

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN SUPERIOR
DE ENSENADA**



**PROGRAMA DE POSGRADO EN CIENCIAS
EN CIENCIAS DE LA COMPUTACIÓN**

**USO DE REDES NEURONALES CON INFORMACIÓN DE VECINDARIOS
PARA ESTIMAR LA UBICACIÓN DE COMPUTADORAS MÓVILES**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de
MAESTRO EN CIENCIAS

Presenta:

LUIS ADRIÁN CASTRO QUIROA

Ensenada, Baja California, México, Septiembre del 2005.

TESIS DEFENDIDA POR
LUIS ADRIÁN CASTRO QUIROA
Y APROBADA POR EL SIGUIENTE COMITÉ

Dr. Jesús Favela Vara
Director del Comité

Dr. José Antonio García Macías
Miembro del Comité

Dr. Hugo Homero Hidalgo Silva
Miembro del Comité

Dr. Luis Armando Villaseñor González
Miembro del Comité

Dr. Pedro Gilberto López Mariscal
*Coordinador del programa de
posgrado en Ciencias de la
Computación*

Dr. Raúl Ramón Castro Escamilla
Director de Estudios de Posgrado

Septiembre de 2005

RESUMEN de la tesis de **LUIS ADRIÁN CASTRO QUIROA**, presentada como requisito parcial para la obtención del grado de **MAESTRO EN CIENCIAS** en **CIENCIAS DE LA COMPUTACIÓN**. Ensenada, Baja California. Septiembre del 2005.

USO DE REDES NEURONALES CON INFORMACIÓN DE VECINDARIOS PARA ESTIMAR LA UBICACIÓN DE COMPUTADORAS MÓVILES.

Resumen aprobado por:

Dr. Jesús Favela Vara
Director de Tesis

El aumento en el uso de dispositivos de cómputo móvil en diversas áreas de trabajo ha sido notable en los últimos años. Para dar un mejor apoyo al usuario en el desempeño de su actividad, las aplicaciones conscientes del contexto adaptan su comportamiento tomando en consideración, principalmente, la identidad, hora del día, la actividad, así como la localización del usuario.

El problema de la localización de usuarios, en especial, ha sido abordado en los últimos años por varios grupos de investigación, de donde han surgido varios enfoques. De particular interés son los sistemas de localización donde se usan las señales de radiofrecuencia (RF) para estimar la posición, ya que hacen uso de la infraestructura de red inalámbrica existente. La intensidad de la señal RF recibida por los dispositivos móviles se puede considerar como una función de la distancia del dispositivo a múltiples puntos de acceso inalámbricos. Sin embargo, esta relación es muy compleja en interiores, donde la estructura física del edificio, muebles, personas, así como otros dispositivos en el ambiente influyen de manera significativa en la intensidad de la señal recibida.

En este trabajo de tesis se propone el uso de redes neuronales para tratar de encontrar dicha relación e inferir la posición del usuario por lo que se diseñó una aplicación para la toma de muestras empíricas de la intensidad de la señal en diferentes ubicaciones dentro de un hospital. Uno de los resultados más interesantes de este trabajo de tesis es que la distancia de error promedio fue de 0.9131 metros con el uso del vecindarios en contraste con los 1.9827 metros obtenidos sin usar este enfoque.

De manera adicional, se realizaron experimentos para observar la sensibilidad del error con respecto a la cantidad de muestras usadas para estimar la posición, tamaño del conjunto de entrenamiento, la arquitectura de red neuronal utilizada, así como cantidad de puntos de acceso utilizados para la estimación. Los resultados de estos experimentos sugieren algunas guías para la implementación de aplicaciones conscientes de la localización. Finalmente, se diseñó e implementó una aplicación para PDA's para probar el enfoque propuesto en este trabajo de tesis.

Palabras clave: Localización de usuarios, seguimiento de usuarios, cómputo consiente del contexto, cómputo consciente de la localización, redes neuronales.

ABSTRACT of the Thesis of **LUIS ADRIÁN CASTRO QUIROA**, presented as partial requirement to obtain the **MASTER in SCIENCES** degree in **COMPUTER SCIENCE**. Ensenada, Baja California. September 2005.

USING NEURAL NETWORKS WITH NEIGHBORHOOD INFORMATION TO ESTIMATE THE LOCATION OF MOBILE COMPUTERS.

Abstract approved by:

Dr. Jesús Favela Vara
Thesis Director

The use of mobile computing devices in several work fields has increased notably in the last few years. To support such activities context-aware applications adapt their behavior taking into account, mainly, time, identity, task and the location of the user.

A considerable amount of research has been conducted to address the problem of estimating the location of users and, as a result, several approaches have been proposed. Of particular interest are radiofrequency-based (RF) location systems since they use the existent wireless network infrastructure. The RF signal strength, received by mobile devices, can be considered to be a function of the distance from the mobile device to multiple wireless access points. However, this relationship becomes complex indoors, where the received signal strength is affected significantly by walls, furniture, people and other electronic devices in the environment.

In this work, we propose the use of artificial neural networks to obtain that relationship and infer the location of users of mobile computers. For this reason, we designed an application to collect empirical samples of the signal strength at different locations within a hospital. One of the most significant results of this work was the reduction of the average error distance from 1.9827 meters using neighborhood information to 0.9131 meters without the use of this approach.

In addition, several experiments were prepared to study the manner in which the accuracy is influenced by the training set size, the architecture of the neural network, and the number of access points used to estimate the location. Also we found out the minimum number of samples required to provide an acceptable estimation of the location. Results of these experiments suggest some design guidelines when deploying context-aware applications. Finally, an application was designed and implemented to test the approach proposed in this work.

Keywords: user location, user tracking, context-aware computing, location-aware computing, neural networks.

Dedicatoria

*A mis padres, José Luis y Ricela,
por su cariño, comprensión y apoyo*

A mis hermanos Joselo y Adriana

Agradecimientos

Un especial agradecimiento a mi director de tesis, Dr. Jesús Favela Vara, por los invaluable consejos e ideas con que me apoyó durante la realización de este trabajo.

A los miembros del comité Dr. José Antonio García Macías, Dr. Hugo Homero Hidalgo Silva y Dr. Luis Armando Villaseñor González por su aportación a este trabajo, así como por su tiempo para la revisión del mismo.

A mis amigos de la generación Mtn, Burciaga, Tentori, Pedrito, David, Chesar, Leo, Liz, Isabel, Diana, Irma, Elisa, Marcel, El Bob y ah! Si Simón Abel, por los momentos inolvidables en estos dos años de estancia en CICESE..

A los “Hermosillos”: El Kareko, El Manuel y La Uzza por hacer parte de mi estancia algo agradable.

A mis otros amigos de CICESE que aunque no son de la generación también pusieron su granito de arena, cada quien a su modo: Doris, Cynthia y Jorge Niebla.

A “clandestino” que algo tuvo que ver con la finalización de la tesis y, claro esta, no podía pasar por alto a las administraciones del famoso “casino”.

A Patty por su cariño incondicional todo este tiempo.

Y finalmente, pero no por eso el menos importante, al CONACYT por su apoyo económico.

Contenido

	Página
Resumen	i
Abstract	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	v
Índice de figuras	ix
Índice de tablas	xii

Capítulo I Introducción

I.1. Antecedentes.....	1
I.2. Planteamiento del Problema.	2
I.3. Objetivos del Estudio.	3
I.3.1. Objetivo General.	3
I.3.2. Objetivos específicos.....	3
I.4. Metodología.....	4
I.5. Contenido de la Tesis.	4

Capítulo II Computación Consciente del Contexto

II.1. Cómputo consciente del contexto.	7
II.1.1. Introducción.	7
II.1.2. ¿Qué es contexto?	8
II.1.3. Cómputo consciente del contexto.	10
II.2. Tecnologías de localización.	11
II.2.1. Sistemas de localización en exteriores.....	11
II.2.2. Sistemas de localización en interiores.....	12

Contenido (continuación)

Página

II.2.3. Sistemas de RF que usan la infraestructura existente de una WLAN.....	14
II.2.4. Sistemas de localización que usan grafos.	16
II.3. Resumen.....	16

Capítulo III Redes Neuronales

III.1. Introducción a las Redes Neuronales Artificiales.....	18
III.1.1. Estructura básica de una red neuronal.	19
III.1.2. Funciones de activación.....	21
III.1.2.1. Función Escalón.....	21
III.1.2.2. Función Identidad.	22
III.1.2.3. Función Sigmoide.....	22
III.2. Perceptrón.	23
III.3. Arquitecturas de Redes Neuronales.....	24
III.3.1. Redes de Capa Sencilla.....	24
III.3.2. Redes Multi-Capa.	25
III.3.3. Redes Recurrentes.....	27
III.4. Proceso de Aprendizaje.	28
III.4.1. Aprendizaje Supervisado.	28
III.4.2. Aprendizaje No Supervisado.	28
III.4.3. Aprendizaje de Refuerzo.	29
III.5. Redes de Retropropagación.	29
III.5.1. Regla Generalizada Delta	30
III.6. Redes Recurrentes.....	31
III.6.1. Redes de Elman.	32
III.6.2. Redes de Jordan.	33
III.7. Resumen.....	36

Contenido (continuación)

Página

Capítulo IV Recolección de Muestras y Preparación de Experimentos

IV.1. Plataforma Experimental.	36
IV.1.1. Descripción del lugar.....	36
IV.1.2. Descripción del hardware.	37
IV.2. Recolección de muestras.	39
IV.2.1. Aplicación para recolección de muestras.	40
IV.2.2. Agrupamiento de muestras para entrenamiento.	43
IV.3. Experimentos realizados con los datos obtenidos.	44
IV.3.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal	45
IV.3.2. Influencia del tamaño del conjunto de entrenamiento en la estimación de la localización	46
IV.3.3. Influencia del número de muestras en la estimación de la localización.....	49
IV.3.4. Influencia del número de puntos de acceso en la estimación de la localización. ...	49
IV.3.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.	50
IV.3.6. Entrenando la red neuronal con información de vecindarios.	50
IV.4. Resumen.	52

Capítulo V Resultados y Discusión

V.1. Resultados de los Experimentos.	53
V.1.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal.....	53
V.1.2. Influencia del tamaño del conjunto de entrenamiento en la estimación de la localización.	57

Contenido (continuación)

Página

V.1.3. Influencia del número de muestras utilizadas en tiempo real para la estimación de la localización.	59
V.1.4. Influencia del número de puntos de acceso en la estimación de la localización.	61
V.1.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.	65
V.2. Entrenando la red neuronal con información de vecindarios.....	67
V.3. Discusión de los Resultados obtenidos.	74
V.3.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal.	74
V.3.2. Influencia del tamaño del conjunto de entrenamiento en la estimación de la localización.	75
V.3.3. Influencia del número de muestras utilizadas en tiempo real para la estimación de la localización.	76
V.3.4. Influencia del número de puntos de acceso en la estimación de la localización.	77
V.3.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.	79
V.3.6. Entrenando la red neuronal con información de vecindarios.....	80
V.4. Resumen.....	81

Capítulo VI Conclusiones y Trabajo Futuro

VI.1. Conclusiones.....	82
VI.2. Aportaciones.	83
VI.3. Limitaciones.....	84
VI.4. Trabajo futuro.	85
Referencias.....	86
Apéndices.....	89

LISTA DE FIGURAS

Figura		Página
1	La imagen es un ejemplo de un patrón complejo	19
2	Este diagrama muestra la estructura típica de una red neuronal	20
3	Función de activación escalón	22
4	Función de activación identidad	22
5	Función de activación logística sigmoïdal	23
6	Función de activación tangente hiperbólica	23
7	Red neuronal de capa sencilla	25
8	Red neuronal multi-capa	26
9	Red neuronal recurrente simplificada	28
10	Red Elman simplificada	33
11	Red Jordan simplificada	34
12	Mapa del edificio donde fueron recolectadas las muestras	37
13	Conectividad de la red inalámbrica en el hospital.	38
14a	Puntos de acceso Airport Extreme usado en las mediciones	38
14b	Puntos de acceso Airport Express usado en las mediciones	38
15	Aplicación para Pocket PC con la que se hizo el registro de las intensidades de la señal de los diferentes puntos de acceso	40
16	Marcas que se usaron para la toma de mediciones	41
17	Intensidades de la señal en el pasillo del área de medicina interna	42
18a	División de los puntos para entrenamiento y prueba	44

LISTA DE FIGURAS (continuación)

Figura		Página
18b	Cada muestra es agrupada de 4 formas diferentes para entrenamiento y prueba.	44
19	Pedestal usado en la toma de lecturas para este experimento	45
20	Localización de los puntos de acceso, así como el pasillo donde fueron tomadas las mediciones	46
21	Figura que muestra los puntos donde se detectaron puntos con lecturas irregulares	49
22	Vecinos (círculos) considerados para entrenamiento de una posición específica (rectángulo)	51
23	Punto de acceso BS1 y la medición de la señal en altura y distancia	54
24	Punto de acceso BS2 y la medición de la señal en altura y distancia	55
25	Punto de acceso “colaborativos” y la medición de la señal en altura y distancia	56
26	Distribución acumulada para el conjunto de 80, 60, 40 y 20 % en los datos de entrenamiento	58
27	Distribución empírica acumulada para diferentes conjuntos de prueba	60
28	Distribución empírica acumulada para tres puntos de acceso	62
29	Distribución empírica acumulada de 4, 5 y 6 puntos de acceso	63
30	Distribución empírica acumulada de 7 arquitecturas diferentes	65
31	Variación de valores de X y Y en diferentes instantes de tiempo	67
32	Distribución acumulada del error del entrenamiento con y sin vecinos.	69

LISTA DE FIGURAS (continuación)

Figura		Página
33	Rutas sobre las que se tomaron mediciones de manera continua en el hospital	70
34	Distribución empírica acumulada del error con y sin vecinos con rutas reales	71
35	Variación de la distancia entre estimaciones consecutivas usando información de vecindarios, así como la variación de la distancia entre estimaciones consecutivas sin usar dicha información de vecindarios.	72
36	Ruta real usada para la prueba de los dos enfoques: sin usar posición anterior (arriba) y usando una posición anterior como parte de la entrada de la red (abajo)	73
37	Pantalla principal del sistema	91
38	Pantalla de configuración	91
39	Barra de progreso	92
40	Mensaje mostrado cuando la tarjeta inalámbrica esta inactiva	92
41	Mensaje para indicar que ya existe el archivo	92
42	Mensaje de confirmación para iniciar con mediciones	93
43	Mensaje de confirmación para guardar mediciones	93
44	Captura de pantalla de la aplicación para la estimación de la localización	94
45	Plano arquitectónico del hospital donde se tomaron las mediciones	95
46	Plano arquitectónico del hospital mostrando las rutas donde fueron tomadas las mediciones.	96

LISTA DE TABLAS

Tabla		Página
I	Distribución de frecuencias en los puntos de acceso	39
II	Tabla que muestra la distribución de puntos en los diferentes conjuntos	47
III	Análisis de Varianza del punto de acceso BS1	55
IV	Análisis de Varianza del punto de acceso BS2	56
V	Análisis de Varianza del punto de acceso Colaborativos	57
VI	Tabla comparativa de los diferentes conjuntos de entrenamiento (80%, 60%, 40% y 20%)	59
VI	Tabla comparativa con los diferentes conjuntos de datos de prueba	61
VIII	Tabla comparativa usando diferente número de puntos de acceso	63
IX	Tabla comparativa de 4, 5 y 6 puntos de acceso	64
X	Tabla comparativa de 7 arquitecturas de red	66
XI	Resultados del entrenamiento con coordenadas vecinas	68
XII	Resultados del entrenamiento con coordenadas vecinas	71

Capítulo I.

Introducción.

I.1. Antecedentes.

La aparición de las computadoras móviles, portátiles ó de bolsillo, así como los avances de la tecnología inalámbrica en los últimos años, han sido un detonante en la variación gradual de los paradigmas de trabajo. La movilidad es inherente a cierto tipo de actividades humanas, por lo que la tecnología de cómputo se ha ido adaptando a ellas para apoyar estas actividades de una mejor manera.

A medida que el cómputo va quedando inmerso en dichas actividades con alto grado de movilidad, van surgiendo nuevas necesidades generadas a partir de la aparición de una nueva variable contextual: la ubicación de los usuarios. Es decir que el saber dónde están los usuarios en determinado momento, se ha vuelto relevante para algunas aplicaciones donde el tipo de información requerida esta ampliamente relacionada con el lugar donde se encuentra el usuario.

Hasta el momento, no se ha encontrado un método de localización adecuado en términos de costo y beneficio. El trabajo previo en el área nos muestra resultados muy buenos con sensores de ultrasonido pero con altos costos de instalación y mantenimiento del equipo especializado (Harter et al., 1999; Priyantha et al., 2000). Otros, como es el caso de *active badge* (Want et al., 1992), pionero en el área de localización de usuarios, también muestra buenos resultados aunque con los mismos inconvenientes que los anteriores. Existen otras tecnologías de las cuales se ha explorado su uso en la ubicación de usuarios. Sin embargo, todas tienen inconvenientes ya sea en términos de escalabilidad, costo, ó algún otro factor importante (Hightower y Borriello, 2001).

Por otro lado, se han hecho esfuerzos en el sentido de aprovechar la infraestructura de red inalámbrica existente de modo que no se tenga que invertir en gran medida en ese rubro. Uno de esos ejemplos es RADAR, un proyecto desarrollado en *Microsoft Research*, en el cual se hace uso de la intensidad de la señal en conjunto con la técnica de los vecinos cercanos (Bahl y Padmanabhan, 2000), obteniéndose un error promedio del orden de los 2 metros. De igual forma, en la Universidad de Trento, en Italia, y en la Universidad Tecnológica de la Mixteca, en Oaxaca, se han hecho estudios usando redes neuronales, al igual que en este trabajo, con errores promedio de 1.8 metros (Battiti et al., 2002) y 2.2 metros en el caso de la segunda (Martinez et al., 2004).

En este trabajo, se propone el uso de redes neuronales para transformar la intensidad de la señal recibida de diferentes puntos de acceso en la localización del usuario donde, además, se plantea la inclusión del vecindario de la posición a estimar para tratar de reducir el error máximo que se puede obtener. El uso de vecindarios se justifica por el hecho que la posición que se esta estimando es una posición muy cercana a donde se estuvo la última vez. De esta forma, se busca que los valores que se obtienen en la estimación sean menos variables.

I.2. Planteamiento del Problema.

El problema de localización de usuarios se puede reducir a dos aspectos: costo y beneficio. En cualquier organización lo que se desea es invertir menos y obtener más a cambio. Aunque la efectividad de cualquier técnica de localización dependerá directamente de la naturaleza de la aplicación para la cual se desea su uso. Para algunas aplicaciones un error aceptable en cuestiones de localización podría ser de 5 metros, no así para otras donde 1 metro puede hacer mucha diferencia, como es el caso de un hospital donde 1 metro puede ser la diferencia de estar junto a la cama de uno u otro paciente y las consecuencias de un error de este tipo podrían ser nefastas.

En este trabajo, se propone usar la infraestructura de red existente, es decir los puntos de acceso inalámbricos que ya existen y que ofrecen sus servicios de red inalámbrica a los usuarios.

La ventaja de este enfoque es el hecho que no se tiene que invertir en sensores adicionales y equipo especializado para obtener resultados aceptables. Es posible hacer esto mediante la medición de la señal inalámbrica desde diferentes puntos del edificio. La idea aquí es que la intensidad de la señal es función de la distancia a la que se encuentra un usuario alejado del punto de acceso. Sin embargo, esto no es completamente cierto con señales de radiofrecuencia en interiores, donde la intensidad de la señal se ve fuertemente afectada por el ambiente. Es decir que, las variaciones de la señal se ven influenciadas por la estructura física del edificio, los materiales con los cuales esta construido, el ancho de las paredes, los muebles que están dentro de él, las personas que transitan dentro de él, así como algunos otros factores externos que afectan en menor medida como lo es la temperatura, humedad, hora del día, etc.

El problema de localización de entidades móviles sigue aún abierto, se ha dedicado mucha investigación tratando de sobrepasarlo. Sin embargo, existe la confianza que se pueden mejorar los resultados hasta ahora publicados y que el problema seguirá siendo abordado en los próximos años.

I.3. Objetivos del Estudio.

En esta sección se presentan los objetivos de este trabajo de tesis.

I.3.1. Objetivo General.

Proponer un método práctico y eficiente para la estimación de la ubicación de computadoras móviles con información de vecindarios usando redes neuronales.

I.3.2. Objetivos específicos.

- Definir un método que permita medir la intensidad de la señal desde diferentes puntos fijos del edificio, así como también tomarlos de una manera continua considerando el tiempo.
- Usar información del pasado para tratar de mejorar los resultados en la estimación de la localización.
- Comparar y evaluar los resultados del método propuesto con métodos reportados en la literatura.
- Diseñar una aplicación que permita probar el enfoque propuesto en este trabajo de tesis.

I.4. Metodología.

La primera parte del proyecto de tesis fue la revisión de literatura, donde se estudiaron proyectos realizados en el área de cómputo consciente del contexto, específicamente en cómputo consciente de la localización. De este modo, se obtuvo una visión general del área en la que el proyecto está inmerso.

Al mismo tiempo que se estaba revisando el trabajo previo en el área, se desarrolló una aplicación que facilitó la toma de mediciones empíricas de la intensidad de la señal en diferentes puntos del edificio.

Una vez con las muestras de la señal, se empezó a estudiar la manera en que dichas muestras dieran mejores resultados al momento de hacer el análisis.

A su vez, se revisó también teoría de redes neuronales, así como trabajo previo en el área. Para verificar cuál sería la manera más adecuada de entrenar una red neuronal con las muestras obtenidas.

Cuando se encontraron resultados adecuados, se implementó la red neuronal en una computadora de bolsillo, de modo que la estimación de la localización se pudiera hacer en tiempo real en dicho dispositivo.

I.5. Contenido de la Tesis.

El contenido y organización de este trabajo se detalla en esta sección. En este primer capítulo se presenta una introducción general donde se trata de ubicar al lector en términos generales de lo que se estará hablando el resto de este trabajo, así como el planteamiento del problema y los objetivos que se siguieron para lograr los resultados esperados.

En el capítulo II se presenta una introducción al cómputo consciente del contexto, que es el área donde se ubica este trabajo de tesis, específicamente dentro de cómputo consciente de la localización. Se describe el trabajo relacionado en esa área de investigación, y se ofrece, también, un bosquejo general de lo que trata dicha área de investigación, así como una idea de la utilidad de que puede ser objeto para determinadas tareas.

En el capítulo III, se presenta una introducción a las redes neuronales, de modo que al finalizar ese capítulo se tenga una noción de la manera en que están constituidas, así como la manera en que aprenden a través de ejemplos.

En el capítulo IV se describe el lugar de trabajo donde se iniciaron las primeras actividades de campo experimental encaminadas a lograr los objetivos propuestos, en este caso el lugar del que se habla es un hospital. Además, se muestra una aplicación diseñada para la toma de lecturas en el lugar. De igual forma, se explica la manera en que las muestras fueron tomadas en el hospital y el método que se siguió para agruparlas y promediarlas para tratar de obtener mejores resultados en la estimación de la localización.

En este mismo capítulo, se explican los experimentos de los que se obtuvieron los resultados más interesantes con las muestras recolectadas en el hospital.

En el capítulo V, se presentan los resultados de los experimentos descritos en el capítulo IV, así como una discusión desde diferentes puntos de vista de los resultados obtenidos.

En el capítulo VI se presentan las conclusiones del trabajo realizado, así como las contribuciones del mismo. Además, en el trabajo futuro se hacen propuestas para la mejora de los resultados obtenidos.

Finalmente, en la sección de apéndices se explica el funcionamiento de las aplicaciones creadas para la toma de mediciones así como para la prueba de la red neuronal.

Capítulo II.

Computación Consciente del Contexto.

En este capítulo se presenta una introducción al Cómputo Consciente del Contexto (*Context-Aware Computing*). Uno de los objetivos de este capítulo es el de dar una idea general de lo que es el Cómputo Consciente del Contexto y cuáles pueden ser sus alcances. En particular, se resalta al Cómputo Consciente de la Localización (*Location-Aware Computing*), que es donde se centra este trabajo de tesis. Además, otro aspecto que se trata en este capítulo es el trabajo previo en el área de cómputo consciente de la localización, de los cuales se presenta una breve reseña.

II.1. Cómputo consciente del contexto.

