras» Birmingham, England : Packt Publishing, 2017.
- [30] H. HSIEH, S. W. PRAKOSA AND J. LEU, *"Towards the Implementation of Recurrent Neural Network Schemes for WiFi Fingerprint-Based Indoor Positioning," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 2018, pp. 1-5.*

Glosario

Overfitting Efecto de sobreentrenar un algoritmo de aprendizaje con unos ciertos datos para los que se conoce el resultado deseado, es decir, el sobreajuste ocurre cuando el modelo o el algoritmo se ajusta demasiado bien a los datos, lo que se traduce en que el modelo o el algoritmo muestra un bajo bias y una alta varianza.. 27

Underfitting Ocurre cuando el modelo o el algoritmo de machine learning, es incapaz de rescatar características relevantes de los datos. Ocurre cuando el modelo no se ajusta bien a los datos mostrando una baja varianza, pero un alto bias. Usualmente un subajuste es debido a la excesiva simpleza del modelo o al reducido numero de datos con que se entrena. Ambos, sobreajuste y subajuste llevan a tener predicciones pobres en sets de datos nuevos.. 27