En esta sección se presenta un bosquejo general del cómputo consciente del contexto. Así como la definición de cómputo consciente de la localización

II.1.1. Introducción.

Durante los últimos 20 años, desde la aparición de las primeras PC's hemos estado acostumbrados a un ambiente donde el cómputo siempre se ha hecho en computadoras de escritorio y portátiles (Moran y Dourish, 2001).

Sin embargo, en los últimos años, han surgido nuevas áreas de estudio que rompen con ese paradigma. Una de esas áreas es el cómputo ubicuo a través de la cual el poder del cómputo ha sido trasladado a una gran diversidad de nuevos dispositivos mucho más ligeros y compactos que las computadoras portátiles. Gran parte de esa versatilidad se debe a la tecnología de comunicación inalámbrica que permite conectar a toda esta amplia gama de dispositivos.

No obstante, con el advenimiento de la tecnología inalámbrica y el cómputo ubicuo surgen una serie de nuevos retos a vencer, así como nuevas necesidades que satisfacer. Una de ellas es la información contextual. La cual se describe más a fondo en la siguiente sección.

II.1.2. ¿Qué es contexto?

Cuando los humanos nos comunicamos unos con otros, somos capaces de transferir ideas de una manera relativamente sencilla y esto se debe principalmente a tres factores: la riqueza del lenguaje con el cual nos comunicamos, el entendimiento mutuo de la forma en que funciona el mundo y el entendimiento implícito de cómo funcionan las cosas (Dey, 2001). Es decir, los humanos tenemos habilidades innatas en el manejo del *contexto* en el cual nos desenvolvemos. Siempre usamos información del ambiente para el manejo de diversas situaciones.

Sin embargo, cuando se trata de transmitir ideas entre humanos y computadoras las cosas se complican un poco. La idea clásica que tenemos en mente cuando hablamos de interacción con una computadora es la de introducir datos a la computadora y esperar un resultado. Es decir, la computadora actúa en base a una entrada explícita por parte del usuario.

El caso contrario es cuando usa información contextual para proveer de información relevante al usuario basándose en las características actuales del medio físico, así como de los medios electrónicos. Se define *contexto* como cualquier interacción entre los usuarios, las aplicaciones y el medio circundante (Dey y Abowd, 2001).

Una de las cosas más relevantes del contexto es saber cómo entenderlo y manejarlo, pero también saber cómo usarlo de una manera que sea transparente para el usuario, de modo que se pueda afectar el comportamiento de la aplicación y, por consecuencia, el ambiente en el cual esta inmerso el usuario.

El concepto ‘consciente del contexto’

Una de las primeras definiciones de **conciencia de contexto** es la que se presenta en (Schilit et al., 1994), en la cual los autores afirman que el *software consciente del contexto* es “aquel que se adapta de acuerdo a la localización física del usuario, personas que estén cerca, computadoras y/o dispositivos accesibles, así como los cambios que se presentan en tales objetos a través del tiempo”. De igual forma, otro de los conceptos clásicos es el presentado en (Brown et al., 1997) en el que introduce el concepto de *aplicaciones conscientes del contexto*. Los autores las definen como “aplicaciones que cambian su comportamiento de acuerdo al contexto” y en el que ellos llaman contexto a: localización, hora del día, temporada del año y temperatura.

Existen muchas otras definiciones de contexto, cada autor tiene su propia concepción de lo que es *contexto*, dependiendo de la naturaleza de la aplicación. Sin embargo, hay una definición en (Dey y Abowd, 1999) que es un poco más general y, gracias a eso, encierra a todas las demás:

“Contexto es cualquier información que se puede usar para caracterizar la situación de una entidad. Una entidad es una persona, lugar u objeto que es considerado relevante a la interpretación entre el usuario y una aplicación, incluyendo el usuario y la aplicación misma”.

Esta definición no es tan específica como las presentadas anteriormente y, sin embargo, se puede aplicar en muchas de las situaciones. Por ejemplo, el software de un teléfono celular para determinar cuándo es conveniente timbrar o no. Como entidades, en este ejemplo, tenemos al usuario y la aplicación. La información de la que habla la definición en este caso podría ser iluminación, carga de la batería, nivel de ruido, etc. La iluminación podría ser un poco relevante ya que puede indicar si se está dentro o fuera de un edificio, pero podría tener menos importancia si se carga dentro de la bolsa. En el caso de la pila podría ser relevante, ya que el vibrar y timbrar al mismo tiempo puede gastar más pila y por último, en el caso del ruido pudiera ser un elemento muy importante ya que si no hay mucha iluminación y sólo está hablando una persona o dos, pudiera pensarse que es una reunión y preferiría que el teléfono sólo vibrara. Si por el contrario el teléfono advierte

que el usuario esta en un carro en movimiento, por el ruido del motor, pudiera ser mejor que timbrara o mostrara algún elemento visual, como podrían ser luces para llamar la atención.

En la siguiente sección se hablará de cómo esa consciencia del contexto puede ser usada en muchos ambientes donde la computacion esta inmersa y como es que cambia la naturaleza de las aplicaciones.

II.1.3. Cómputo consciente del contexto.

Una definición clásica del cómputo consciente del contexto es la encontrada en (Dey, 2001), que dice textualmente:

Un sistema es consciente del contexto si usa información contextual para proveer información y/o servicios relevantes al usuario, donde la relevancia depende de la tarea que esté haciendo el usuario.

El concepto de información contextual va implícito en la definición anterior. Como se puede ver, es un amplio espectro de factores que pueden ser relevantes o no para una persona en una tarea determinada. Es por eso que la complejidad de los ambientes conscientes del contexto es de gran magnitud. Para alcanzar la visión que se tiene y se espera de un ambiente consciente de contexto aún falta mucho. Se tienen grandes problemas que atacar, entre los que destacan la privacidad del usuario (Tentori, 2005), los tecnológicos y, además, romper con los paradigmas actuales de uso de la tecnología.

Como se vio, información contextual es toda aquella que resulta relevante dependiendo de la actividad que este realizando el usuario. Un tipo de esa información clásica en cómputo consciente del contexto es la *localización de entidades* dentro de un área determinada. Esto con la finalidad de proveer dicha localización como herramienta de decisión a la aplicación para presentar información relevante al usuario.

En la siguiente sección se presentan una serie de tecnologías y de sistemas que han sido propuestos para tratar de estimar la posición de entidades, algunas de ellas innovadoras en su tiempo y con muy buenos resultados.

II.2. Tecnologías de localización.

Esta sección presenta una serie de tecnologías desarrolladas por diversos grupos de investigación para tratar de atacar el problema de la localización de entidades, uno de los problemas más populares en la computación consciente del contexto.

II.2.1. Sistemas de localización en exteriores.

En este apartado se describen sistemas que abordan el problema de la localización de usuarios en exteriores, es decir, fuera de las paredes o la estructura física de un edificio. La mayoría de ellos a gran escala.

GPS (Sistema de Posicionamiento Global).

Sin lugar a dudas, es el sistema de localización más ampliamente difundido y conocido. Debido a la comercialización de los dispositivos, su uso esta siendo cada vez más común e incluso en combinación con otro tipo de tecnologías como lo es la telefonía celular.

El sistema de localización GPS o Sistema de Posicionamiento Global es un sistema basado en una red de 24 satélites establecida, en sus orígenes, con fines militares por el Departamento de la Defensa de Estados Unidos (Getting, 1993). Los receptores GPS reciben señales de dichos satélites, las cuales son procesadas y usadas para estimar la posición del usuario mediante triangulación. El error de la estimación en la localización de esta tecnología es alrededor de 10 metros, sin embargo no funciona en interiores debido a que la señal no penetra las paredes de los edificios.

E911.

La Comisión Federal de Comunicaciones de Estados Unidos (US FCC)¹ obliga, por medio de una iniciativa de ley, a los proveedores de servicios de telefonía inalámbrica a desarrollar una forma de localizar cualquier teléfono celular que marque el número de emergencia 911.

El requerimiento por parte de la US FCC es que el dispositivo móvil pueda ser ubicado con un error estimado de 150 metros en un esquema basado en dispositivos con GPS y hasta 300 metros en un esquema donde la estimación es parte de la infraestructura de la red celular. No es un sistema de localización como tal todavía, sin embargo, se espera que dentro de poco tiempo ya exista la infraestructura necesaria para soportarlo.

II.2.2. Sistemas de localización en interiores.

Active Badge.

Es uno de los primeros sistemas de localización y se desarrolló en lo que hoy se conoce como los laboratorios AT&T en Cambridge (Want et al., 1992). Cada persona ó entidad a ser localizada usa un dispositivo que emite un código único a través de pulsos infrarrojos cada 15 segundos por un periodo de una décima de segundo. Las señales emitidas son recolectadas por una red de sensores posicionados en todo el edificio y procesadas por un servidor central. Esta tecnología tiene el inconveniente de que en un cuarto con ventanas el rendimiento puede bajar debido a que la luz solar afecta considerablemente los pulsos infrarrojos, además de que su alcance es relativamente pequeño.

Active Bat.

En los mismos laboratorios AT&T en Cambridge unos años más tarde se desarrolló una técnica de localización que usa el ultrasonido (Harter et al., 1999). Cada persona usa un pequeño dispositivo que emite pulsos ultrasónicos como respuesta a un mensaje enviado

¹ US FCC, <http://www.fcc.gov/911/enhanced/>, Octubre del 2004.

por una estación base. El dispositivo emite un pulso ultrasónico a una red de sensores en el techo, los cuales envían los datos recibidos a una unidad central para hacer la triangulación de los datos. La triangulación puede hacerse debido a que el tiempo que le toma a la señal en llegar se puede traducir en una distancia debido a que la velocidad en que viaja el sonido es prácticamente la misma bajo ciertas condiciones. La precisión de esta técnica es excelente, ya que en el 95% de los casos el error fue de aproximadamente 9 cm. Sin embargo, los costos de implantación y mantenimiento de esta tecnología son relativamente altos debido a la infraestructura que necesita para operar.

Cricket.

Desarrollado en el MIT, Cricket es un sistema que complementa al Active Bat System (Priyantha et al., 2000). La diferencia se halla en que el emisor y receptor de los pulsos ultrasónicos radican en los dispositivos móviles. De este modo, cada dispositivo realiza sus cálculos para sacar su posición. Además, usa radiofrecuencia para sincronizar los tiempos de los dispositivos. Cricket, aunque se basa en Active Bat, en la práctica es mucho menos preciso (regiones de 4x4 pies cuadrados). Sin embargo, una de las ventajas es la privacidad que ofrece al realizarse los cálculos en el dispositivo móvil, aunque eso implica una carga adicional de procesamiento para los dispositivos móviles.

Motion Star.

El uso de ondas electromagnéticas es un método clásico para el seguimiento de la posición (Raab et al., 1979). Estos sistemas usan pulsos electromagnéticos que se reciben desde antenas en posiciones fijas. El sistema calcula la posición y orientación de las antenas midiendo la respuesta en tres ejes ortogonales del pulso transmitido. Sistemas como éste, estiman de manera precisa la posición relativa a la antena transmisora. La precisión y exactitud de los resultados en este tipo de sistemas tienen resoluciones espaciales del orden de 1mm, 1ms temporales y, finalmente, 0.1° de capacidad de orientación. Algunas de las desventajas son los altos precios de implementación y, además, la necesidad de conectar el objeto a seguir con una unidad de control. Otras desventajas son que los objetos deben estar

de 1 a 3 metros del transmisor y la exactitud se reduce considerablemente con la presencia de objetos metálicos.

Smart Floor.

Esta tecnología se desarrolló en Georgia Tech (Orr y Abowd, 2000), la cual usa sensores de presión en el piso para seguimiento y reconocimiento de usuarios. Las desventajas de esta técnica son los costos en que se incurren en la instalación de estos sensores de presión en todos los pisos donde desea que se implemente además de la escalabilidad.

Easy Living.

Este sistema basado en tecnología de visión por computadora fue desarrollado en los laboratorios de investigación de Microsoft. En él se usan cámaras 3D para proporcionar una capacidad de localización basada en visión estéreo (Brumitt et al., 2000). Una de las desventajas de este tipo de sistemas es el poder de procesamiento requerido para el análisis de las imágenes capturadas por las cámaras.

II.2.3. Sistemas de RF que usan la infraestructura existente de una WLAN.

En este tipo de sistemas se hace uso de la infraestructura de red inalámbrica existente, es decir que no se requiere de gasto adicional en equipo especializado y para muchas organizaciones esto puede ser adecuado dados los gastos adicionales en mantenimiento. Por otro lado, el uso de las redes Wi-Fi es cada día más común. En unos años, asistir a un lugar público donde no exista este tipo de infraestructura resultará casi inconcebible. Este trabajo de tesis entra en esta categoría.

Entre los proyectos de mayor relevancia en esta área destacan el proyecto RADAR de los laboratorios de investigación de Microsoft, así como algoritmos desarrollados en la Universidad de Carnegie Mellon y, finalmente, un enfoque con redes neuronales de la Universidad de Trento en Italia. En los siguientes párrafos se describen dichos sistemas.

RADAR.

Como se mencionó anteriormente, es un proyecto de Microsoft Research que usa la tecnología IEEE 802.11b para redes inalámbricas (WLAN). En este proyecto se usa la intensidad de la señal recibida (RSSI, *Received Signal Strength Indicator*) y la razón Señal-a-Ruido (SNR, *Signal-to-Noise Ratio*), las cuales son traducidas a una posición 2D. La ventaja de este enfoque es que se necesita solo la infraestructura de red existente, sin gastos adicionales en instalación y mantenimiento de sensores (Bahl y Padmanabhan, 2000).

CMU-PM y CMU-TMI

Estos algoritmos fueron propuestos por la Universidad de Carnegie Mellon y en ellos se usa la intensidad de la señal para entrenar dichos algoritmos. La cantidad de datos de entrenamiento depende de cual de los dos algoritmos es el que se use.

CMU-PM usa el emparejamiento de patrones, es decir que se comparan las señales recibidas en el dispositivo móvil con las señales usadas en el entrenamiento en posiciones físicas conocidas para determinar la posición del dispositivo, las más parecidas se convierten en la posición estimada.

CMU-TMI necesita, por el contrario, conocer la ubicación física de todos los puntos de acceso. Se utiliza, además una función para convertir la señal a distancia, dicha función se creó en base a observaciones. Con base a esta fórmula se crean contornos alrededor de los puntos de acceso y se buscan las intersecciones entre dichos contornos, se determinan un grupo de probables posiciones. Una vez hecho esto, se promedian y se obtiene la posición del dispositivo en el espacio de señal, después se transforma dicha posición a su posición física estimada (Smailagic y Kogan, 2002).

Redes Neuronales.

La Universidad de Trento en Italia utilizó las redes neuronales para hacer una transformación no lineal de la intensidad de la señal recibida de 3 puntos de acceso a coordenadas físicas en 2D (Battiti et al., 2002). Se tienen que tomar mediciones empíricas

de la señal en el lugar donde se desea efectuar la estimación para después entrenar la red neuronal con los datos obtenidos. Al final, se prueba la red neuronal con datos diferentes a los de su entrenamiento.

Este trabajo de tesis presenta un enfoque similar a este trabajo, la diferencia principal radica en que se cambia la forma de entrenar la red neuronal para considerar posiciones pasadas de la ubicación del usuario.

II.2.4. Sistemas de localización que usan grafos.

Existen otros sistemas, como lo es SmartMoveX de Microsoft (Krumm et al., 2002) y el sistema de seguimiento de usuarios Voronoi de la Universidad de Washington (Liao et al., 2003) que hacen el uso de grafos para restringir el paso de usuarios a través de paredes. De este modo, la forma que toma el grafo queda restringida al espacio libre en el edificio.

La diferencia básica entre estos dos sistemas es que en el caso de Microsoft hacen uso de cadenas de Markov como dinámica del sistema y que el espacio de localización queda *discretizado*, ya que los nodos son las únicas ubicaciones posibles donde puede ser estimada la ubicación. Cada nodo puede ser asignado a un cuarto ó un pasillo, dependiendo de la granularidad que se requiera. (Krumm et al., 2002).

Por otro lado, en la Universidad de Washington usan filtros bayesianos para la estimación de la posición. Además, una ventaja con respecto al anterior es que la estimación puede darse a lo largo de los arcos del grafo. Es decir, que no queda limitado a la cantidad de nodos que tenga el grafo (Liao et al., 2003).

En ambos casos, la estimación de la posición mejoro como consecuencia del uso de dichos grafos.

II.3. Resumen.

En este capítulo se presentaron conceptos introductorios para situar al lector en el dominio de aplicación que se esta hablando. Al principio del capítulo se introduce el

concepto de contexto, y cómo ese concepto puede estar inmerso en muchas aplicaciones, con lo cual se llega al cómputo consciente del contexto, el cual se explica brevemente y se provee un ejemplo.

Finalmente, se presenta el cómputo consciente de la localización y algunos de los trabajos más relevantes en el área desarrollados por importantes grupos de investigación, de los cuales se presenta un bosquejo.

En este capítulo, se habla de donde esta situada esta tesis, es decir del tipo de trabajo que se esta desarrollando. Dentro del cómputo consciente de la localización, estimar la posición de un dispositivo móvil usando redes neuronales tomando en cuenta posiciones del pasado del usuario.

En el siguiente capítulo se presenta una introducción a las redes neuronales, en la cual, de manera breve y concisa se explicará la forma en que funcionan y la manera en que *aprenden* dichas redes a través de ejemplos.

Capítulo III.

Redes Neuronales.

Tratar de encontrar una relación entre la distancia y la intensidad de la señal recibida es una función muy compleja en interiores. La estructura física del edificio, los muebles, y las mismas personas afectan considerablemente la fuerza de la señal de una manera difícil de predecir. Es por eso que las redes neuronales representan una buena alternativa para tratar de obtener esa relación ya que pueden aprender de ejemplos y transformar patrones de entrada en patrones de salida aún cuando la relación entre estos conjuntos de datos no pueda ser modelada por un proceso matemático.

En este capítulo se representa una introducción a las redes neuronales artificiales. Se responden preguntas como ¿Qué son las redes neuronales? ¿Cómo están formadas? ¿Cuál es su proceso de aprendizaje?

III.1. Introducción a las Redes Neuronales Artificiales.

El hombre siempre ha querido caracterizar problemas del mundo real a través de las computadoras. Algunos problemas son tan complejos, que para una computadora convencional no es posible resolver, problemas no matemáticos, como lo es el reconocimiento de imágenes y patrones (Freeman y Skapura, 1991).

En la figura 1 podemos apreciar un perro dálmata con su cabeza inclinada hacia abajo. Este es un ejemplo de un patrón muy complejo para tratar de comprenderlo con un programa computacional convencional ¿es posible distinguir entre el perro y su entorno? ¿Por qué al observar la imagen podemos descubrir relativamente fácil que hay un perro en ella? Difícilmente una computadora convencional podría hacer esta distinción.

Los componentes computacionales son, por varios órdenes de magnitud, más rápidos que nuestro sistema neurobiológico. Sin embargo, la estructura de nuestro cerebro y

la manera en que están interconectadas las neuronas le dan cierta ventaja sobre las computadoras más avanzadas de nuestros días (Freeman y Skapura, 1991).

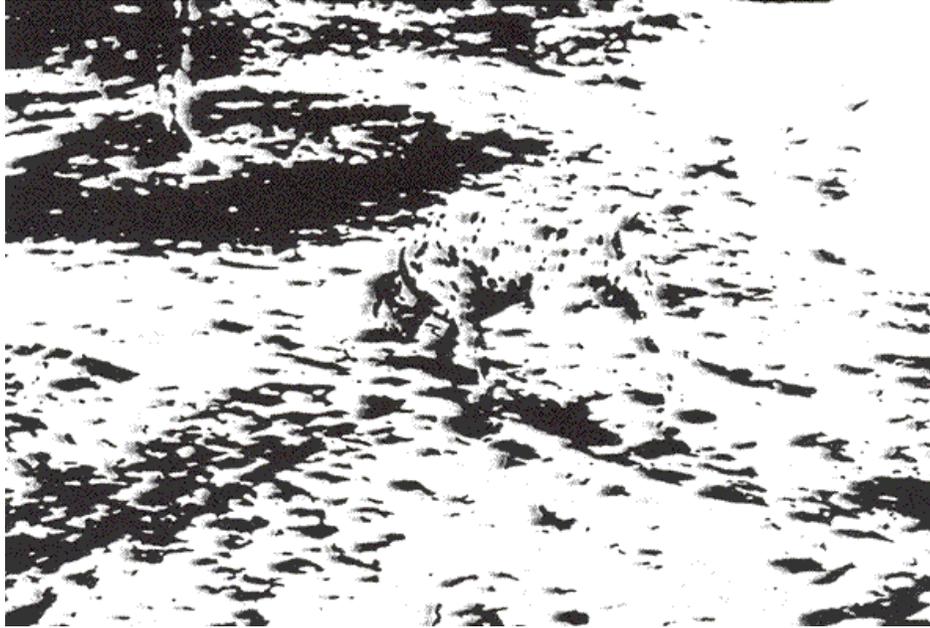


Figura 1. La imagen es un ejemplo de un patrón complejo. *Fuente: (Freeman y Skapura, 1991)*

Por ende, las ciencias computacionales han adoptado la teoría de la fisiología para modelar ciertos problemas del mundo real y tratar de resolverlos de la forma en que lo hace el cerebro humano. De esta manera, problemas complejos pueden ser resueltos, hasta cierto punto, de una manera adecuada.

Los Sistemas de Neuronas Artificiales (ANS, por sus siglas en inglés) nacen a raíz de esta problemática. Algunos de los problemas para los que más comúnmente se usan las redes neuronales son para el reconocimiento de patrones, clasificación de datos y predicción de valores futuros.

III.1.1. Estructura básica de una red neuronal.

La estructura básica de una red neuronal consta de una colección de *procesadores paralelos* conectados juntos en la forma de grafo dirigido, organizado de tal forma que la

estructura de la red se presta para resolver el problema en cuestión (Freeman y Skapura, 1991). Si se observa la figura 2, se puede apreciar que se representan esquemáticamente cada **elemento de procesamiento** (o **unidad**) en la red como un **nodo**, con conexiones entre ellos representados por los **arcos**. La dirección en que fluye la información se indica a través de flechas en las conexiones.

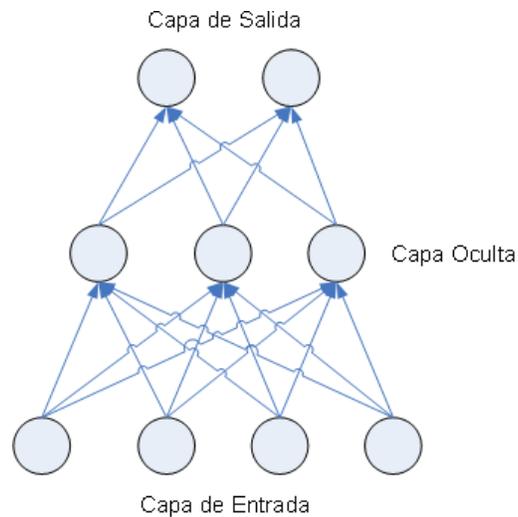


Figura 2. Este diagrama muestra la estructura típica de una red neuronal

Las unidades de procesamiento están unidas a través de conexiones, cada una de las cuales tiene un peso asignado a ella.

Una red neuronal puede definirse de la siguiente manera (Haykin, 1999):

Una red neuronal es un procesador distribuido paralelo formado por simples unidades de procesamiento propensa, de manera natural, a almacenar conocimiento experimental y hacerlo disponible para su uso. La red neuronal se asemeja al cerebro humano en dos aspectos:

1. *El conocimiento se adquiere por un proceso de aprendizaje*
2. *Las fuerzas de conexión entre neuronas, conocidas como pesos sinápticos, son utilizados para almacenar el conocimiento.*

La neurona es la unidad de procesamiento básica de una red neuronal, en ella se pueden identificar algunos elementos básicos, los cuales se mencionan a continuación:

1. *Un conjunto de ligas sinápticas.* Cada una de ellas tiene un peso asociado. Una señal de entrada x_j conectada en la sinapsis j de la neurona k es multiplicada por el peso sináptico w_{kj} .
2. *Sumador ó combinador lineal.* Para sumar o agregar las señales de entrada, que previamente fueron multiplicadas por los respectivos pesos sinápticos w_{kj} .
3. *Una función de activación $g(\)$.* Limita la amplitud de la señal de salida de la neurona.

III.1.2. Funciones de activación.

Como se menciona en la sección anterior, uno de los elementos básicos de una neurona es la función de activación, la cual se encarga de normalizar la salida de la unidad de procesamiento o neurona (Haykin, 1999). A continuación se presentan las funciones de activación más comunes:

III.1.2.1. Función Escalón.

La función escalón ó *Heaviside step* se describe mediante la siguiente ecuación:

$$g(a) = \begin{cases} 0 & \text{cuando } a < 0 \\ 1 & \text{cuando } a \geq 0 \end{cases} \quad (1)$$

Donde a es el nivel de actividad de la neurona y_k y esta definida por (2). En la figura 3 se puede apreciar gráficamente la representación de (1). Estas funciones también son conocidas como *umbral*.

$$y_k = \sum_{i=1}^p w_{ki} x_i + w_{k0} \quad (2)$$

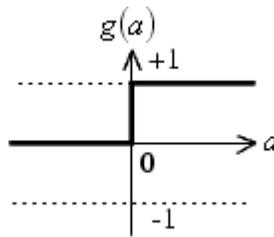


Figura 3. Función de activación escalón

III.1.2.2. Función Identidad.

La función identidad se define como:

$$g(a) = a \quad (3)$$

Donde a esta definida, al igual que en la función anterior, por (2). Esta función se utiliza cuando se desea conservar el nivel de actividad de la neurona. En la figura 4 se ilustra gráficamente.

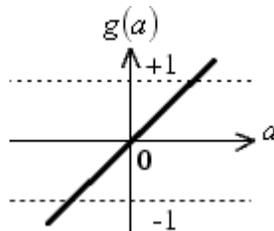


Figura 4. Función de activación identidad

III.1.2.3. Función Sigmoide.

Esta función de activación, cuya gráfica tiene forma de S , es, por mucho, la más utilizada en la construcción de redes neuronales (Haykin, 1999). Un ejemplo de función sigmoide es la función *logística*:

$$g(a) = \frac{1}{1 + e^{-a}} \quad (4)$$

La cual se puede apreciar gráficamente en la figura 5. Esta función produce salidas en el rango (0, 1).

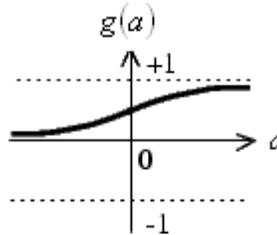


Figura 5. Función de activación logística sigmoideal

Otro ejemplo de función sigmoideal es la *tangente hiperbólica*, cuyas salidas son del rango (-1, 1).

$$g(a) = \tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (5)$$

De la cual la representación gráfica se muestra en la figura 6.

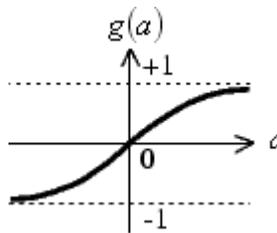


Figura 6. Función de activación tangente hiperbólica.

III.2. Perceptrón.

El perceptrón es la forma más simple de una red neuronal, la que sólo consta de una unidad de procesamiento con una función de activación de escalón (Bishop, 1995). Están limitados a la clasificación de patrones que son *linealmente separables* (Haykin, 1999). Rosenblatt fue el primero que introdujo el término *perceptrón* en el año de 1962. Se usaron principalmente para la clasificación de imágenes binarias de caracteres ó figuras simples.

Al mismo tiempo que Rosenblatt estaba trabajando en el perceptrón, Widrow y sus colaboradores trabajando en la misma línea con sistemas conocidos como *adalines*. El término proviene de ADaptive LINnear Element y se refiere a unidades de procesamiento con *umbral* no lineal de la misma forma que el perceptrón.

La salida de un perceptrón esta denotada por:

$$y = g\left(\sum_{j=0}^M w_j \phi_j(x)\right) = g(W^T \phi) \quad (6)$$

Donde ϕ denota el vector formado por las activaciones ϕ_0, \dots, ϕ_M . La función de activación de la unidad de salida es una versión anti-simétrica de la función de activación escalón definida de la forma:

$$g(a) = \begin{cases} -1 & \text{when } a < 0 \\ +1 & \text{when } a \geq 0 \end{cases} \quad (7)$$

Por lo que la salida de la unidad de procesamiento del perceptrón se puede clasificar en dos, dependiendo de su valor de salida: +1 para una clase y -1 para la otra clase. Se puede encontrar mas información al respecto en (Bishop, 1995; Haykin, 1999) así como en (Freeman y Skapura, 1991).

III.3. Arquitecturas de Redes Neuronales.

III.3.1. Redes de Capa Sencilla.

Es la arquitectura de red más sencilla, ya que contiene sólo una capa de entrada y una de salida. En la figura 7 se ilustra un ejemplo de este tipo de arquitectura.

Considerando unidades de salida lineales, se pueden expresar las salidas de la red neuronal en términos de los elementos del vector de pesos sinápticos $\{w_k\}$ por medio de la siguiente expresión:

$$y_k(x) = \sum_{i=0}^d w_{ki} x_i + w_{k0} \quad (8)$$

Donde cada línea que conecta una entrada i con una salida k corresponde a un peso sináptico w_{ki} . Una vez que la red es entrenada, se aplica un nuevo vector de entrada a la red, se calculan las activaciones de las unidades de salida y el vector de entrada se asigna a la clase cuya unidad de activación tenga el más alto valor de activación (Bishop, 1995).

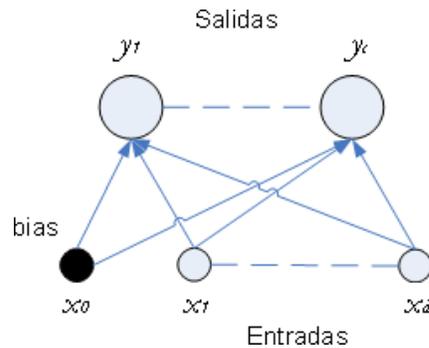


Figura 7. Red neuronal de capa sencilla

III.3.2. Redes Multi-Capa.

En esta arquitectura se cuenta, además de las capas de entrada y salida, con capas adicionales, las cuales se denominan *capas ocultas*. En la figura 8 se ilustra una arquitectura de este tipo donde se pueden apreciar las unidades ocultas, aunque en la imagen sólo se muestra una capa oculta.

Redes de este tipo son capaces de aproximar cualquier función continua cuando sus funciones de activación son sigmoideas (Duda et al., 2001), las cuales también son comúnmente conocidas como *Perceptrones Multi-Capa* (MLP, por sus siglas en ingles)..

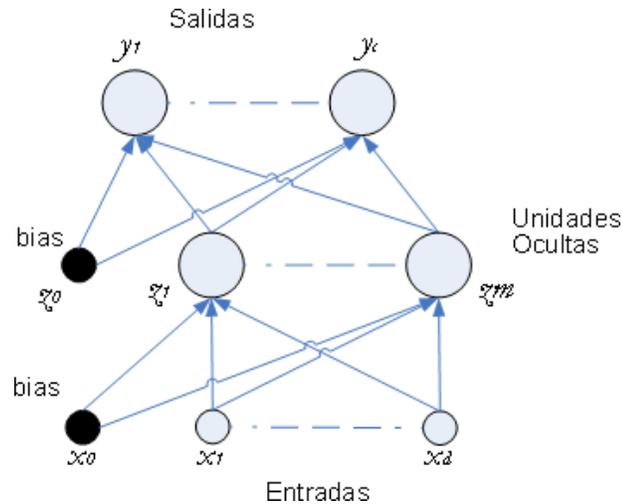


Figura 8. Red neuronal multi-capas

En la figura 8 se puede apreciar que las unidades que no son ni de entrada ni de salida son llamadas *unidades ocultas*. En este ejemplo, la red tiene d entradas, m unidades ocultas y c unidades de salida. La salida de la j -ésima unidad oculta se obtiene formando una combinación lineal de los d valores de entrada y, además, sumando el *bias*, cuya función es la de incrementar o disminuir la entrada de la función de activación, dependiendo si es positivo o negativo (Haykin, 1999). A continuación se describe el valor de la j -ésima unidad oculta:

$$a_j = \sum_{i=1}^d w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad (9)$$

Aquí $w_{ji}^{(1)}$ denota un peso en la primera capa, yendo desde la entrada i hasta la unidad oculta j , y $w_{j0}^{(1)}$ denota el bias para la unidad oculta j . La activación de la unidad oculta j se obtiene de transformar la sumatoria usando una función de activación $g(\cdot)$ en (9) para dar como resultado:

$$z_j = g(a_j) \quad (10)$$

Las salidas de la red se obtienen de transformar las activaciones de las unidades ocultas usando una segunda capa de unidades de procesamiento. Por lo que, para cada

unidad de salida k se hace una combinación lineal de las salidas de las unidades ocultas de la forma:

$$a_k = \sum_{j=1}^M w_{kj}^{(2)} z_j + w_{k0}^{(2)} \quad (11)$$

La activación de la k -ésima unidad de salida se obtiene haciendo una transformación no lineal de esta combinación lineal, para dar como resultado:

$$y_k = \tilde{g}(a_k) \quad (12)$$

Aquí se usó la notación $\tilde{g}(\cdot)$ para hacer énfasis de que no es necesario que sea la misma función de activación a la usada en las unidades ocultas. Si se combinan (9), (10), (11) y (12), se obtiene la expresión para la función completa representada por el diagrama de la red presentada en la figura 8 (Bishop, 1995).

$$y_k = \tilde{g} \left(\sum_{j=1}^M w_{kj}^{(2)} g \left(\sum_{i=0}^d w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right) \quad (13)$$

III.3.3. Redes Recurrentes.

Las redes recurrentes, a diferencia de las redes anteriores, tienen una capa de retroalimentación en la que las redes mantienen el estado de instantes de tiempo anteriores, por lo que les permite detectar y generar patrones que cambian en el tiempo (Elman, 1990). En la sección III.6. se elabora un poco más sobre este tipo de arquitectura de redes neuronales. En la figura 9 se puede apreciar un ejemplo de una red neuronal simplificada con conexiones recurrentes.

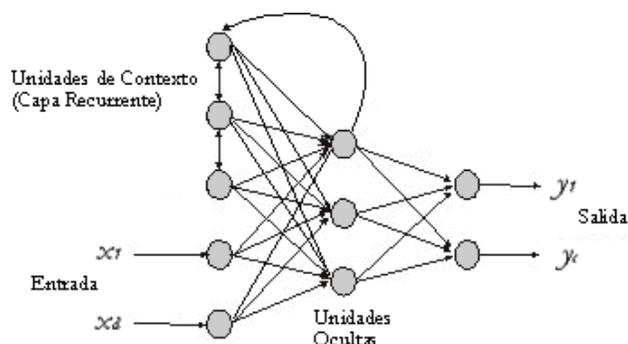


Figura 9. Red neuronal recurrente simplificada

III.4. Proceso de Aprendizaje.

En sentido amplio, todo método que incluya datos de entrenamiento, necesariamente debe incluir un proceso de aprendizaje (Duda et al., 2001). El proceso de aprendizaje es iterativo, en el cual los pesos sinápticos se van ajustando, al igual que los *bias*. El aprendizaje se refiere a algún tipo de algoritmo para reducir el error en un conjunto de datos de entrenamiento. Existen tres tipos de aprendizaje:

III.4.1. Aprendizaje Supervisado.

Para cada entrada del conjunto de datos de entrenamiento se provee de una salida. Es decir, que cada patrón de entrada a la red tiene un maestro ó salida deseada. Durante el proceso de aprendizaje, la salida generada por la red es comparada con el *profesor* y el resultado de esta comparación (error) repercute en el ajuste de los pesos sinápticos de la red. Este trabajo de tesis emplea este enfoque en el aprendizaje de los datos.

III.4.2. Aprendizaje No Supervisado.

Este tipo de aprendizaje también es conocido como *clustering*. Aquí no existe un *profesor* de manera explícita, es decir que no hay ejemplos de donde pueda aprender la red neuronal. En este tipo de aprendizaje, la red neuronal descubre las *regularidades* o características de los datos de entrada, formando *clusters* o “agrupaciones naturales” de los patrones de entrada (Duda et al., 2001).

III.4.3. Aprendizaje de Refuerzo.

Este tipo de aprendizaje es parecido al aprendizaje supervisado, la diferencia radica que mientras en el aprendizaje supervisado el *profesor* muestra la “respuesta verdadera”, en el aprendizaje de refuerzo el *profesor* sólo responde si la respuesta es correcta ó incorrecta. Es decir, que la salida de la red neuronal es clasificada como correcta o incorrecta, pero no se comunica en qué medida es incorrecta (Duda et al., 2001).

III.5. Redes de Retropropagación.

En la sección III.1 se explicó la importancia de las redes neuronales en la solución de determinados problemas. Las redes de retropropagación (BPN) son un ejemplo de ellas, en las que estas *aprenden* las relaciones entre un conjunto de patrones de ejemplo, y son capaces de aplicar la misma relación para nuevos patrones de entrada nunca antes vistos (Freeman y Skapura, 1991).

En esta sección se verá la forma en que las redes de retropropagación resuelven problemas complejos de reconocimiento de patrones como los presentados en la sección III.1.

La red neuronal *aprende* un conjunto predefinido de pares de datos de entrada-salida usando un ciclo de dos fases: propagación y adaptación. Después de que el primer patrón de entrada ha *estimulado* la primera capa de las unidades de la red, la señal se propaga a través de todas las capas posteriores hasta que se produce una salida. Dicha salida, obtenida por la red, se compara con la salida deseada y se calcula el error producido para cada unidad de salida.

Las señales de error producidas se transmiten hacia atrás desde la capa de salida hasta cada nodo en las capas intermedias que contribuyen de alguna manera a la salida. Sin embargo, cada unidad en las capas intermedias recibe sólo una porción del error total de la capa de salida, tomando en cuenta la contribución de dicha unidad a la salida original. Basándose en la señal de error que recibieron, todas las unidades de las capas intermedias

actualizan los pesos sinápticos de sus conexiones para tratar de que la red considere todos los patrones de entrenamiento, de modo que se vaya *generalizando* su salida.

La importancia de este proceso radica en que, a medida que la red es entrenada, los nodos de las capas intermedias se auto-organizan de forma que aprenden a reconocer de manera independiente diferentes aspectos del total de los patrones de entrada. Después del entrenamiento, cuando se presente una entrada incompleta o *ruidosa* a la red neuronal, ésta responderá con una salida adecuada si la entrada contiene algún patrón que le *recuerde* las particularidades que cada unidad aprendió a reconocer durante el entrenamiento.

De esta forma, la red desarrolla una representación interna a través de relaciones entre sus nodos de las capas ocultas, que pueden ser o no visibles para el ojo humano. Sin embargo, lo importante aquí es que ayudan a generar las salidas deseadas cuando se presentan los patrones de entrenamiento. Esta misma representación puede ser aplicada a nuevos patrones de entrada nunca antes vistos por la red y, aun así, obtener buenos resultados. Esto es posible dado que comparten algunas ó las mismas particularidades de los patrones de entrenamiento (Freeman y Skapura, 1991).

III.5.1. Regla Generalizada Delta

Esta sección presenta el algoritmo de aprendizaje de las BPN: La Regla Generalizada Delta (GDR, por sus siglas en inglés). Las redes de retropropagación son redes con capas, de alimentación hacia delante en las que sus nodos en capas adyacentes están conectados hacia delante. No hay conexiones de retroalimentación ni conexiones que sobrepasen su capa posterior. Por lo general se usan sólo 3 capas, aunque se pueden usar más de 1 capa oculta.

Una red neuronal es capaz de identificar una relación funcional entre sus entradas y sus salidas. Por ejemplo, si el valor de entrada de la red neuronal es el valor de un ángulo y el valor de salida es el coseno de dicho ángulo, la red implementa la transformación de $\theta \rightarrow \cos(\theta)$. Para esta transformación no es necesaria una red neuronal. Sin embargo, se

puede usar para transformaciones más complejas donde no se conoce con anterioridad la relación funcional pero si se conocen algunos ejemplos del *mapeo* correcto.

El procedimiento básico para el entrenamiento de una red neuronal consta de los siguientes pasos:

1. Aplicar un vector de entrada, propagarlo hacia delante y calcular sus respectivas salidas.
2. Comparar las salidas obtenidas con las salidas deseadas (correctas) y determinar el error obtenido.
3. Determinar el monto por el cual cambiar los pesos sinápticos para reducir el error.
4. Hacer las correcciones a los pesos hacia atrás.
5. Repetir los pasos 1 al 4 con todos los vectores de entrenamiento hasta que el error para todos los vectores del conjunto de entrenamiento se reduzca a un valor aceptable.

La descripción matemática formal de este algoritmo se puede encontrar en (Freeman y Skapura, 1991; Bishop, 1995; Haykin, 1999).

III.6. Redes Recurrentes.

Las redes recurrentes son aquellas que cuentan con conexiones de retroalimentación entre sus capas (Skapura, 1996). Este tipo de redes se usan principalmente para el manejo de patrones temporales, así como espaciales (Hüsken y Stagge, 2003). Se usan tanto en la clasificación, como en la predicción de valores futuros (Hüsken y Stagge, 2003). Un ejemplo de predicción es el del clima, donde se toman como muestras los valores de temperatura, humedad, así como otras condiciones atmosféricas y se tratan de predecir valores futuros en base a los registros históricos que se tienen de esas variables.

En el caso de este trabajo de tesis, se exploró el uso de este concepto para estimar la posición del usuario en base a donde estuvo anteriormente, es decir que se toma en cuenta las posiciones anteriores para disminuir el error en la estimación de la localización.

En las siguientes sub-secciones se habla de dos tipos de redes recurrentes, de las cuales se ha explorado su uso en trabajos anteriores (Elman, 1990; Dorffner, 1996; Hüsken y Stagge, 2003) para abordar el problema de predicción, así como el de clasificación.

III.6.1. Redes de Elman.

Las redes Elman fueron propuestas en (Elman, 1990) y son una variación del MLP. Se incluye retroalimentación de la salida de la capa oculta hacia la primera capa. Esta conexión recurrente permite a la red neuronal aprender patrones temporales. En la figura 10 se muestra una red neuronal Elman simplificada.

En este tipo de redes se agregan unidades a la capa de entrada llamadas *unidades de contexto*. Este tipo de unidades duplican la actividad de las unidades de la capa oculta (Skapura, 1996) y gracias a su conexión recurrente es posible guardar valores del instante de tiempo anterior $t - 1$ que pueden ser usados en el instante de tiempo t .

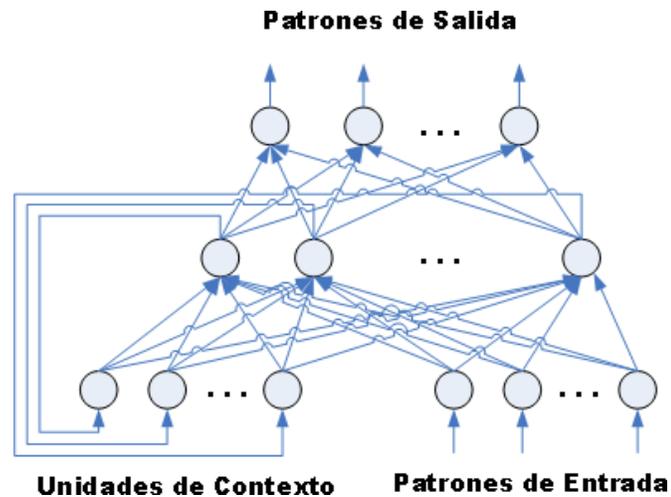


Figura 10. Red Elman simplificada

El procesamiento de las redes Elman consiste en la siguiente secuencia de eventos: en el tiempo t , las unidades de entrada reciben los primeros patrones de entrada. Las unidades de entrada, así como las unidades de contexto activan las unidades ocultas, para luego estas ser las que activen las unidades de salida. Además, dichas unidades ocultas retroalimentan a las unidades de contexto. Si no se presenta aprendizaje el error se propaga hacia atrás para ajustar los pesos sinápticos. Al siguiente instante de tiempo $t+1$, la secuencia anterior se repite. Sin embargo, esta vez las unidades de contexto tendrán valores idénticos a los encontrados en la capa oculta en el tiempo t , de modo que se hace una copia de la capa oculta. Estas unidades de contexto son las que proveen la memoria a la red (Elman, 1990).

III.6.2. Redes de Jordan.

Otro tipo de redes recurrentes son las redes de Jordan donde, al igual que las redes de Elman, también se obtiene retroalimentación con la diferencia de que ésta se provee por la capa de salida en vez de la capa oculta. En la figura 11 se ilustra de manera gráfica la manera en que se conecta este tipo de red.

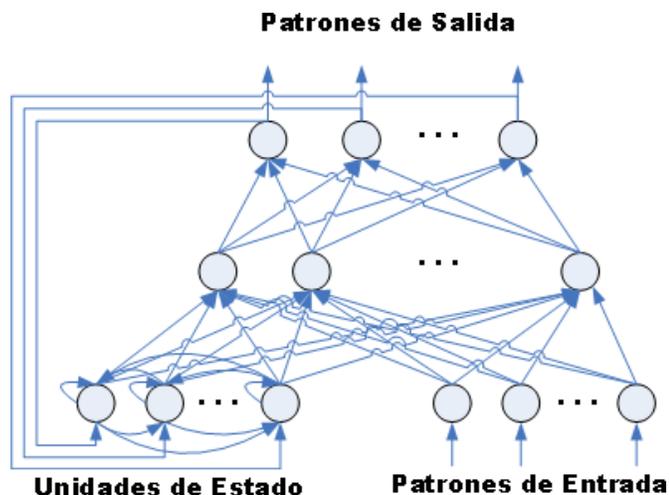


Figura 11. Red Jordan simplificada

Este tipo de redes es apropiado para aprender secuencias de patrones ya que la salida en el instante de tiempo anterior $t - 1$ es parte de la entrada a la red en el tiempo t . De esta forma, las redes Jordan son adecuadas en aplicaciones donde los patrones de entrenamiento están relacionados unos con otros en una secuencia específica de tiempo (Skapura, 1996).

III.7. Resumen.

En este capítulo se presentó un bosquejo general de las redes neuronales. Al inicio se presentó, de manera general, como pueden ser aplicadas en problemas complejos para los que no existe una representación matemática cerrada.

También se presentó una reseña sobre la arquitectura básica de una red neuronal, así como los tipos de funciones de activación más comunes. Se describen, de igual forma, las arquitecturas típicas de las redes neuronales, el proceso de aprendizaje y, finalmente, se abordan las redes de retropropagación así como las redes recurrentes y los dos tipos más comunes de éstas.

En el siguiente capítulo se describen actividades, así como *software* y herramientas diseñadas para llevar a cabo la recolección de muestras, así como la descripción de los experimentos que se hicieron con las muestras obtenidas durante dicha fase.

Capítulo IV.

Recolección de Muestras y Preparación de Experimentos

El método que se propone en esta tesis esta basado en la toma de muestras empíricas de la señal desde diferentes puntos físicos del edificio. Con ellas, se entrena una red neuronal para estimar la posición de un dispositivo móvil.

En este capítulo se presenta la descripción del lugar en donde se llevaron a cabo la toma de mediciones de la señal inalámbrica, así como la forma en que se realizó su recolección. De igual forma, se describe el hardware y software usados en la recolección de las mismas.

En la última sección de este capítulo se describen una serie de experimentos que se realizaron con las muestras obtenidas con el objetivo de mejorar resultados obtenidos en trabajos previos en la estimación de la localización.

IV.1. Plataforma Experimental.

En esta sección se describen las características del lugar en donde se tomaron las muestras de la señal inalámbrica, así como las características del equipo usado en la recolección de las mismas.

IV.1.1. Descripción del lugar

La toma de muestras se llevó a cabo en el tercer piso del Hospital General del Instituto Mexicano de Seguro Social (IMSS) Zona IV No. 8 de Ensenada en el área de Medicina Interna. Las medidas del piso en cuestión son 46.8 m. de largo por 13.3 m. de ancho. El piso incluye 10 cuartos para pacientes, de los cuales 3 son para pacientes aislados y los restantes tienen 3 camas por cuarto. Además, hay un cuarto de curación, dos baños y una central de enfermeras. En la figura 12 se puede apreciar el plano arquitectónico del edificio

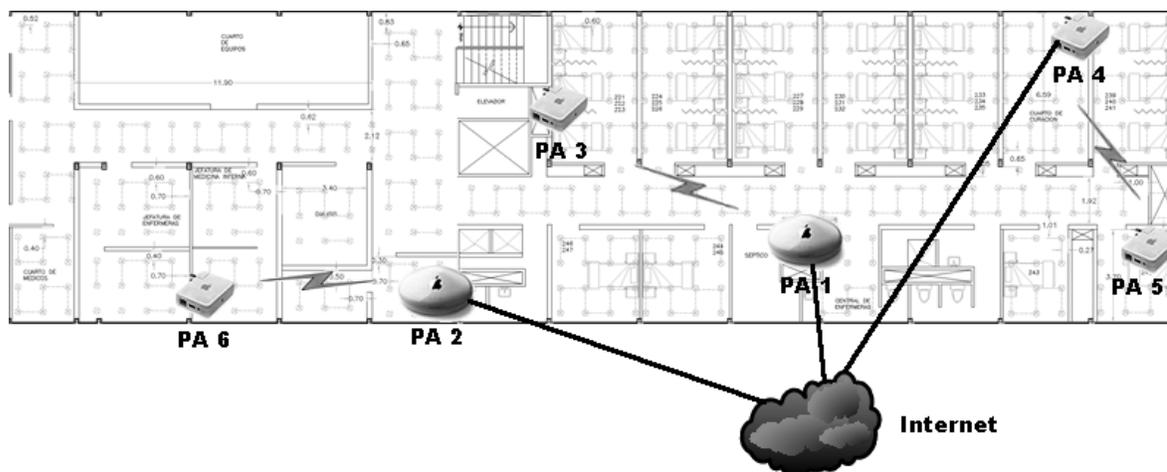


Figura 13. Conectividad de la red inalámbrica en el hospital.

Los puntos de acceso a la red inalámbrica utilizados son del fabricante Apple, de los cuales dos son modelo Airport Extreme (PA 1 y PA 2) y los cuatro restantes son Airport Express. En la figura 14 se muestran imágenes que ilustran a los puntos de acceso mencionados.

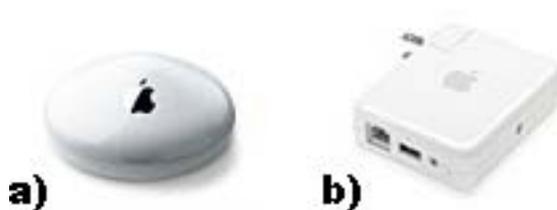


Figura 14. Puntos de acceso usados en las mediciones. a) Airport Extreme; b) Airport Express

La red inalámbrica opera en la banda de 2.4 GHz, la cual es de licencia libre ISM (Industrial, Científica y Médica) y tiene un rango de alcance de 160 m, 50 m y 25 m para espacios abiertos, semiabiertos y cerrados, respectivamente. La banda de los 2.4 GHz esta dividida en 11 canales (estándar de USA), de los cuales se utilizan 3 canales: 1, 6 y 11.

Los puntos de acceso con frecuencias adyacentes fueron colocados lo más lejos posible para minimizar la interferencia entre ellos. La utilización de los canales se muestra en la tabla I.

Tabla I. Distribución de frecuencias en los puntos de acceso

No.	Nombre	Canal	Frecuencias	BSSID (MAC)
1	AP1	11	1 2 5.5 11	00:11:24:1F:92:71
2	AP2	1	1 2 5.5 11	00:11:24:21:15:A2
3	AP3	11	1 2 5.5 11	00:11:24:0B:EC:87
4	AP4	6	1 2 5.5 11	00:11:24:0B:53:A7
5	AP5	6	1 2 5.5 11	00:11:24:0B:AF:C1
6	AP6	1	1 2 5.5 11	00:11:24:EB:DB:49

Para la toma de mediciones se usó un PDA (Asistente Digital Personal) modelo Dell Axim X3 con procesador Intel XScale de 400MHz, sistema operativo Pocket PC, provisto con 64Mb de memoria RAM. El dispositivo trae integrada la tarjeta de red inalámbrica (WNIC).

IV.2. Recolección de muestras.

La red neuronal fue entrenada con las lecturas recolectadas de los 6 puntos de acceso localizados en los lugares previamente mostrados en la figura 12. Se tomaron mediciones en 323 puntos en todo el piso de medicina interna. En cada punto se tomaron 100 muestras de la señal recibida por el dispositivo móvil con una periodicidad de una por segundo en la que cada muestra es igual a 6 valores de los distintos puntos de acceso. En todos los puntos se registró el valor de la intensidad de la señal de al menos 4 puntos de acceso. En los casos donde la señal no era lo suficientemente fuerte para ser medida el valor que se asignó a dicho punto de acceso fue de -99dBm. Normalmente la señal detectada por el PDA varía en un rango de -25dBm a -94dBm. Sin embargo, cada fabricante tiene especificaciones diferentes.

Al final, se obtuvieron 32,300 muestras en 4 sesiones, las cuales tomaron alrededor de 9 horas contando sólo el tiempo de estar registrando las lecturas de la tarjeta inalámbrica.

IV.2.1. Aplicación para recolección de muestras.

Para facilitar la recolección de muestras en el hospital, se diseñó una aplicación para Pocket PC. En la cual se tiene almacenado un plano arquitectónico del hospital a escala, de modo que se tienen las medidas exactas del hospital. La aplicación cuenta con una interfaz sencilla en la que con solo tocar la pantalla con la pluma que trae incluida el PDA se inicia el registro de valores de intensidad de la señal en la localización física señalada en el PDA. En la figura 15 se puede apreciar una fotografía de la aplicación siendo utilizada en la recolección de muestras.



Figura 15. Aplicación para Pocket PC con la que se hizo el registro de las intensidades de la señal de los diferentes puntos de acceso.

Cada vez que el usuario toca un lugar del mapa en el PDA, aparece un mensaje de confirmación para iniciar el registro de las intensidades de la señal de los puntos de acceso en el punto señalado. Se tomaron 100 muestras en cada lugar señalado del mapa. Sin embargo, dichos puntos son solo una referencia ya que en la práctica, hubo varios obstáculos tales como camas, mesas, cajas con medicamentos, etc. Si no era posible tomar

la lectura en el punto señalado, simplemente se colocaba en la posición físicamente más cercana al punto y se marcaba la posición correspondiente en la aplicación. En el apéndice A se encuentra el manual de usuario de la aplicación diseñada para la recolección de muestras.

Además, para agilizar la toma de muestras en el hospital, se diseñó una red de 3 m cuadrados, la cual se iba posicionando de acuerdo a los puntos marcados en el mapa del PDA, en cada vértice de la red se iba colocando una marca para, después de colocadas varias marcas, empezar las lecturas. En la figura 16, se ilustra como es que dicha red se utilizó para el posicionamiento de las marcas antes de ser utilizadas como referencia en la toma de muestras.



Figura 16. Marcas que se usaron para la toma de mediciones.

En la figura 17 se aprecia de una manera gráfica la forma en que la señal disminuye dependiendo de la ubicación del punto de acceso.

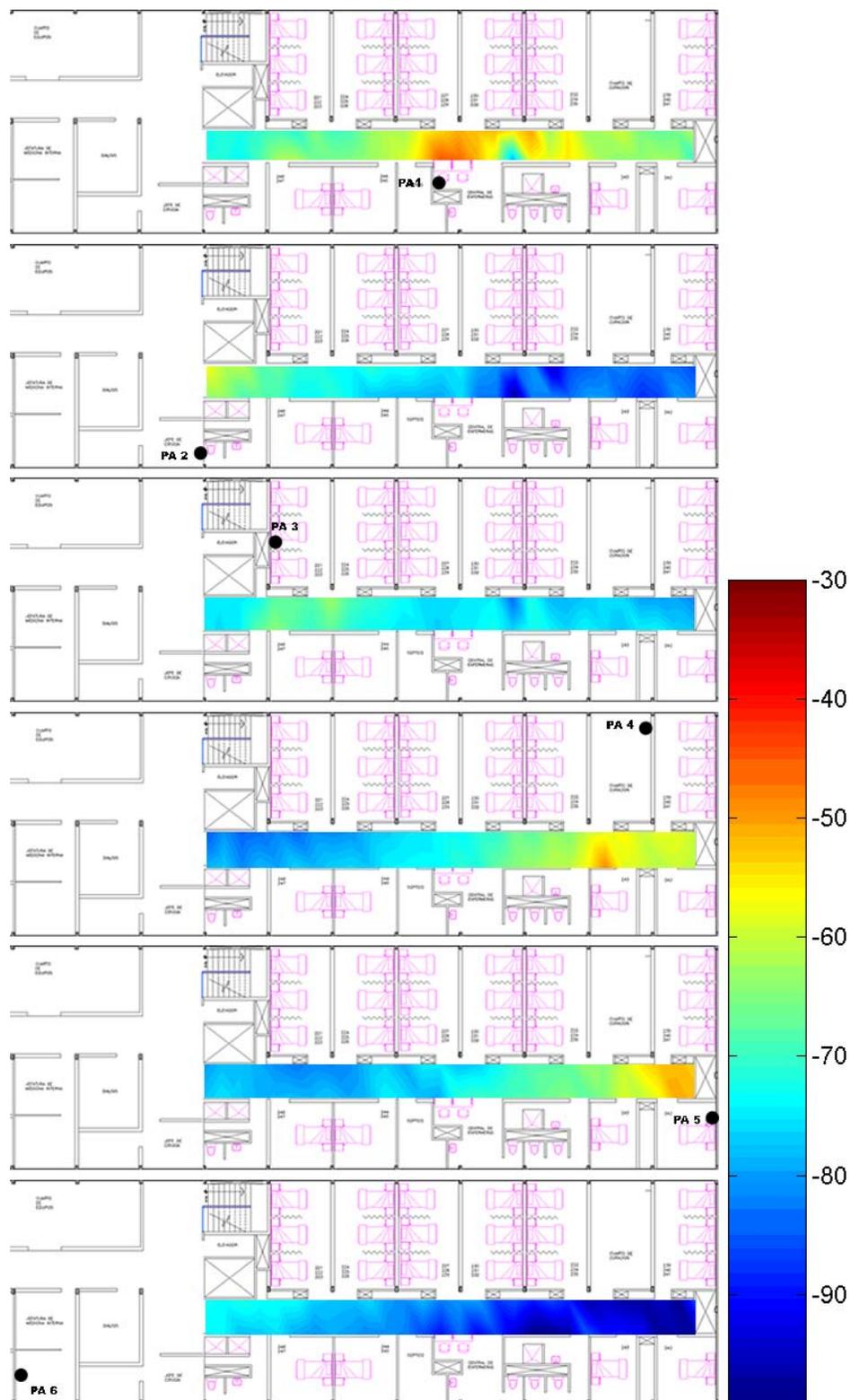


Figura 17. Intensidades de la señal en el pasillo del área de medicina interna.

IV.2.2. Agrupamiento de muestras para entrenamiento.

En todos los puntos en donde se recolectaron muestras se obtuvieron registros de al menos cuatro puntos de acceso. Con el objetivo de ver si se obtenían mejores resultados en el entrenamiento se decide agrupar las muestras en diversos conjuntos como se explica a continuación.

Del conjunto original de muestras que se obtuvieron de los 323 puntos se decide eliminar una parte de ellos (26 puntos) por considerarse que eran puntos que afectarían el entrenamiento de la red. El criterio de eliminación se explica más adelante en la sección IV.3.2. De los 297 puntos restantes se tomó el 20% de ellos para prueba (59 puntos), los cuales fueron elegidos de manera aleatoria.

El 80% restante de las muestras (238 puntos) se tomó para entrenamiento, de los cuales se formaron otros subgrupos de datos de entrenamiento, con el 60% de los datos, el 40% y, finalmente el 20% de los datos. En la figura 18a se muestra la forma en que se agruparon los datos.

Además, de las 100 muestras de cada punto de acceso se formaron cuatro grupos de 10 elementos cada uno, como se muestra en la figura 18b. Por ejemplo, para el caso de los grupos promediados de 3 muestras, se toman las muestras numeradas del 1 al 3, después se toman las numeradas del 11 al 13, y así sucesivamente hasta juntar los 10 grupos y se obtiene un promedio aritmético de cada grupo. Lo mismo para los demás grupos. En el único caso donde no se promedia, por obvias razones, es en el caso donde se toma una sola muestra: es decir, se toma la número 1, la número 11, la 21 y así sucesivamente.

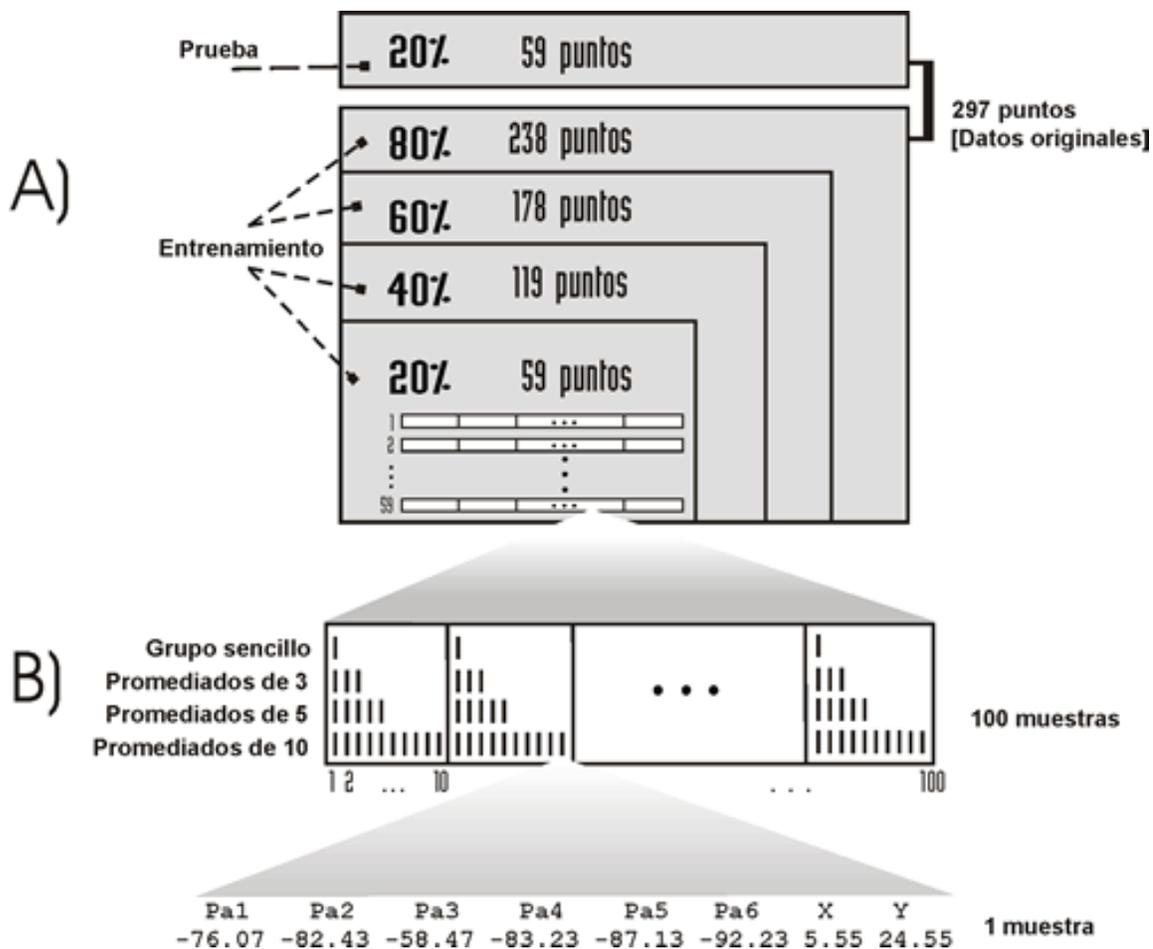


Figura 18. a) División de los puntos para entrenamiento y prueba; b) cada muestra es agrupada de 4 formas diferentes para entrenamiento y prueba.

IV.3. Experimentos realizados con los datos obtenidos.

En esta sección se describen los diversos experimentos efectuados durante la realización de este trabajo. Cada experimento arroja resultados sobre diferentes aspectos tales como número de puntos de acceso a considerar, densidad de mediciones, etc. que en algunos casos pudiera ser relevantes sobre todo por cuestiones de costo-beneficio. Los resultados de tales experimentos se presentan en el capítulo V.

IV.3.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal

En este experimento se busca conocer si la altura del dispositivo influye de manera significativa en el valor obtenido de la intensidad de la señal, es decir para personas mas altas que otras cargando un PDA, la altura puede influir de manera directa en los valores de la señal, ya sea por rebotes de la señal en el piso o la misma complejión del usuario que carga el dispositivo.

Este experimento se llevó a cabo en las instalaciones de CICESE, en el 3er piso del edificio de Física Aplicada a lo largo de un pasillo. Se tomaron en cuenta 3 puntos de acceso, de los cuales se pueden apreciar las posiciones en la figura 20.

Para llevar a cabo este experimento se diseñó un pedestal (figura 19) para capturar las intensidades de la señal desde diferentes alturas sin que varíe el punto de referencia. El pedestal tiene 4 alturas fijas en las cuales se toman las mediciones: a 0.5 m, a 1 m, a 1.5 m y, por último, a 2 m del piso.

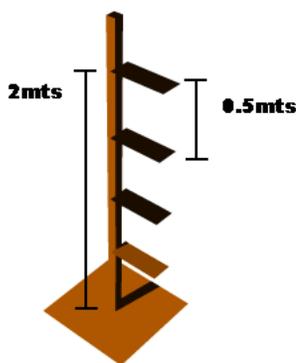


Figura 19. Pedestal usado en la toma de lecturas para este experimento.

Los modelos de los puntos de acceso utilizados en este experimento son Airport Extreme (802.11b/g) de Apple para BS1 y BS2, y para el caso del punto de acceso denominado “colaborativos” se usó un punto de acceso WAP11 (802.11b) del fabricante LinkSys.

con una distancia de separación de 1 metro uno de otro. Sin embargo, tomar mediciones a 1 metro requiere de mucho esfuerzo, si se toma en cuenta que por cada punto se requiere de alrededor de 2 minutos. Por lo tanto, tomando en cuenta que son 323 puntos, se requieren alrededor de 11 horas sólo para estar parado en un lugar registrando las intensidades de la señal.

Todo esto, sin tomar en cuenta el tiempo invertido en hacer las mediciones físicas en el edificio, así como el tiempo de espera cuando algún médico esta haciendo alguna inspección física a los pacientes, o cualquier otro procedimiento en los que, indudablemente, nuestra presencia pudiera incomodar a los pacientes, familiares o personal médico.

Tabla II. Tabla que muestra la distribución de puntos en los diferentes conjuntos

Conjuntos de datos	No. de puntos	PCT (%)
Original	323	100
Puntos irregulares	26	8
Conjunto Original menos puntos irregulares	297	92
Conjunto de Prueba (20%)	59	18
Entrenamiento (80%)	238	74
Entrenamiento (60%)	178	55
Entrenamiento (40%)	119	37
Entrenamiento (20%)	59	18

Por las cuestiones mencionadas en los párrafos anteriores se decide hacer un experimento eliminando parte del conjunto de entrenamiento. Los datos en cada conjunto quedan como se expresa en la tabla II.

Los porcentajes de los datos de prueba y los datos de entrenamiento son respecto al número total de puntos menos los puntos que fueron descartados, es decir 297. El criterio que se toma en cuenta para la eliminación de los puntos mencionados se explica a continuación.

Se tienen 100 muestras por cada uno de los 323 puntos en el mapa y se tienen lecturas de 6 puntos de acceso, lo cual da un total de 1938 distribuciones de datos por cada sitio y punto de acceso. Se decide calcular la media y la desviación estándar de dichos puntos y como criterio de eliminación se descartan aquellos en los que uno o más de los datos registrados de los 6 diferentes puntos de acceso den como resultado una desviación estándar mayor a 10. La fórmula de la desviación estándar que se usa:

$$S_{N-1} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (14)$$

Donde \bar{x} es la media de las N intensidades de cada uno de los 6 puntos de acceso registradas en un punto determinado. N es igual a 100, es decir el número de muestras tomadas en un punto determinado.

Después de efectuar el cálculo correspondiente se eliminaron los mencionados 26 puntos, curiosamente la mayoría de los puntos descartados para entrenamiento pertenecen al punto de acceso 2 (PA 2) y a dos cuartos de pacientes. En la figura 21 se muestra gráficamente la ubicación de los sitios descartados y a qué punto de acceso se atribuye su eliminación del conjunto de entrenamiento.

Como se puede observar en la figura 21, la mayoría de los puntos pertenece a dos cuartos de pacientes con un punto de acceso en específico. Sin embargo, no es trivial suponer las causas que originaron tales problemas aunque suponemos que la causa puede ser la cantidad de paredes que se interponen entre el mencionado punto de acceso y el dispositivo móvil al momento de estar tomando las lecturas. Sin embargo, existen muchos otros sitios donde el número de paredes es igual o mayor a los de los sitios en cuestión, de tal forma, que resulta un poco difícil de creer que éste sea el problema. Una de las causas que posiblemente afectó, fue que al momento de estar haciendo las mediciones con el dispositivo móvil, las puertas de dichos cuartos permanecieron cerradas, ya que los pacientes en esos cuartos eran pacientes aislados. No fue posible corroborar si en la oficina donde está el punto de acceso la puerta también estaba cerrada al momento de hacer las mediciones.

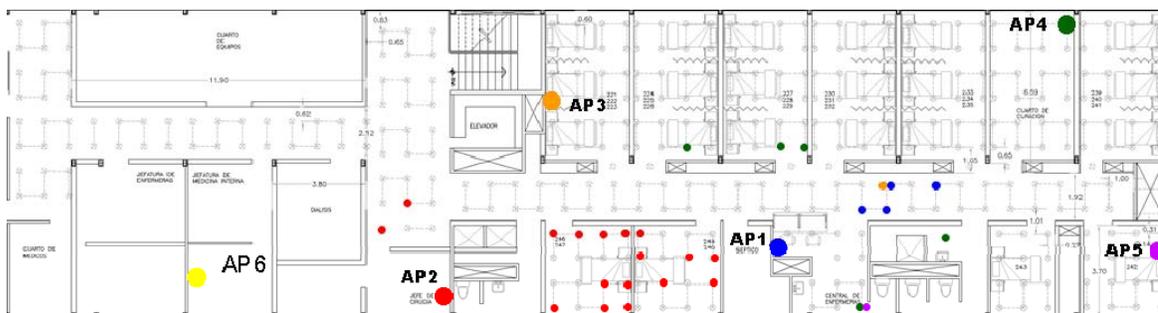


Figura 21. Sitios donde se detectaron lecturas irregulares.

IV.3.3. Influencia del número de muestras en la estimación de la localización.

El objetivo de este experimento es el de identificar cuantas muestras deben ser consideradas al momento de estimar la localización de modo que se obtenga una estimación confiable. Dicho de otra manera, es tomar como datos de entrada a la red neuronal 1 muestra (1 muestra = 6 lecturas de puntos de acceso) o, en otro caso, tomar como datos de entrada 3 muestras promediadas y ver cual de las dos formas proporciona un mejor resultado en la estimación de la localización.

Se sabe que la señal de radiofrecuencia varía mucho, con este experimento lo que se busca es suavizar las lecturas, es decir que si una de cada 10 lecturas, por ejemplo, es un dato fuera de los niveles normales (outliers) se pueda disminuir el efecto que tiene éste en la estimación de la localización.

Para este experimento se considera el agrupado y promediado de conjuntos de 10, 5, 3 y 1 muestra. El criterio de selección de los grupos es similar al de la sección IV.2.2. en donde se van seleccionando de grupos de 10, la primera muestra, las primeras 3 muestras, las primeras 5 o las 10 muestras y promediarlas, según sea el caso.

IV.3.4. Influencia del número de puntos de acceso en la estimación de la localización.

El objetivo de este experimento es el de ver que tanto afecta el tener determinado número de puntos de acceso. Las mediciones se realizaron con 6 puntos de acceso, en

posiciones fijas. Sin embargo, es relativamente fácil, el eliminar los datos de un determinado punto de acceso, entrenar la red neuronal y ver como es que influye en el error obtenido en la estimación de la localización.

Se realizaron pruebas con 3, 4, 5 y 6 puntos de acceso. En el caso donde se dejaron solo 3 puntos de acceso, se probaron con tres diferentes combinaciones de puntos de acceso, es decir, que algunas veces se dejaron los puntos de acceso No. 3, 4 y 5, en otro los puntos de acceso No. 1, 4 y 5 y, finalmente un experimento con los puntos de acceso 1, 3 y 5 (ver figura 12)

Obviamente, el objetivo de este experimento es conocer la sensibilidad del error con respecto a la cantidad de puntos de acceso usados para la estimación para ver si es posible reducir el número de ellos y aun así obtener una aproximación de la posición aceptable.

IV.3.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.

En este experimento lo que se desea es ver las variaciones en el error obtenido por las redes neuronales después de ser entrenadas con diferentes arquitecturas. Básicamente, lo que se hace en este sentido es cambiar el número de unidades en la capa oculta. Mientras más unidades ocultas tiene una red neuronal, más será su tiempo de entrenamiento, así como el de prueba, aunque el tiempo en realidad no importa en estos casos ya que se hacen *off-line*, es decir que el entrenamiento puede durar horas e incluso días sin afectar de gran manera el desempeño al simular la red en tiempo real.

La diferencia radica en el dispositivo que se desea simular la red neuronal, en una computadora de escritorio al hacer la simulación toma un tiempo de centésimas de segundo, en cambio en un PDA toma alrededor de 1 segundo.

IV.3.6. Entrenando la red neuronal con información de vecindarios.

En este experimento lo que se intenta es tratar de reducir el error máximo obtenido al momento de hacer la estimación. Es decir, que tomando en cuenta la posición que se

desea estimar, se agrega como dato de entrada, además de los valores de la intensidad de la señal de los 6 puntos de acceso, la coordenada de un vecino a dicha posición. Es decir, que se entrena la red neuronal con todos los lugares desde donde es accesible la posición a estimar. En la figura 22 se ilustra un ejemplo de una posición así como sus vecinos considerados como entrada.

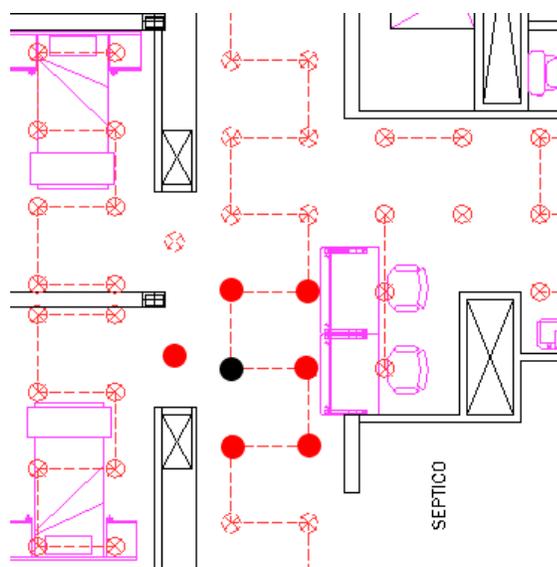


Figura 22. Vecinos (rojo) considerados para entrenamiento de una posición específica (negro).

Es preciso señalar que cada una de estas posiciones representa una entrada diferente. Es decir que las tuplas de entrenamiento y prueba en este experimento son de la forma: $(X_v, Y_v, PA1, PA2, PA3, PA4, PA5, PA6, X, Y)$. Donde X_v y Y_v son las coordenadas de uno de los vecinos. $PA1, PA2, PA3, PA4, PA5$ y $PA6$ son los valores de la intensidad de la señal y, finalmente, las coordenadas x y y son las de la posición a estimar. La estructura física del edificio influye directamente en la cantidad de vecinos que se eligen para cada posición. Por ejemplo, si la posición que se desea estimar es una esquina en determinado cuarto, la cantidad de vecinos podría ser 3, en comparación con los vecinos elegidos en la figura 22, donde el número de vecinos es 6.

IV.4. Resumen.

En el presente capítulo se describió el lugar en donde se tomaron muestras de la intensidad de la señal inalámbrica, las cuales serán la entrada para la red neuronal con la que se hace la estimación de la localización de los dispositivos móviles.

Se describe también el hardware y el software usado para la toma de lecturas. Así como la forma en que se agruparon los conjuntos de muestras para el entrenamiento y prueba de la red neuronal.

De igual forma, se describen una serie de experimentos que se realizaron con las muestras obtenidas, cada uno con objetivos muy específicos orientados a obtener y comprobar diferentes aspectos señalados en la sección correspondiente a cada experimento.

En el siguiente capítulo se describen los resultados de los experimentos mencionados en el presente. Además, se incluye una discusión acerca de los resultados obtenidos.

Capítulo V.

Resultados y Discusión.

En este capítulo se presenta la descripción de los resultados obtenidos así como una discusión de los mismos desde tres puntos de vista: los resultados aislados de los experimentos, de igual forma se presentan los resultados obtenidos en la literatura en experimentos similares y, finalmente, la utilidad de los resultados obtenidos en términos de una aplicación consciente del contexto en un ambiente hospitalario.

V.1. Resultados de los Experimentos.

En esta sección se describen los resultados de los experimentos descritos en la sección IV.3. Para comparar entre las distintas configuraciones, el error se calcula en base a la distancia euclidiana entre la posición real y la posición estimada por la red neuronal. Para comparar los resultados se calculó la Distancia de Error Promedio:

$$DEP = \frac{1}{P} \sum_{p=1}^P \sqrt{(tx_p - ox_p(w))^2 + (ty_p - oy_p(w))^2} \quad (15)$$

Donde $ox_p(w)$ y $oy_p(w)$ son las coordenadas x y y obtenidas por la red neuronal y tx_p y ty_p son los patrones de salida correctos. P es la cantidad total de patrones de entrada.

V.1.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal.

En este experimento, como se menciona en la sección IV.3.1., se desea conocer si la altura es un factor determinante en la variación de la intensidad de la señal. Los resultados que se obtienen de este experimento se describen en los siguientes párrafos.

Para iniciar, se formulan dos hipótesis:

H₀: La altura no es un factor determinante en la intensidad de la señal

H₁: La altura es un factor determinante en la intensidad de la señal

En la figura 23 se muestran las intensidades recolectadas a las diferentes alturas en el punto de acceso denominado BS1. Como puede observarse a simple vista, la intensidad de la señal disminuye a medida que uno se aleja del punto de acceso, el cual se encuentra ubicado unos 3 metros a la izquierda del inicio de la gráfica.

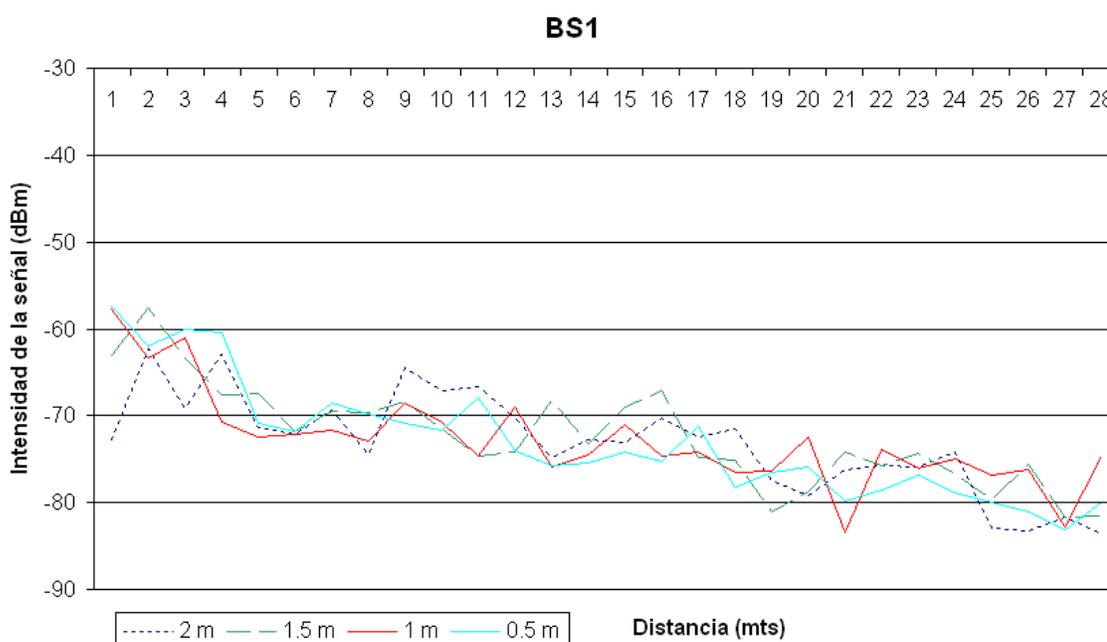


Figura 23. Punto de acceso BS1 y la medición de la señal en altura y distancia

De igual forma, se puede ver que el comportamiento de la gráfica, en general, es el mismo para las cuatro alturas en las que se tomaron lecturas. Aunque a simple vista pueden parecer comportamientos similares es necesario hacer un análisis estadístico de los datos para tener mayor certeza.

Se hizo un análisis de varianza (ANOVA) de las lecturas obtenidas a diferentes alturas y se llegó a la conclusión de que estadísticamente no existe una diferencia significativa entre los cuatro grupos. En la tabla III se aprecia el resultado del análisis.

Tabla III. Análisis de Varianza del punto de acceso BS1

Fuente de Variación	Suma de Cuadrados	Grados de libertad	Media de Cuadrados	F	P	F crítica
Entre Grupos	11.30062	3	3.766872	0.105386	0.956771	2.688691
Dentro de los Grupos	3860.3	108	35.74352			

En la tabla III se puede apreciar el valor F y el valor de F-crítica, cuando el primer valor es superado por el segundo quiere decir que la hipótesis nula (H_0) se acepta. Además, el valor P nos indica el nivel de certeza, en este caso es un valor por arriba del 95%. Por lo que se puede afirmar que la altura no es un factor determinante en la intensidad de la señal.

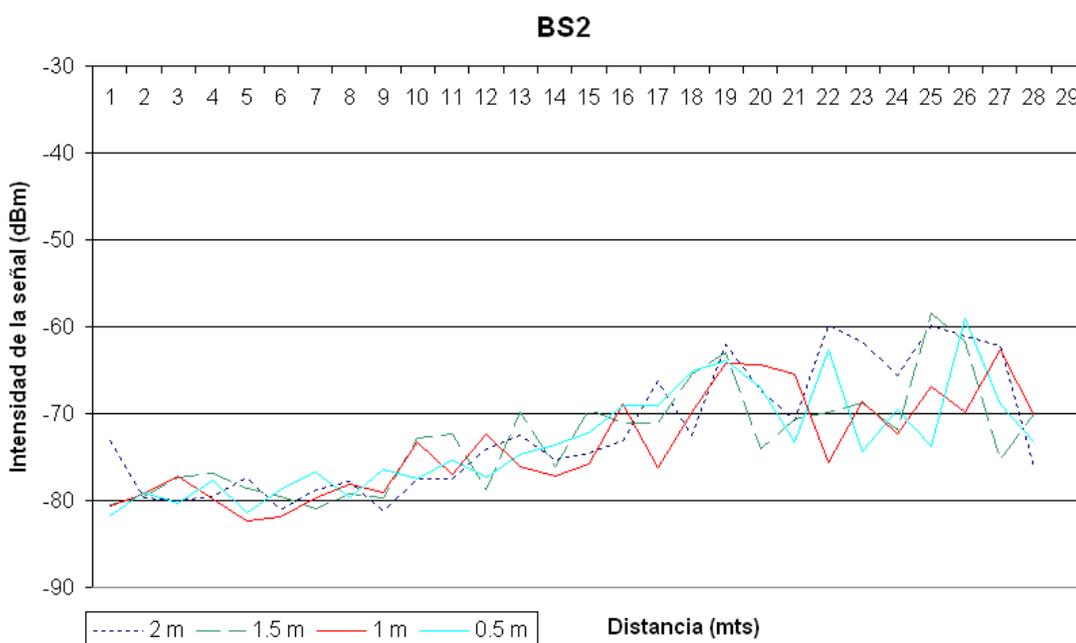


Figura 24. Punto de acceso BS2 y la medición de la señal en altura y distancia

Por otro lado, para el punto de acceso BS2, se puede apreciar en la figura 24 la forma en que la fuerza de la señal va aumentando a medida que el instrumento de medición se aproxima al punto de acceso, el cual está ubicado en el metro 22 aproximadamente. En este caso, al igual que el anterior se hizo un ANOVA, para ver si la diferencia de los datos es estadísticamente significativa, el cual se muestra en la tabla IV. De nuevo, del valor de F y F-crítica es fácil darse cuenta que F-crítica es mayor que la F, por lo que en este caso la

hipótesis nula (H_0) también se acepta. Por lo tanto, se puede afirmar que para este punto de acceso la altura no afecta el comportamiento de la intensidad de la señal con una certeza muy cercana al 80%.

Tabla IV. Análisis de Varianza del punto de acceso BS2

Fuente de Variación	Suma de Cuadrados	Grados de libertad	Media de Cuadrados	F	P	F crítica
Entre Grupos	39.79045	3	13.26348	0.341291	0.795506	2.688691
Dentro de Grupos	4197.17	108	38.86269			

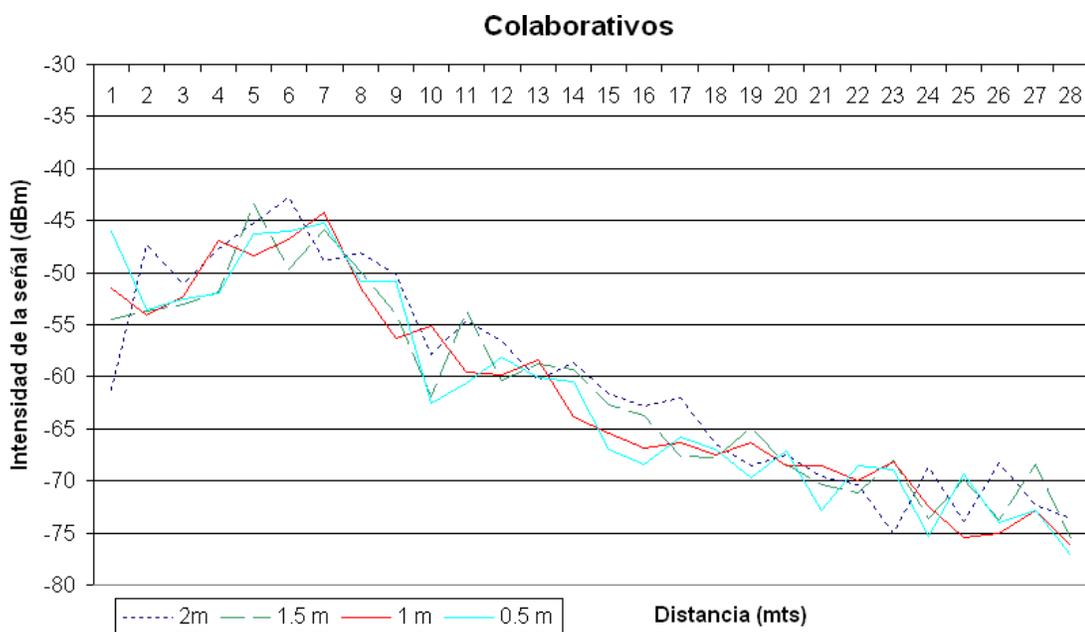


Figura 25. Punto de acceso “colaborativos” y la medición de la señal en altura y distancia

En el caso del punto de acceso denominado “colaborativos” (ver figura 25), como se menciona en la sección IV.3.1., se utiliza un fabricante diferente y se maneja un estándar de la IEEE (803.11b) que difiere (aunque compatible) con el usado en los puntos de acceso anteriores (802.11b/g).

El punto de acceso se encuentra situado en el metro 5 aproximadamente, que es precisamente donde se encuentran los valores más altos de la intensidad de la señal.

En la tabla V se muestra el ANOVA del punto de acceso colaborativos, cuyos resultados nos indican que no existe evidencia estadística de que la altura influya directamente en la intensidad de la señal. En esta tabla se ven los valores F y F-crítica, en el cual el segundo supera al primero. Por lo que, de nuevo, la hipótesis nula se acepta para este caso con una certeza de casi el 95%.

Tabla V. Análisis de Varianza del punto de acceso Colaborativos

Fuente de Variación	Suma de Cuadrados	Grados de libertad	Media de Cuadrados	F	P	F crítica
Entre Grupos	33.64502	3	11.21501	0.120811	0.947676	2.688691
Dentro de Grupos	10025.78	108	92.83133			

En general, se puede apreciar que las mediciones para los tres puntos de acceso son muy irregulares, aún en las lecturas en metros adyacentes. Algunas veces la intensidad de la señal fue más fuerte a una altura que en otra, pero esta diferencia no es constante, es decir que la señal más fuerte no siempre se encuentra la misma altura.

Aunque no se hizo un estudio exhaustivo de la propagación de la señal ni las causas que pudieran originar estas discrepancias en las lecturas, en general, estas diferencias se atribuyen a la estructura física del edificio, así como a los muebles, personas que pasaron al momento de estar tomando las lecturas. Además, el ambiente también se vio influenciado por el ruido provocado por las señales de otros puntos de acceso (de 4 a 8 puntos de acceso dependiendo la zona) que se encuentran configurados para transmitir en canales de frecuencias cercanos a los canales utilizados durante este experimento.

V.1.2. Influencia del tamaño del conjunto de entrenamiento en la estimación de la localización.

En este experimento, como se menciona en la sección IV.3.2., se desea determinar la diferencia en la estimación de la localización si se descartan ciertos datos del conjunto de entrenamiento.

En la sección IV.3.2 se menciona la forma en que los datos de entrenamiento y de prueba son divididos de modo que se obtienen 4 subgrupos de datos de entrenamiento y de prueba.

Este experimento se llevó a cabo utilizando las redes Elman, con una arquitectura 6-16-2, aunque no se obtienen buenos resultados en términos del error, este experimento nos da una idea de la sensibilidad del error derivado del tamaño del conjunto de entrenamiento.

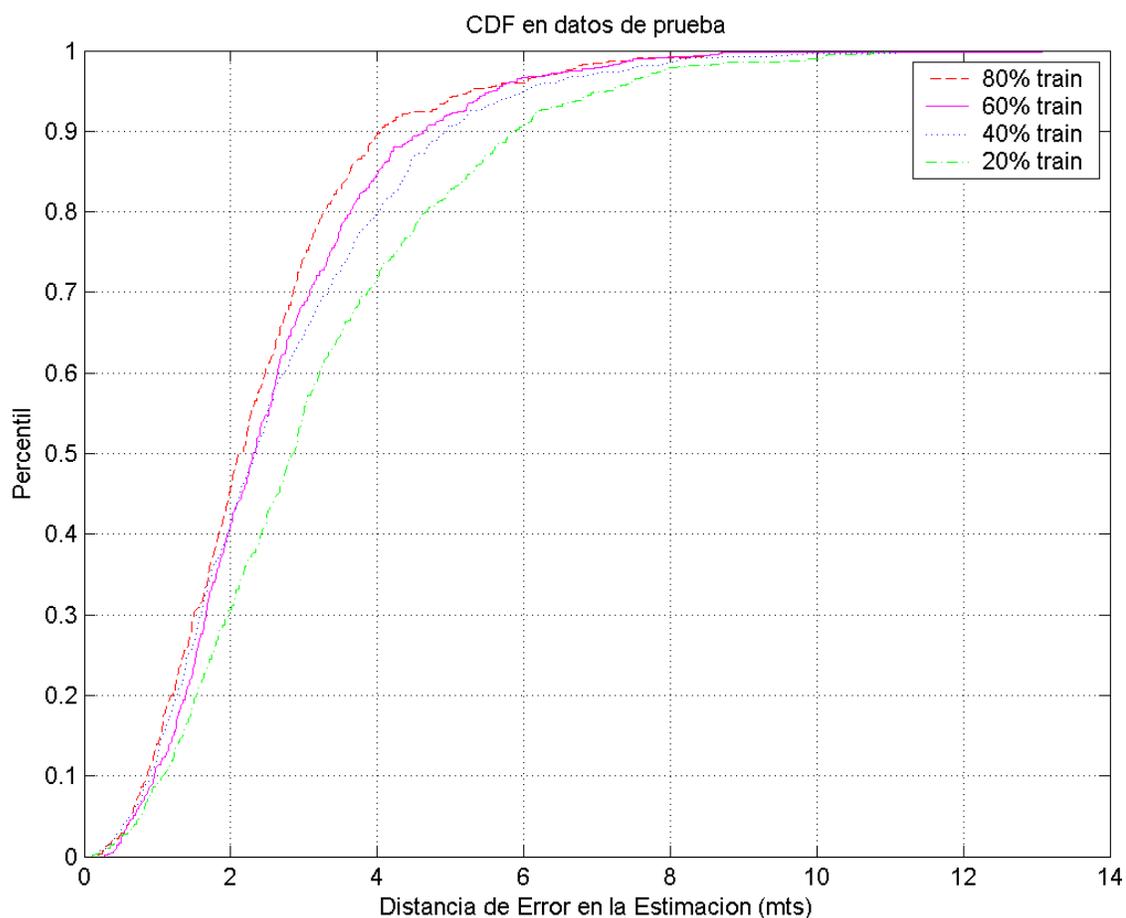


Figura 26. Distribución acumulada para el conjunto de 80, 60, 40 y 20 % en los datos de entrenamiento.

En la figura 26 se muestra la distribución acumulada con cuatro conjuntos de entrenamiento (80%, 60%, 40% y 20%), y con el conjunto de prueba del grupo promediado de 3 muestras. Como se puede observar, a simple vista se puede determinar que la mejor de

las estimaciones corresponde al conjunto de entrenamiento donde se incluye el 80% de los datos. En la tabla VI, se pueden apreciar de una mejor manera los resultados de estas pruebas.

Tabla VI. Tabla comparativa de los diferentes conjuntos de entrenamiento (80%, 60%, 40% y 20%)

Datos de entrenamiento	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
80%	2.4000	1.4731	1.3539	2.1196	3.0337	4.0521	8.7321
60%	2.6105	1.5620	1.5303	2.3159	3.3620	4.6344	13.0783
40%	2.7076	1.7508	1.4080	2.3134	3.6440	4.9473	11.0970
20%	3.2063	1.9694	1.7382	2.8491	4.2666	5.8580	10.9731

De las columnas presentadas en la tabla VI las de mayor relevancia son la de Error (1.0), Media y Desviación Estándar. La de Error (1.0) nos muestra el error máximo en metros que se estimó con los datos de prueba, lo conveniente en este caso es que el error máximo se reduzca de modo que al estimar una posición se tenga una certeza de que la localización real esta en un radio, por ejemplo, de 2 metros, con lo cual, dependiendo de la estructura física del edificio, se podría estar hablando de lo ancho de una habitación.

Con estos resultados se puede hacer un balance entre el tiempo y esfuerzo que se esta dispuesto a invertir en la recolección de muestras a cambio de una mayor precisión en la estimación de la posición. Si se recuerda, la separación de las posiciones durante la toma de lecturas fue de 1 metro, no obstante, el resultado de este experimento sugiere que tomando mediciones a una separación de 2 metros puede ser adecuado para muchas aplicaciones.

V.1.3. Influencia del número de muestras utilizadas para la estimación de la localización.

Como se menciona en la sección IV.3.3., el objetivo de este experimento es el de encontrar una cantidad ideal de muestras para ser promediadas y que sirvan como entrada a la red neuronal. Entre menor sea esta cantidad es mejor, ya que una cantidad mayor

necesariamente obliga a recolectar dichas muestras, lo cual implica un mayor uso de la tarjeta inalámbrica y, en consecuencia, mayor consumo de procesamiento, carga de batería y memoria del dispositivo móvil.

En la figura 27 se muestran los resultados de la estimación usando una red neuronal de Elman con una arquitectura (6-16-2), el entrenamiento se llevó a cabo con los conjuntos promediados de 5 muestras y para las pruebas, como se puede observar, se utilizaron 4 conjuntos.

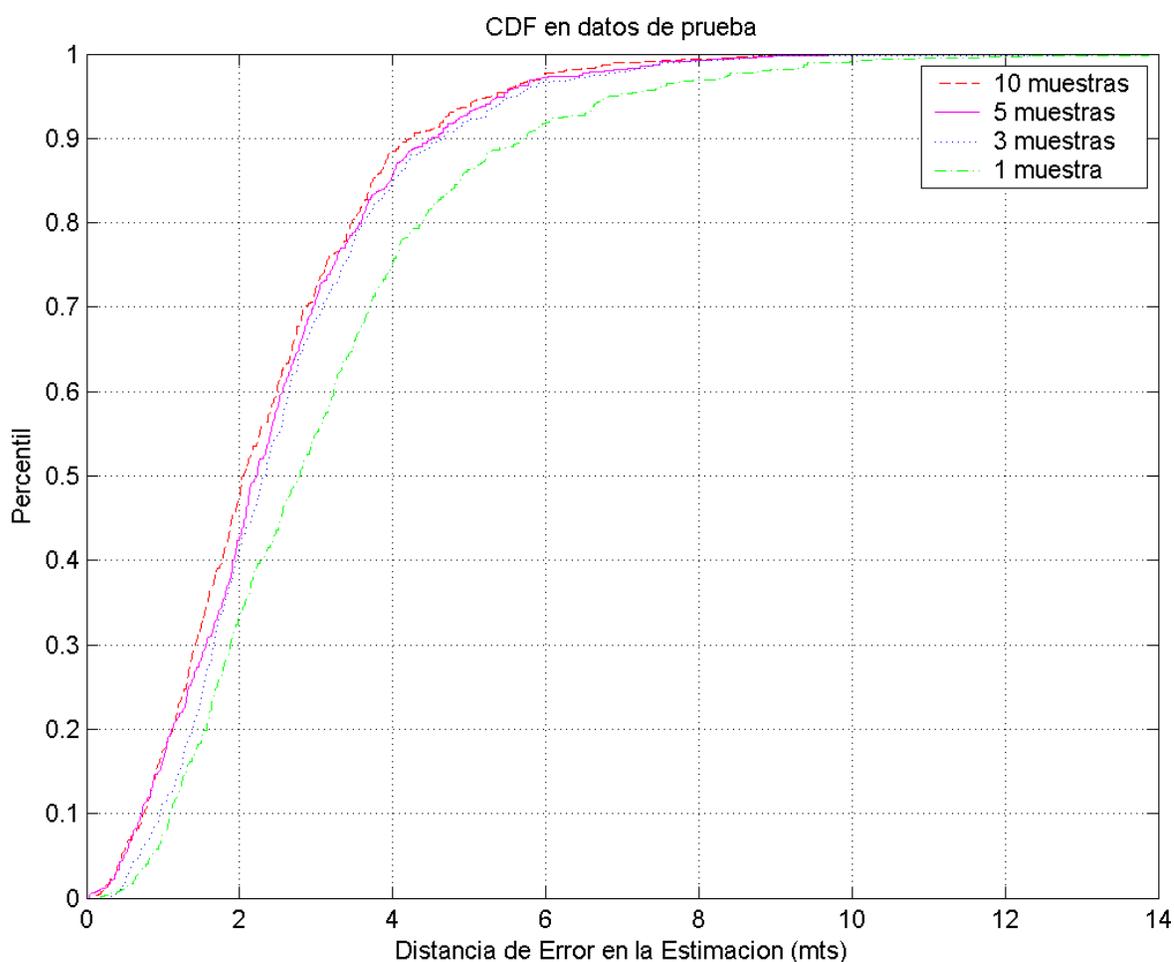


Figura 27. Distribución empírica acumulada para diferentes conjuntos de prueba

En la tabla VII se muestran los resultados obtenidos, teniendo como mejor estimación el conjunto de las muestras promediadas de 10, como era de esperarse. Sin

embargo, por cuestiones prácticas es más conveniente considerar los conjuntos de 3 muestras dado que se usan menos recursos del dispositivo en la obtención de las mismas y los resultados no son muy distintos a cuando se usan 10 muestras.

Tabla VII. Tabla comparativa con los diferentes conjuntos de datos de prueba

Datos de prueba	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
1	3.1355	1.9781	1.7046	2.7729	3.9866	5.6990	13.8718
3	2.6105	1.5620	1.5303	2.3159	3.3620	4.6344	13.0783
5	2.4723	1.5498	1.3386	2.2150	3.2307	4.5140	9.6924
10	2.3629	1.4882	1.3022	2.0338	3.1339	4.2053	9.0570

En la sección V.3.3. se analizan a más detalle estos resultados desde diferentes puntos de vista. A continuación se presentan resultados de eliminar las muestras obtenidas de determinados puntos de acceso.

V.1.4. Influencia del número de puntos de acceso en la estimación de la localización.

La toma de mediciones en el hospital se llevó a cabo con 6 puntos de acceso, sin embargo, hay áreas con menor extensión donde con un número un poco más reducido de puntos de acceso se puede dar un servicio excelente en cuestiones de cobertura y rendimiento de la red.

El objetivo de este experimento es el de corroborar si con menos puntos de acceso aún se podría obtener una estimación de la localización aceptable, dependiendo de la naturaleza de la aplicación.

Este experimento se realizó con redes de Elman, con una arquitectura 6-16-2, los cuales se describen a detalle en la sección IV.3.4. En la figura 28 se muestran los resultados de entrenar con 3 puntos de acceso utilizando, para cada caso, tres puntos de acceso inalámbricos diferentes. La leyenda indica los puntos de acceso que se utilizaron en cada uno de los casos.

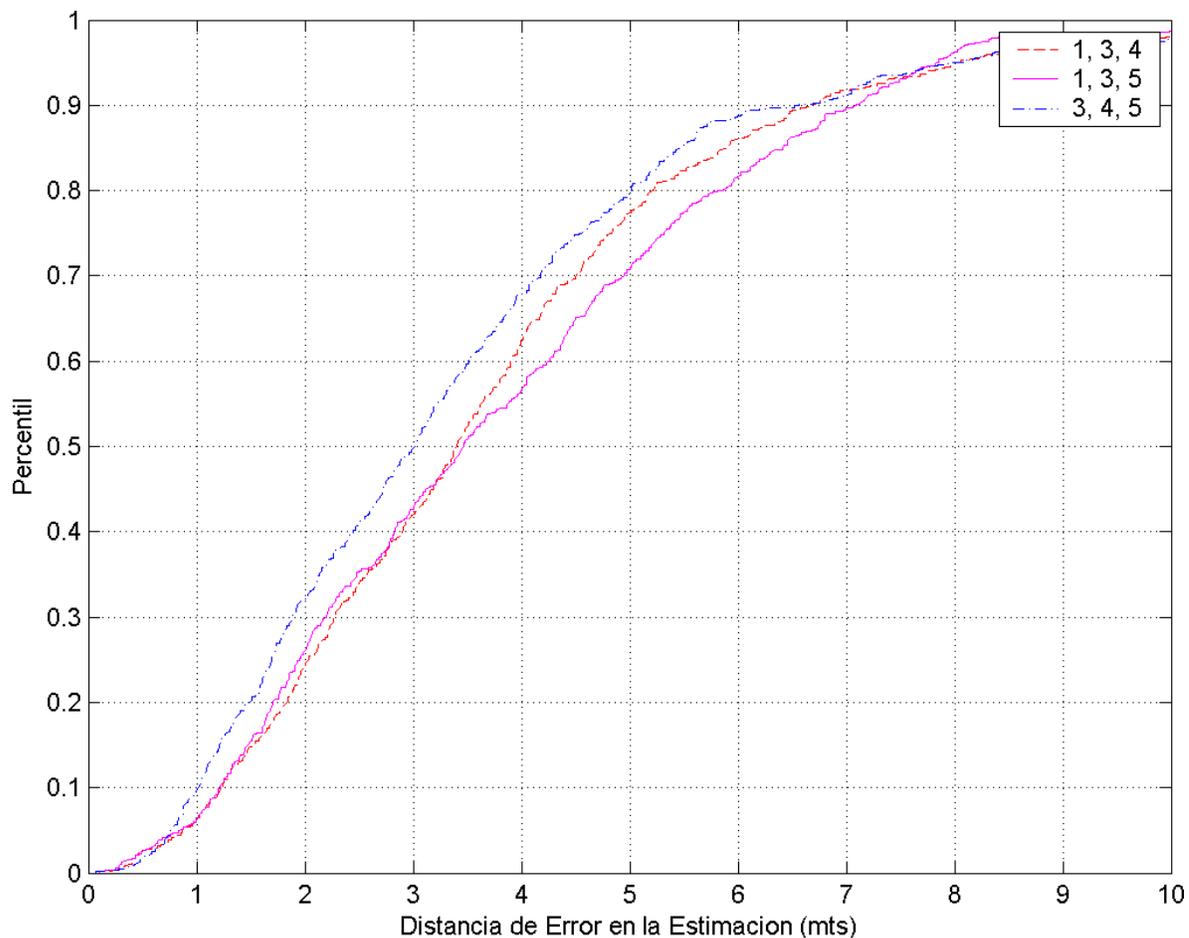


Figura 28. Distribución empírica acumulada para tres puntos de acceso.

En la figura 28 se observa a simple vista que las funciones de distribución son muy parecidas pero si observamos la tabla VIII, veremos que en el caso del error máximo (error 1.0) la función generada a partir de los puntos de acceso 3, 4 y 5, es con mucho la peor, sin embargo si la comparación se hace en términos de la media, es la mejor. Como se menciona anteriormente, la utilidad y superioridad de una distribución de error sobre otra depende de varios factores que están directamente relacionados con la naturaleza de la aplicación en la que se desea utilizar. Para efectos de una mejor apreciación de la gráfica se usan solos valores menores a 10 metros.

Tabla VIII. Tabla comparativa usando tres diferentes puntos de acceso

Puntos de acceso	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
1, 3, 4	3.7485	2.3605	2.0238	3.3986	4.8287	6.6454	16.6599
1, 3, 5	3.8447	2.3659	1.9492	3.4599	5.3223	7.0761	17.7161
3, 4, 5	3.5903	3.1350	1.6956	3.0083	4.5714	6.5232	28.1383

Además del anterior, también se hizo la comparación con 4, 5 y 6 puntos de acceso, de la cual se pueden apreciar los resultados en la figura 29. De igual forma, la leyenda indica los puntos de acceso que se utilizaron en cada uno de los casos.

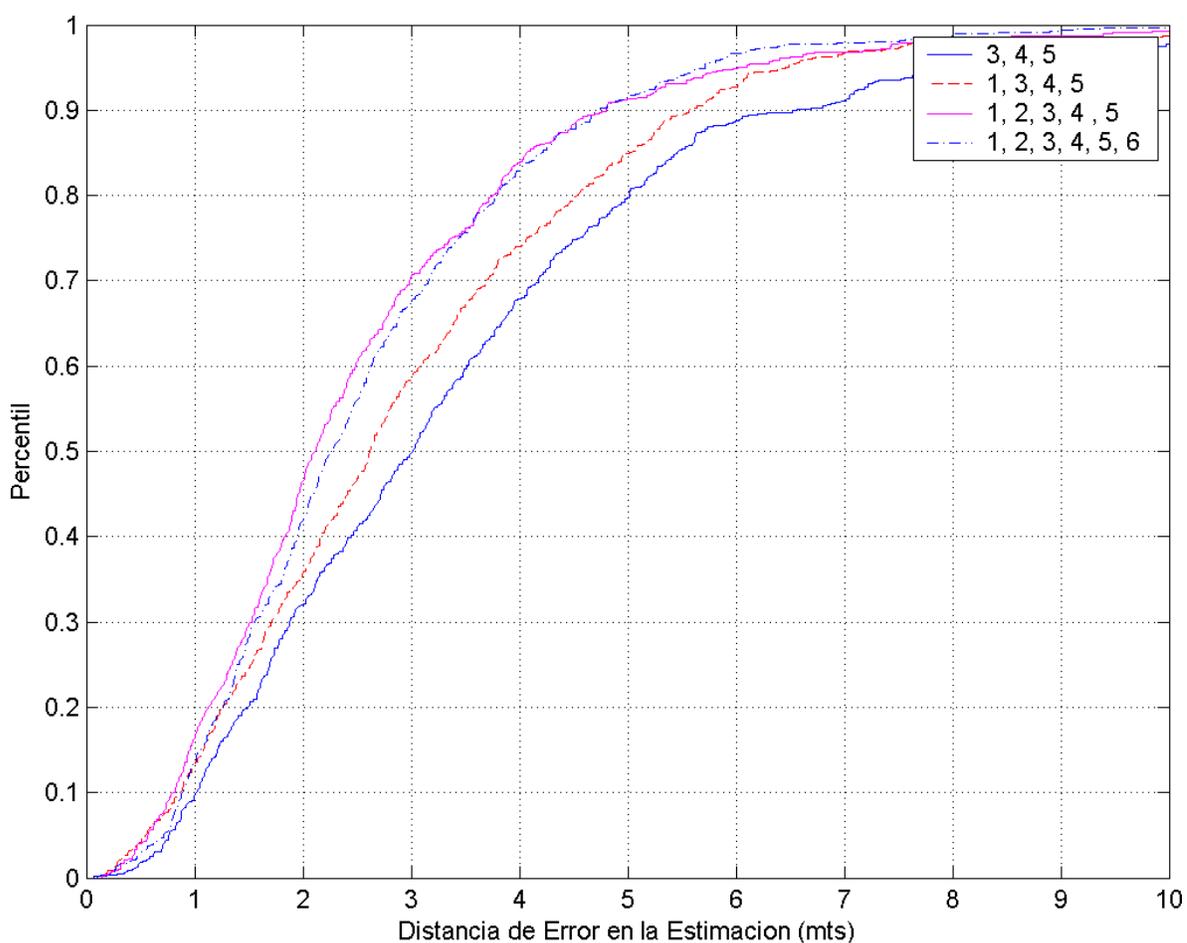


Figura 29. Distribución empírica acumulada de 3, 4, 5 y 6 puntos de acceso

En la figura 29 se puede apreciar que el deterioro que sufrió la estimación de la posición de usar 5 y 6 puntos de acceso fue mínimo. Sin embargo, dicha diferencia se puede atribuir a que el punto de acceso descartado se eligió de manera arbitraria, ya que la posición que mantiene el punto de acceso respecto a la zona de interés en la estimación no tiene mucha relevancia. La tabla IX nos muestra las diferencias entre usar 4, 5 y 6 puntos de acceso.

Tabla IX. Tabla comparativa de 4, 5 y 6 puntos de acceso

Puntos de acceso	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
1, 3, 4, 5	3.0931	2.6255	1.5171	2.6169	4.0723	5.5695	26.2473
1, 2, 3, 4, 5	2.6202	2.1254	1.3427	2.1124	3.3943	4.8051	19.0827
1, 2, 3, 4, 5, 6	2.6191	1.6635	1.3954	2.2515	3.4504	4.7116	13.0596

La naturaleza de la aplicación influye directamente en la cantidad de puntos de acceso a usar. Si se observan las tablas VIII y IX se puede apreciar que el deterioro en la estimación entre el usar 3 y 6 puntos de acceso fue de 1 metro aproximadamente en la media, por lo que para ciertas aplicaciones, usar 3 puntos de acceso en la estimación de la ubicación puede ser adecuado.

V.1.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.

En este experimento lo que se desea medir, como ya se mencionó en la sección IV.3.5., es la variación de la estimación de la localización usando diferentes arquitecturas de red. La figura 30 ilustra la función de distribución empírica acumulada de 7 arquitecturas usadas. En ella se puede apreciar, a simple vista que el que mejores resultados arrojó fue el de 128 unidades en la capa oculta. Sin embargo hay consideraciones prácticas que tienen que hacerse, como lo es el tomar en cuenta el dispositivo que será usado para estimar la posición, es decir ejecutar la red neuronal con los pesos sinápticos que resultaron del entrenamiento.

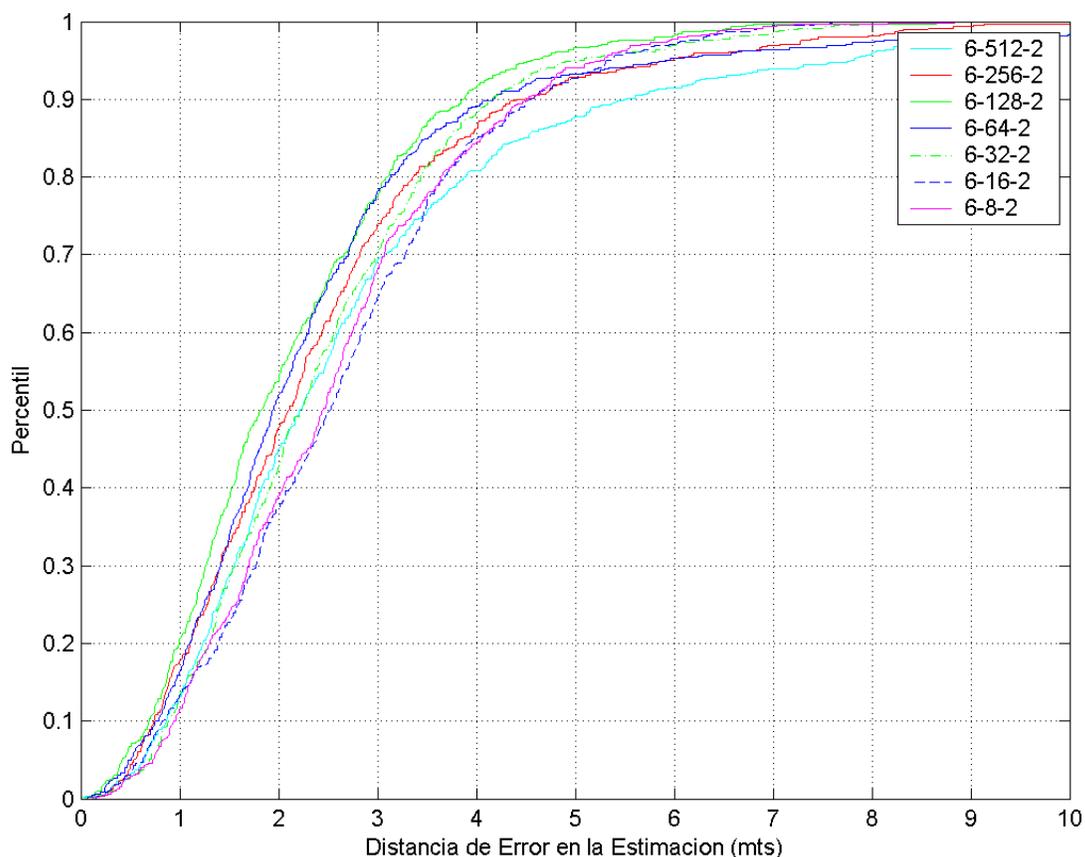


Figura 30. Distribución empírica acumulada de 7 arquitecturas diferentes

Para computadoras de escritorio ordinarias, donde el poder de procesamiento cada día se incrementa, la ejecución de una red neuronal de 128 unidades ocultas se realiza en centésimas de segundo, no así para dispositivos de bolsillo como PDA o celulares donde el poder de procesamiento y memoria son menores y una simulación de este tipo podría tardar alrededor de 5 segundos, sin contar el gasto de batería que involucra hacer los cálculos necesarios.

En la tabla X se muestran los resultados que se obtuvieron de las 7 distribuciones de error mostradas en la figura 30. Para efectos de una mejor apreciación, la gráfica sólo muestra valores no mayores a 10 metros en el eje X. Sin embargo, los resultados completos se pueden apreciar en la mencionada tabla.

Tabla X. Tabla comparativa de 7 arquitecturas de red

Arquitectura	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
6-512-2	2.8131	2.2044	1.3977	2.2141	3.4830	5.4875	14.3289
6-256-2	2.4383	1.7026	1.2769	2.1037	3.0876	4.4012	11.5595
6-128-2	2.1164	1.3386	1.1632	1.8221	2.8669	3.8555	8.7089
6-64-2	2.4012	1.9654	1.2479	1.9408	2.8514	4.1695	14.1684
6-32-2	2.4509	1.4123	1.4113	2.2172	3.2055	4.1445	8.4369
6-16-2	2.6331	1.4429	1.5961	2.5202	3.4625	4.5355	7.6426
6-8-2	2.5800	1.4064	1.5684	2.4476	3.3492	4.5093	8.8311

En este caso, la que mejor resultado ofrece es el usar 128 unidades en la capa oculta, sin embargo se tiene que considerar la capacidad de procesamiento del dispositivo móvil. En nuestro caso, la red neuronal esta implementada en un PDA, por lo que se tomó la decisión de hacer uso de 16 unidades ocultas dado el tiempo que toma el PDA en hacer la simulación con 128 unidades ocultas y que usando 16 unidades ocultas se puede obtener una precisión aceptable en la estimación de la ubicación.

V.2. Entrenando la red neuronal con información de vecindarios.

En este trabajo se busca disminuir el error máximo que se puede obtener. Para esto, se utilizó la idea de que la estimación de la localización debería estar basada en la historia de las estimaciones previas. Es decir, tomar en cuenta las últimas 5 posiciones, por mencionar un ejemplo, y predecir la siguiente ubicación basándose en la trayectoria ó ruta que el usuario estaba siguiendo. De este modo, la posición generada sería una posición muy cercana a las 5 estimaciones pasadas.

La idea detrás de este enfoque es que si se toman las coordenadas como dos variables, x y y , que cambian a medida que el tiempo transcurre pueden ser consideradas como una serie de tiempo. En la figura 31 se puede apreciar la variación de las dos variables en un recorrido de un poco más de 20 metros que hace un médico interno desde la Central de Enfermeras hacia la Jefatura de Medicina Interna.

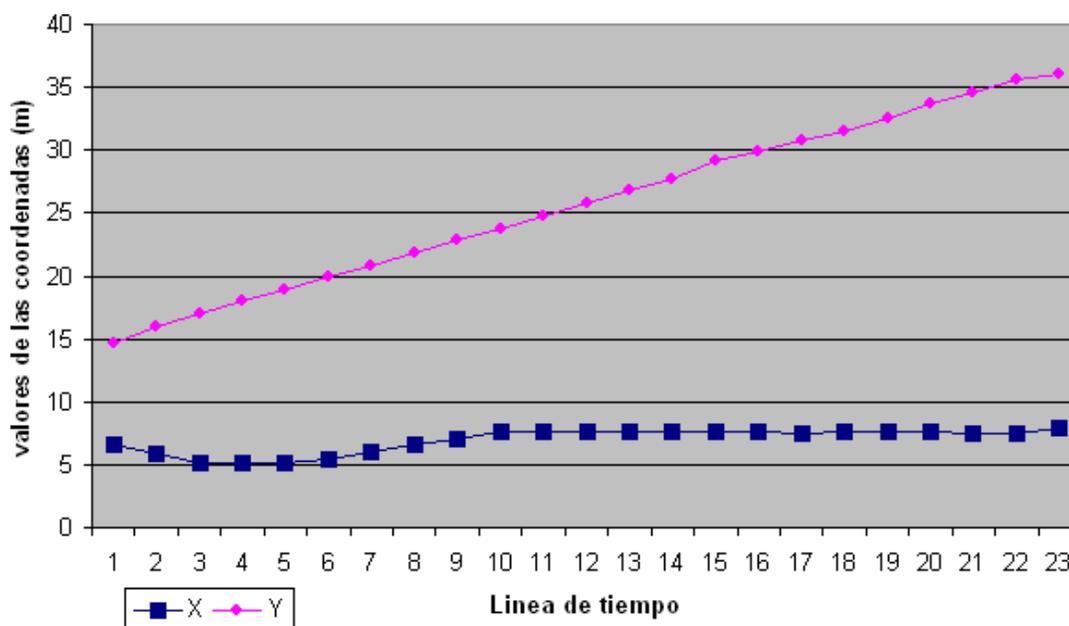


Figura 31. Variación de valores de X y Y en diferentes instantes de tiempo.

Para este experimento se seleccionaron las rutas más probables a ser cubiertas por el personal médico, esto, en base a observación durante las mediciones. Sin embargo, la

movilidad del personal médico es algo compleja y el modelar rutas suficientes para entrenamiento y prueba resulta algo complicado. Los resultados que se obtuvieron fueron basados en rutas no reales. Es decir, que no se hizo un seguimiento de un médico o una enfermera para trazar las rutas sino que estas rutas fueron diseñadas de acuerdo a la estructura física del edificio y en base a simple observación sin hacer un estudio a fondo de los movimientos efectuados por el personal.

Debido a esto, se decidió usar un enfoque más práctico y en cuyo caso, el objetivo era el mismo que el anterior: lograr una estimación más *constante*, es decir, asegurarse que el error que ofrezca como salida la red neuronal tenga un máximo de pocos metros, 3 por ejemplo.

Para lograr esto se preparó un experimento como se describe en la sección IV.3.6. en el que, como se menciona, se usan coordenadas vecinas a la posición que se desea estimar. De este modo, se entrena a la red indicándole que la posición que esta estimando es muy cercana a donde estuvo el usuario la última vez.

Tabla XI. Resultados del entrenamiento con coordenadas vecinas.

Vecinos	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
BP Con	0.9131	0.4045	0.6212	0.9242	1.2185	1.4272	2.1761
BP Sin	1.9827	1.0685	1.1952	1.8595	2.5726	3.3883	6.9550
Elman Con	0.8888	0.4059	0.5914	0.9059	1.1748	1.4050	2.3592
Elman Sin	2.1633	1.3289	1.2532	1.9099	2.7216	3.6847	8.8095

En la tabla XI se muestran los resultados con los datos de prueba, resultado del entrenamiento con vecinos, además de los resultados que se obtuvieron sin usar coordenadas vecinas. Ambos experimentos se hicieron con redes de retropropagación (BP) con arquitectura 8-16-2 y Elman con arquitectura 8-16-2.

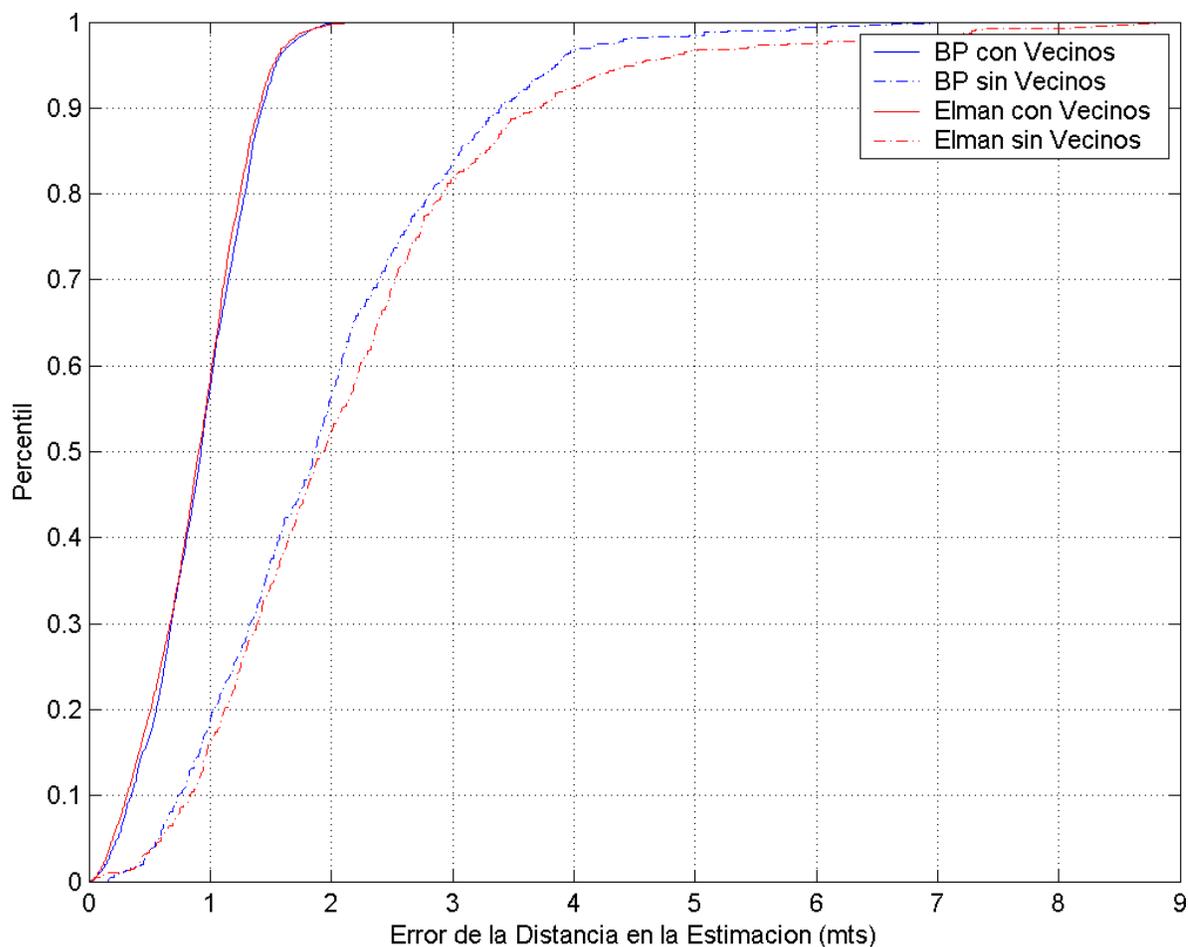


Figura 32. Distribución acumulada del error del entrenamiento con y sin vecinos.

En la figura 32 se muestran los resultados con los datos de prueba, cabe señalar que los vecinos en los datos de prueba fueron seleccionados, al igual que en los datos de entrenamiento, es decir, de manera arbitraria.

En este experimento se obtuvieron muy buenos resultados, con un error promedio por debajo del 1 metro cuando se utilizan coordenadas vecinas. Con lo cual se diseñó e implementó una aplicación para probar este enfoque, la cual se describe en el apéndice B.

Este enfoque también se probó con mediciones tomadas en el lugar de prueba, sólo que en lugar de ser mediciones en diversos puntos estáticos, éstas fueron tomadas siguiendo

rutas predefinidas similares a las que siguen los médicos internos cuando hacen sus rondas (véase figura 33).

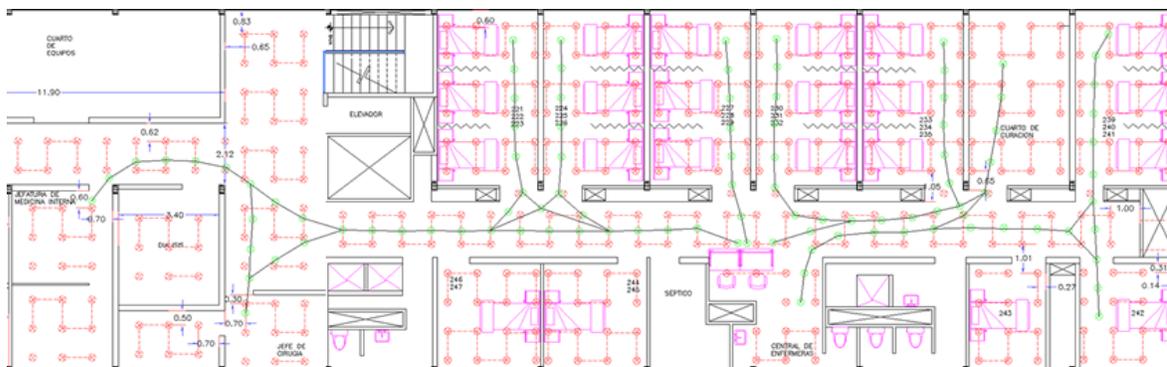


Figura 33. Rutas sobre las que se tomaron mediciones de manera continua en el hospital.

Para la prueba se usaron dos redes neuronales de retropropagación con arquitecturas 6-16-2 y 8-16-2. La primera es la encargada de proporcionar la primera estimación usando sólo la intensidad de los 6 puntos de acceso. Una vez estimada dicha posición es proporcionada como parte de la entrada de la segunda red neuronal junto con las 6 intensidades de la señal. A partir de entonces, cada salida generada por la red neuronal se convierte en parte de la entrada de la misma red en el siguiente instante de tiempo. De este modo, las posiciones estimadas son posiciones muy cercanas a donde se estuvo por última vez.

En la figura 34 se puede apreciar la función de distribución empírica de esta prueba usando información de vecindarios, así como también se presenta una distribución donde no se usó la información del vecindario. Se usaron las redes de retropropagación y las redes Elman.

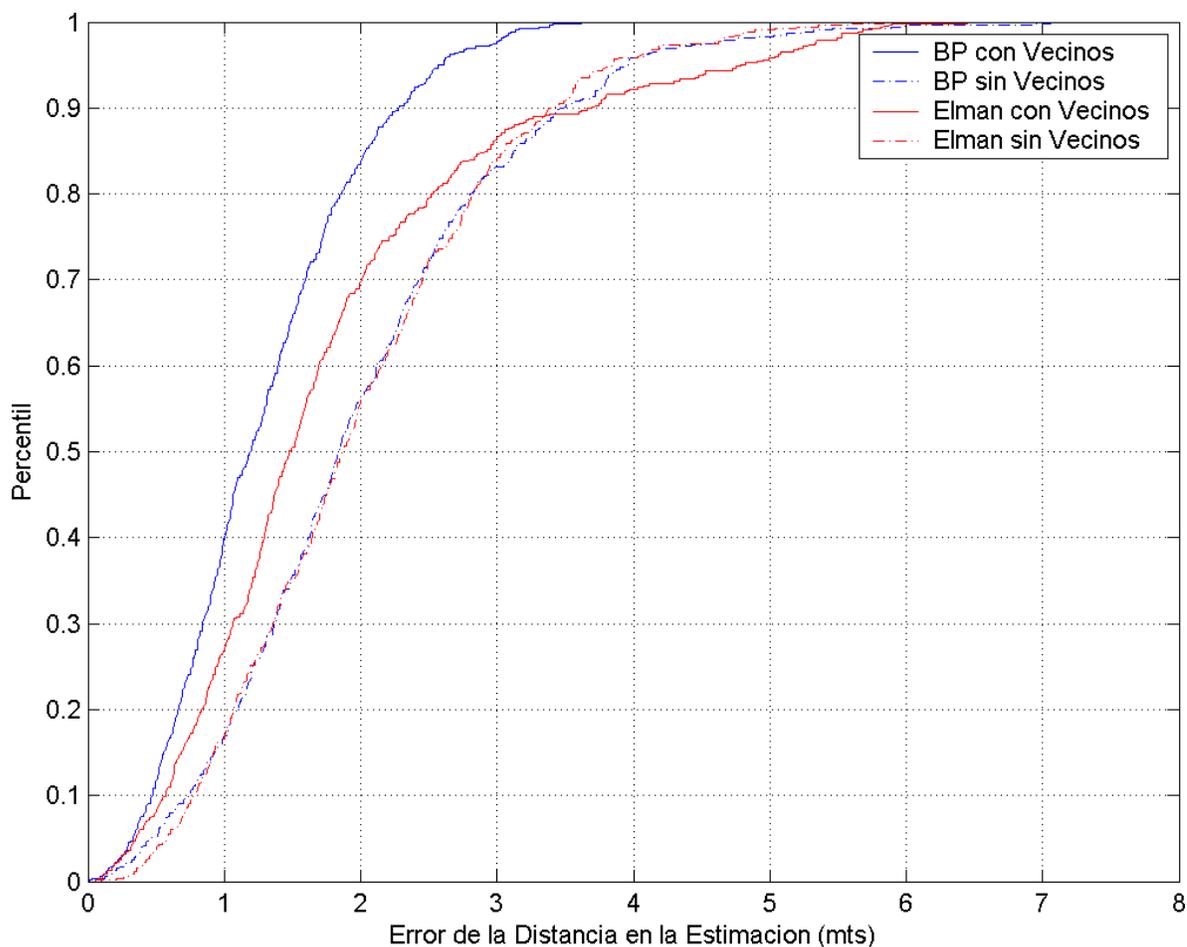


Figura 34. Distribución empírica acumulada del error con y sin vecinos con rutas reales.

En la tabla XII se muestran los resultados con los datos de la ruta real, resultado del entrenamiento con vecinos, además de los resultados que se obtuvieron sin usar la posición anterior para los dos tipos de redes.

Tabla XII. Resultados del entrenamiento con coordenadas vecinas.

Vecinos	Media	Desv. estándar	Error (0.25)	Error (0.5)	Error (0.75)	Error (0.9)	Error (1.0)
BP Con	1.2889	0.6991	0.7666	1.1953	1.7273	2.271	3.6204
BP Sin	2.0086	1.1023	1.2111	1.8432	2.5989	3.484	7.0575
Elman Con	1.7941	1.2501	0.9443	1.4763	2.2191	3.6633	6.4498
Elman Sin	2.0044	1.0304	1.1920	1.8620	2.6710	3.3804	5.8396

Los datos de la tabla XII presentan evidencia de que, efectivamente, usando una posición anterior mejora la estimación de la posición, no sólo en la media, sino en la *estabilidad* de las estimaciones. Si se observa la figura 35, se puede apreciar la diferencia en metros entre estimaciones consecutivas que se obtuvieron siguiendo una ruta predefinida.

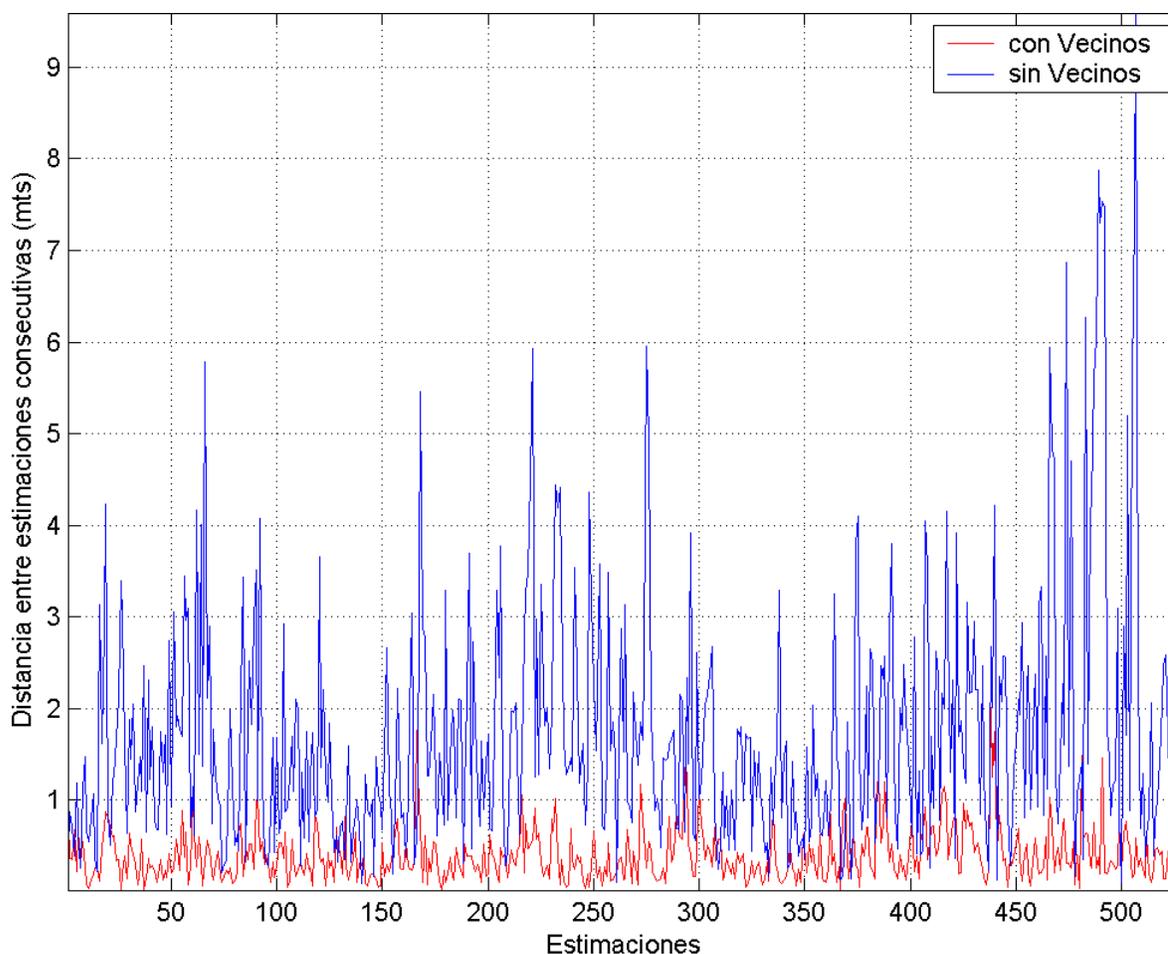


Figura 35. Variación de la distancia entre estimaciones consecutivas usando información de vecindarios, así como la variación de la distancia entre estimaciones consecutivas sin usar dicha información de vecindarios.

En el eje X se grafican el número de estimaciones que se hicieron durante la ruta mencionada. Por otro lado, en el eje Y se muestra la distancia entre estimaciones consecutivas. Lo que esta gráfica muestra es la *estabilidad* de la estimación al incluir la

posición anterior como parte de la entrada en la red neuronal. Al hacer esto, se evita que la red este obteniendo resultados que, al observarlos en una animación por computadora, hagan parecer que el usuario *salta*, en cuestión de segundos, entre ubicaciones alejadas por varios metros. La distancia promedio entre estimaciones consecutivas al usar una posición anterior fue de 0.3976 metros, mientras que sin usar la posición anterior fue de 1.7347 metros.

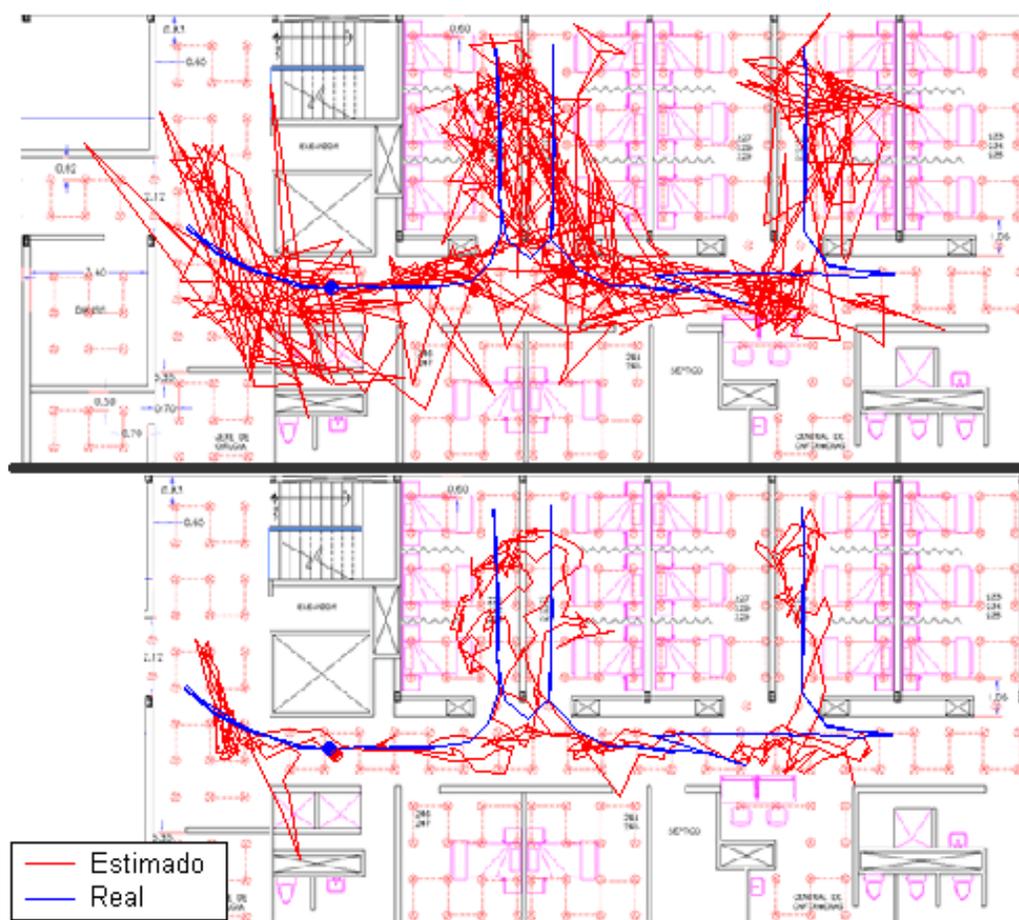


Figura 36. Ruta real usada para la prueba de los dos enfoques: sin usar posición anterior (arriba) y usando una posición anterior como parte de la entrada de la red (abajo)

En la figura 36 se aprecia de manera gráfica las variaciones y los *saltos* de la estimación cuando no se toma en cuenta la última estimación (arriba), de igual forma se

presenta la ruta cuando se toma en cuenta la última estimación como parte de la entrada de la red neuronal (abajo). La línea roja representa la ruta estimada y la ruta azul representa la trayectoria real donde se tomaron las mediciones. La ruta azul no es una ruta siempre hacia adelante, es decir que durante las mediciones se hicieron pausas en determinadas posiciones, así como también se retrocedió en algunos puntos. Esto para tratar de simular el comportamiento de una persona cuando se detiene a conversar con alguien por algunos segundos.

V.3. Discusión de los Resultados obtenidos.

En esta sección se discuten los resultados obtenidos y descritos en la sección V.1. y V.2 desde tres perspectivas: los resultados aislados de los experimentos, los resultados reportados en la literatura en experimentos similares y, por último, los resultados de los experimentos en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

V.3.1. Influencia de la altura del dispositivo móvil en la intensidad de la señal.

Los resultados descritos en la sección V.1.1. nos demuestran que, efectivamente, no existe una diferencia estadísticamente significativa en la intensidad de la señal tomada a diferentes alturas. Lo cual, aplicado en cuestiones prácticas, nos sirve para tener una estimación confiable y similar para personas de diferente estatura.

Resultados reportados en la literatura en experimentos similares.

En RADAR (Bahl y Padmanabhan, 2000), se hace también un análisis sobre la orientación del usuario, se concluye que la orientación tiene un efecto importante sobre la intensidad de la señal. Dejan claro que la orientación influye debido a la obstrucción del cuerpo de la persona que carga el dispositivo móvil, además hacen experimentos que demuestran que el usar información de orientación completa puede mejorar en promedio 70%, en comparación de si no se usara.

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

En este experimento se consideraron cuatro alturas diferentes sólo para conocer si existe una variación significativa que nos pudiera afectar al momento de estar estimando la posición en tiempo real. Esto, debido a que existen personas más altas que otras. Sin embargo, el uso normal del dispositivo hace pensar que la altura promedio donde se estará usando el dispositivo será entre 1 metro y 1.5 metros del suelo por lo que el efecto que puede tener en la estimación es mínimo.

V.3.2. Influencia del tamaño del conjunto de entrenamiento en la estimación de la localización.

Aunque no se muestran todos los resultados de los experimentos que se hicieron en este rubro, se presentan los más significativos. Este experimento sirve de fondo para descubrir la sensibilidad del error con respecto al tamaño del conjunto de entrenamiento. La cual es conveniente conocer para el caso donde se desea hacer una recolección de muestras más rápida, de menor costo y, a la vez, obtener una estimación confiable dependiendo de la naturaleza de la aplicación.

Resultados reportados en la literatura en experimentos similares.

En RADAR (Bahl y Padmanabhan, 2000) se hace un experimento en el cual se reduce la cantidad de posiciones físicas que se usan en la estimación de la localización. Hacen varias corridas usando posiciones en un rango desde 2 hasta 70 elegidas de manera aleatoria. Ellos concluyen, al final, que, para el lugar donde se hicieron las pruebas, su algoritmo obtendría resultados muy similares si se incluyen solamente 40 de los 70 puntos en los que hicieron mediciones. Algo muy similar a lo que se encontró en este trabajo, en el que con el 60% de las posiciones físicas originales se pueden obtener resultados muy buenos y similares a los que se obtienen usando el 80% de las localizaciones originales.

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

La gente que se encuentra en un hospital, pacientes o familiares de los mismos, esta muy sensible dadas las circunstancias, por lo que una reducción en la toma de mediciones, es decir, en el tiempo que se pasa al lado de las camas y dentro de los dormitorios puede ser de gran ayuda para evitar molestia en los pacientes y/o familiares.

Por otra parte, para efectos de su aplicación en un sistema consciente del contexto, debe tomarse en cuenta a qué detalle se desea estimar la posición y el error que se esta dispuesto a tolerar. Si mencionamos el ejemplo donde la posición esta definida por el paciente con el que se encuentra determinado médico, será muy importante diferenciar a un paciente de otro. Sin embargo, si nos interesa saber sólo el área donde se encuentra dicho médico, se pudiera considerar el hacer una etapa de medición de la señal inalámbrica rápida y tomar mediciones sólo en algunos puntos en el área donde se desea estimar la posición.

V.3.3. Influencia del número de muestras utilizadas para la estimación de la localización.

Los resultados de este experimento nos muestran que se puede obtener una estimación de la localización confiable a partir de 3 muestras promediadas. Aumentar la cantidad de muestras que se promedia no incrementa considerablemente la exactitud de la estimación. Sin embargo, es un aspecto que se debe tomar en cuenta considerando el uso que tendrá, en los párrafos subsecuentes se describe la utilidad de los resultados en términos de su aplicación.

Resultados reportados en la literatura en experimentos similares.

En RADAR (Bahl y Padmanabhan, 2000), al igual que en nuestro trabajo, se hace una evaluación del impacto que tienen la cantidad de muestras utilizadas en la estimación. Para estimar la localización usan la media de todas las muestras recolectadas en una determinada localización. Sin embargo, mencionan la dificultad de tomar 20 o más lecturas

en tiempo real, es decir durante el uso real del sistema, por lo que hacen un análisis con una simple muestra, con dos muestras y con tres. Dando como resultado que el usar una sola muestra es aproximadamente 30% peor que usando todas. Con dos muestras es 11% peor que usando todas las muestras y con 3 muestras es alrededor de 4% peor.

No se menciona si se hicieron más experimentos, pero es posible, dados los resultados que utilicen, al igual que nosotros, los promedios de grupos de 3 muestras.

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

El diseño de aplicaciones conscientes del contexto para hospitales incluye una gran variedad de dispositivos, entre los que se encuentran PDA's, los cuales, como se sabe, son dispositivos limitados en cuanto a su capacidad de almacenamiento, velocidad de procesamiento y duración de la batería.

El procesamiento es uno de los aspectos que más energía consume de la batería del PDA, al igual que la recepción de señal en la tarjeta inalámbrica. Por lo que al efectuar estimaciones es necesario tomar en cuenta cuántas muestras serán consideradas para el cálculo de la estimación en el PDA.

Los resultados nos muestran que la diferencia en la estimación entre tomar en cuenta 3 y 10 muestras es de 0.2476 metros. Lo cual es despreciable si se toma en cuenta que el área ocupada por una persona de pie en una habitación es de 0.3 x 0.5 metros aproximadamente. Sin embargo, el dejar de tomar 7 muestras nos deberá ahorrar una cantidad de batería considerable en el transcurso de minutos.

V.3.4. Influencia del número de puntos de acceso en la estimación de la localización.

La cantidad de puntos de acceso usados en una red inalámbrica muchas veces esta determinado por su costo. Por lo que lo ideal es tener un número reducido de ellos y, a la

vez, tener una aproximación de la localización aceptable dependiendo de la naturaleza de la aplicación, que es, en última instancia, quien decide el máximo error tolerable.

Otro punto a considerar en este caso es que al tener más puntos de acceso, son más las entradas que se tienen a la red neuronal, con lo que se incrementa el costo de procesamiento.

Resultados reportados en la literatura en experimentos similares.

En (Prasithsangaree et al., 2002) realizaron un experimento de este tipo con el fin de ver la tolerancia a fallos de su algoritmo. Lo que hicieron fue deshabilitar un punto de acceso cuando su algoritmo estaba leyendo las señales de los puntos de acceso. Como era de esperarse, el error aumentó, sin embargo solo aumentó muy poco en relación a si no hubiera habido un fallo en un punto de acceso. La razón de la robustez es que hay muy pocas localizaciones en la base de datos que concuerden con la posición del dispositivo móvil.

En la segunda versión de RADAR (Bahl y Padmanabhan, 2000) lo que desean saber es, al igual que en este trabajo, la sensibilidad del error con respecto a la cantidad de puntos de acceso usados para la estimación de la posición. Hicieron 5 experimentos donde utilizaron desde 1 hasta 5 puntos de acceso para hacer la estimación. Concluyen que hay poco beneficio en términos de disminución de error a partir de 3 puntos de acceso.

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

Los resultados de este experimento nos muestran la sensibilidad del error con respecto a la cantidad de puntos de acceso usados en la estimación de la localización. Sin embargo, no implica necesariamente que pueda adaptarse dinámicamente para trabajar, en determinado momento, con un punto de acceso menos, en caso de que falle uno, por ejemplo. Para este caso, es necesario hacer un entrenamiento con cada una de las

configuraciones posibles de los puntos de acceso y la aplicación consciente del contexto es la que se encargará de decidir o detectar cuantos puntos de acceso están en operación.

V.3.5. Influencia de la arquitectura de la red neuronal en la estimación de la localización.

Los resultados de este experimento nos muestran que no es necesario tener tantas unidades ocultas para obtener una aproximación adecuada. De hecho, se puede apreciar en la sección V.1.5 que no es necesario tener una cantidad muy grande de unidades ocultas, el experimento sugiere que basta con 32 unidades ocultas para tener una buena estimación.

Resultados reportados en la literatura en experimentos similares.

En (Martinez et al., 2004) se tiene una comparación de algoritmos de entrenamiento, dentro del cual, para demostrar la superioridad de un algoritmo sobre otro, se decide reducir la cantidad de unidades ocultas en la red neuronal en determinados algoritmos. Aunque sólo se hace el experimento con 2 arquitecturas diferentes (8 y 16 unidades ocultas), es importante señalar que el tener más unidades ocultas no necesariamente implica un mejor resultado. El tener más unidades ocultas favorece el sobre-entrenamiento, dando lugar a un mayor error de prueba (Bishop, 1995). De los 7 algoritmos presentados en tal experimento, la arquitectura 5-8-2 fue igual o mejor que otros 4 algoritmos al 90% del percentil.

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

Este experimento es muy importante para efectos de la aplicación práctica de una red neuronal, sobre todo en dispositivos como PDA donde, como se menciona en la sección V.2.3., las restricciones en almacenamiento, procesamiento y batería son determinantes al momento de decidir, en este caso, la cantidad de unidades ocultas, ya que son ciclos adicionales de procesamiento que se deben ejecutar y, por ende, el consumo de poder de procesamiento, pila y memoria se incrementa. Es por eso que ésta debe ser también una

consideración a tomarse en cuenta cuando se este planeando llevar a la práctica un trabajo como este.

V.3.6. Entrenando la red neuronal con información de vecindarios.

El propósito de este experimento es el de ayudar a reducir el error máximo en el que, en estimaciones consecutivas hacen pensar que el usuario pasa de una posición a otra alejadas por varios metros en cuestión de segundos. Los resultados obtenidos en este experimento son muy buenos ya que se ha podido reducir el error máximo a un poco más del doble del error promedio (0.9799 metros). Asimismo, en el caso donde se probó con una ruta real el error máximo fue de 3.6204 metros, una cantidad muy cercana a la mitad de lo que se obtuvo sin usar el vecindario donde fue de 7.0575 metros.

Resultados reportados en la literatura en experimentos similares.

En la segunda versión de RADAR (Bahl y Padmanabhan, 2000) se implementó un algoritmo, cuyo objetivo era el de considerar estimaciones del pasado para reducir el error en la estimación. La idea detrás de este algoritmo es el de evitar que el usuario *salte* grandes distancias de una posición a otra en cuestión de segundos. Es muy probable que el usuario se encuentre en una posición muy cercana a donde se encontraba unos instantes de tiempo antes. Uno de los beneficios de este enfoque es el de disminuir la confusión en el espacio de la señal de dos posiciones separados por una distancia considerable.

Lo interesante de este enfoque es que ofrece resultados que son aproximadamente 50% mejor que NNSS (*Nearest Neighbor in Signal Space*), algoritmo que fue desarrollado por ellos mismos pero sin considerar información del pasado. La descripción del algoritmo se puede encontrar en (Bahl y Padmanabhan, 2000).

Los resultados en términos de su empleo en aplicaciones conscientes del contexto en hospitales.

El resultado alcanzado en este experimento es de gran importancia para una aplicación consciente del contexto en un ambiente de hospital. Esto, debido que las camas

de un paciente miden aproximadamente 90 cm. de ancho., pero el espacio ocupado por un paciente en realidad es de 2 x 3 metros aproximadamente. Tomando en cuenta esto es posible conocer la identidad de un paciente dependiendo de la posición estimada y, de esta forma, acceder a información relacionada a él.

V.4. Resumen.

En este capítulo se presentaron los resultados de los diferentes experimentos descritos en el capítulo anterior tal como son la altura del dispositivo móvil en la intensidad de la señal, cantidad de muestras que se deben tomar en cuenta para tener una aproximación de la localización aceptable, cantidad de muestras que deben ser tomadas en cuenta al momento de estimar la posición, la variación en el error dependiendo del número de puntos de acceso utilizados y , al final, se elabora sobre un experimento en el cual se obtuvieron muy buenos resultados: el entrenamiento de la red con la ayuda de coordenadas vecinas al punto que se desea estimar.

Estos experimentos nos ayudan cuando se desea planear la toma de mediciones en algún lugar así como tomar consideraciones de diseño de una aplicación consciente del contexto. Uno de los aspectos en los que estos experimentos pueden ser de mayor utilidad es en la disminución de costo, ya sea por la cantidad de puntos de acceso para efectuar la estimación, al igual que la disminución de tiempo en la toma de mediciones.

En el siguiente capítulo se presentan las conclusiones obtenidas respecto a este trabajo de tesis, de las cuales algunas pueden ser tomadas como consideraciones prácticas para cuando se quiera realizar un trabajo como este para aplicaciones no solo en hospitales sino en cualquier otro lugar. De igual forma, se presentan las contribuciones y trabajo futuro propuesto para la mejora de los resultados reportados en este trabajo.

Capítulo VI.

Conclusiones y Trabajo Futuro.

VI.1. Conclusiones.

En este trabajo de tesis se presentó un estudio en el cual el objetivo principal fue el de elaborar un método práctico de localización de usuarios. Para esto, se tomaron mediciones en varias partes de un edificio de manera que se pudieran tomar en cuenta de manera continua considerando tiempos y lugares.

La utilización del vecindario como ayuda en el entrenamiento de la red neuronal disminuyó el promedio de error en la distancia en un 53.94% en comparación con experimentos realizados sin el uso de los vecinos para el caso de las redes de retropropagación y en un 58.91% en el caso de las redes Elman. Por lo tanto, la utilización de una posición anterior como entrada de la red neuronal al momento de estimar la posición de una computadora móvil disminuye notablemente el promedio de error en la distancia.

Se realizaron experimentos con diferente número de patrones de entrenamiento. El conjunto donde se uso el 80% de los datos fue el que obtuvo el mejor resultado. Sin embargo, si se observan los resultados del 80% y el 20% el error promedio es 4 y 6 metros respectivamente, pero los resultados que se obtuvieron con los conjuntos de 60% y 40% son muy similares a los obtenidos con el conjunto de 80%, lo que sugiere que tomando mediciones a una resolución de 2 metros puede ser adecuado para muchas aplicaciones.

De igual forma se presentó un análisis con diferentes conjuntos generados a partir de los datos de prueba. Se generaron 4 grupos de prueba, con conjuntos de muestras promediadas de 1, 3, 5 y 10 muestras. Se encontró, en base a los resultados reportados, que es necesario tener al menos 3 lecturas para obtener una estimación confiable. Sin embargo, no es recomendable hacer una estimación con más de 3 muestras promediadas ya que

incrementar de 3 a 10 muestras no reduce el error de manera significativa. Y, por otro lado, el costo adicional en uso de la batería, y el poder de procesamiento podría verse afectado.

Finalmente, las redes neuronales son un enfoque adecuado para estimar la posición de usuarios móviles en interiores. Sin embargo, la forma de entrenar la red neuronal utilizada tiene un inconveniente, es decir que es posible prácticamente atravesar paredes, lo cual, en la vida real es imposible. Por lo que se sugiere complementar el entrenamiento con alguna técnica con la cual se puedan hacer ciertas restricciones en los lugares por donde no es posible transitar en el edificio. Dicho enfoque ya ha sido utilizado en otros trabajos (Krumm et al., 2002; Liao et al., 2003) obteniendo como resultado la disminución de error y consiste básicamente en trazar un grafo conectado sobreponiéndolo a la estructura física del edificio de modo que los arcos representan lugares por donde es posible transitar.

VI.2. Aportaciones.

El objetivo principal de este trabajo era el de proponer un método por medio del cual se obtenga, de una manera práctica y eficiente, una estimación de la localización aceptable en términos del error que se desea obtener. En ese sentido, la principal aportación es la descripción del procedimiento que debe seguirse para la obtención de un error en la aproximación de la posición real de usuario del orden de 1 metro.

A continuación se listan otras aportaciones que se hicieron en el transcurso de la realización este trabajo de tesis:

- Una aplicación con la cual tomar mediciones de la intensidad de la señal, la única adecuación que tiene que hacerse, sin necesidad de reprogramar, es copiar un plano arquitectónico del lugar donde desean hacerse las mediciones. Además, ajustar los valores de *largo* y *ancho* del edificio donde han de hacerse tales mediciones.
- El diseño de una herramienta práctica con la cual señalar físicamente el lugar donde han de tomarse las mediciones.

- Un análisis estadístico en la comparación de la toma de mediciones desde diferentes alturas. Es decir, que se tomaron lecturas desde diferentes alturas para ver si existe una diferencia debido a factores como interferencia en la señal debido al techo y al piso.
- La manera de entrenar la red neuronal para aprovechar la última posición en la que estuvo el usuario. Con este enfoque se obtienen mejores resultados ya que se le está indicando a la red neuronal que la posición actual que desea estimar es muy cercana a la última posición conocida.
- Un estudio sobre la sensibilidad del error con respecto a la cantidad de puntos de acceso, así como tamaño del conjunto de entrenamiento y, finalmente, la influencia de la arquitectura de la red neuronal.
- De la misma manera, un estudio sobre la sensibilidad del error con respecto a la cantidad de muestras para promediar que han de tomarse en cuenta al momento de estar haciendo la estimación de la localización.
- Una aplicación con la cual se puede hacer la estimación de la posición en una computadora de bolsillo. En esta aplicación, lo que se hace, básicamente es implementar las formulas descritas en la sección III.3.2.
- La publicación de un artículo de investigación que describe parte de los resultados obtenidos en este trabajo: Luis A. Castro, Jesús Favela. "Continuous Tracking of User Location in WLANs Using Recurrent Neural Networks". En las memorias del Sexto Encuentro Internacional de Ciencias de la Computación (ENC 2005), IEEE Computer Press, Puebla, México.

VI.3. Limitaciones.

Este trabajo posee ciertas limitaciones inherentes debido a las propiedades de la señal inalámbrica. A continuación enumeran algunas de las principales:

- Algunos cambios radicales en la estructura física del edificio provocará que el sistema se comporte de manera inadecuada.
- La variación en la cantidad de personas al momento de hacer la estimación puede impactar en la estimación. Por ejemplo, en la hora de la visita de familiares.
- La batería del dispositivo limita, hasta cierto punto, el seguimiento continuo de usuarios.
- La capacidad de procesamiento del dispositivo usado para las estimaciones limita la cantidad de muestras que se pueden tomar al momento de estar haciendo la estimación.

VI.4. Trabajo futuro.

Los resultados obtenidos en la serie de experimentos que se realizaron durante el presente trabajo de tesis fueron muy buenos, sin embargo, en ese sentido, aún queda trabajo por hacer. A continuación se enumeran una serie de adecuaciones que deberían hacerse y que permitirían obtener resultados mejores a los reportados en este trabajo.

- Encontrar una manera de representar esquemáticamente lugares físicos, en especial, interiores para explorar la idea de que puedan limitarse los lugares por los que se puede transitar.
- Explorar el uso de grafos, para evitar que se puedan atravesar paredes, por mencionar un ejemplo. Lo cual ya ha sido aplicado en trabajo previo mejorándose los resultados obtenidos hasta ese momento (Krumm et al., 2002; Liao et al., 2003).
- Hacer un estudio de las rutas que siguen los usuarios para ver si esta información puede ser usada para predecirlas.

Referencias.

- Bahl, P. y V. N. Padmanabhan (2000). Enhancements to the RADAR User Location and Tracking System. Seattle, WA, USA, Microsoft Research.
- Bahl, P. y V. N. Padmanabhan. 2000. RADAR: An In-Building RF-Based User Location and Tracking System. In the Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000). IEEE Press. 26-30 March, 2000. Tel-Aviv, Israel. p 775-784.
- Battiti, R., A. Villani y T. L. Nhat. 2002. Neural Network Models for Intelligent Networks: Deriving the Location from Signal Patterns. Proceedings of The First Annual Symposium on Autonomous Intelligent Networks and Systems UCLA (AINS 2002). May 8-9, 2002. Los Angeles, CA, USA.
- Bishop, C. M. 1995. Neural Networks for Pattern Recognition. Oxford, Clarendon Press. 482 pp.
- Brown, P. J., J. D. Bovey y X. Chen. 1997. Context-aware applications: from the laboratory to the marketplace. IEEE Personal Communications. **4**(5): 58-64.
- Brumitt, B., B. Meyers, J. Krumm, A. Kern y S. Shafer. 2000. EasyLiving: Technologies for Intelligent Environments. Handheld and Ubiquitous Computing (HUC2K). Springer-Verlag. September 2000. Bristol, UK. p 12.
- Dey, A. K. 2001. Understanding and Using Context. Personal and Ubiquitous Computing. **5**(1): 4-7.
- Dey, A. K. y G. D. Abowd. 1999. Towards a Better Understanding of Context and Context-Awareness. 1st International Symposium on Handheld and Ubiquitous Computing (HUC '99). September 27-29, 1999. Karlsruhe, Germany. p 304-307.
- Dey, A. K. y G. D. Abowd. 2001. A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. Anchor article of a special issue on context-aware computing in the Human-Computer Interaction (HCI) Journal. **16**(2-4): 97-166.
- Dorffner, G. 1996. Neural Networks for Time Series Processing. Neural Network World. **6**(4): 447-468.
- Duda, R. O., P. E. Hart y D. G. Stork. 2001. Pattern Classification. N.Y., John Wiley & Sons. 654 pp.

- Elman, J. L. 1990. Finding Structure In Time. *Cognitive Science*. **14**: 179-211.
- Freeman, J. A. y D. M. Skapura. 1991. *Neural Networks: Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing company, Inc. 401 pp.
- Getting, I. A. 1993. The Global Positioning System. *Spectrum, IEEE*. **30**(12): 36-38, 43-47.
- Harter, A., A. Hopper, P. Steggles, A. Ward y P. Webster. 1999. The anatomy of a context-aware application. 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom 1999). ACM Press. August 1999. Seattle, WA. p 59-68.
- Haykin, S. 1999. *Neural Networks: A Comprehensive Foundation*, Prentice Hall. 842 pp.
- Hightower, J. y G. Borriello. 2001. Location Systems for Ubiquitous Computing. *IEEE Computer*. **34**(8): 57-66.
- Hüsken, M. y P. Stagge. 2003. Recurrent Neural Networks for Time Series Classification. *Neurocomputing*. **50**(C): 223-235.
- Krumm, J., L. Williams y G. Smith. 2002. SmartMoveX on a Graph - An Inexpensive Active Badge Tracker. *Ubiquitous Computing: 4th International Conference (UbiComp 2002)*. Springer-Verlag. September 29 - October 1, 2002. Göteborg, Sweden. p 299-307.
- Liao, L., D. Fox, J. Hightower, H. Kautz y D. Schulz. 2003. Voronoi tracking: Location estimation using sparse and noisy sensor data. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE Press. October 2003. Las Vegas, Nevada. Retrieved from <http://www.cs.washington.edu/homes/kautz/papers/voronoi-tracking.ps> in November, 2004.
- Martinez, E. A., J. Favela y R. Cruz. 2004. Estimating User Location in a WLAN Using Backpropagation Neural Networks. *Advances in Artificial Intelligence – IBERAMIA 2004: 9th Ibero-American Conference on AI*. Springer-Verlag. November 22-25, 2004. Puebla, Mexico. p 737-746.
- Moran, T. y P. Dourish. 2001. Introduction to This Special Issue on Context Aware Computing. *Human Computer Interaction*. **16**: 2-8.
- Orr, R. J. y G. D. Abowd. 2000. The Smart Floor: A Mechanism for Natural User Identification and Tracking. *Conf. Human Factors in Computing Systems (CHI 2000)*. ACM Press. 1-6 april 2000. The Hague, The Netherlands. Retrieved from <http://www.cc.gatech.edu/fce/pubs/floor-short.pdf> in November, 2004.
- Prasithsangaree, P., P. Krishnamurthy y P. K. Chrysanthis. 2002. On Indoor Position Location with Wireless LANs. *13th IEEE International Symposium on Personal, Indoor,*

and Mobile Radio Communications. IEEE Press. 15-18 Sept. 2002. Lisbon, Portugal. p 720- 724.

Priyantha, N. B., A. Chakraborty y H. Balakrishnan. 2000. The Cricket Location-Support System. 6th Ann. Int'l Conf. Mobile Computing and Networking (Mobicom 00). ACM Press. New York. p 32-43.

Raab, F., E. Blood, T. Steiner y H. Jones. 1979. Magnetic position and orientation tracking system. IEEE Transactions on Aerospace and Electronic Systems. **15**(5): 709-717.

Schilit, B. N., N. Adams y R. Want. 1994. Context-aware Computing Applications. Proceedings of the 1st International Workshop on Mobile Computing Systems and Applications. IEEE Press. December, 1994. Los Alamitos, CA. p 85-90.

Skapura, D. M. 1996. Building Neural Networks, ACM Press Books. 286 pp.

Smailagic, A. y D. Kogan. 2002. Location Sensing and Privacy in a Context Aware Computing Environment. IEEE Wireless Communications. **9**(5): 10-17.

Tentori, M. E. 2005. Tesis Cómputo Consciente de la Privacidad en Ambientes Médicos Ubicuos. Depto. de Ciencias de la Computación. Ensenada, México. CICESE. 167 pp.

Want, R., A. Hopper, V. Falcão y J. Gibbons. 1992. The active badge location system. ACM Trans. On Information Systems. **10**(1): 91-102.

Apéndice A.

Sistema para mediciones.

Se desarrolló un sistema para PDA para hacer las mediciones de la fuerza de la señal de una manera más rápida y efectiva. A continuación se describe el sistema mencionado.

El sistema fue desarrollado en VB.Net de Microsoft, además se usa un *framework* llamado *Smart Device Framework*², el cual es un marco de clases desarrollado por un grupo de programadores bajo el movimiento de código abierto que extiende el actual Compact Framework de Microsoft.

Descripción del sistema

El sistema está compuesto por una interfaz sencilla. La ventana principal muestra el plano arquitectónico del lugar en el cual se hacen las mediciones. La cual se muestra en la figura 37.

En la parte inferior del sistema está el menú y una pequeña barra de herramientas. El menú consta de una opción que es para cambiar los parámetros de configuración del sistema, la cual se describe enseguida:

La pantalla de Configuración consta de varios campos donde se tiene que indicar los siguientes datos:

1. El nombre del archivo en el que se desea guardar las mediciones tomadas. El nombre por defecto es *medicion.txt*.
2. El nombre de la imagen del plano arquitectónico donde se tomarán las mediciones. El nombre por defecto es *imss.png*.

² <http://www.opennetcf.org>, septiembre del 2004

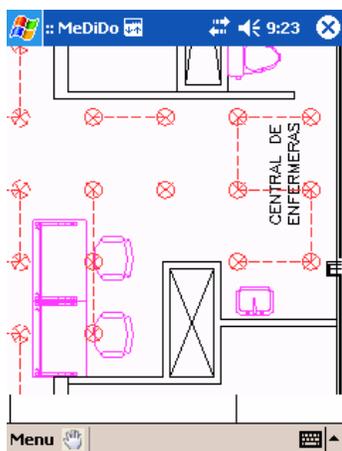


Figura 37. Pantalla principal del sistema

- Los dos últimos campos guardan las medidas reales del edificio al que pertenece el plano arquitectónico. Por defecto trae los valores de 43.2 m para el campo Largo y 33.45 m para el campo Ancho. Estos datos sirven para hacer la conversión de pixeles a metros.

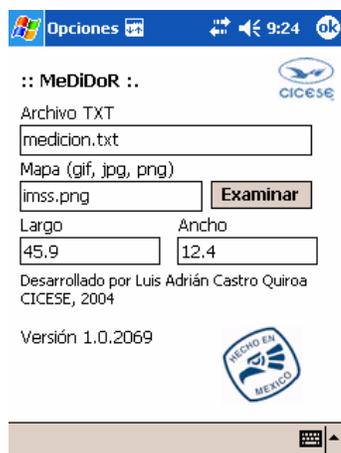


Figura 38. Pantalla de configuración.

La barra de herramientas consta de un solo botón [] permite hacer el arrastre de la imagen para tomar mediciones en otros puntos que no sean los que aparecen en pantalla. Cuando la toma de mediciones ha iniciado aparece una barra para indicar el progreso de la toma de 100 mediciones por punto. En la figura 3 se puede ver dicha barra.



Figura 39. Barra de progreso

Inicio de la aplicación.

Al iniciar, el sistema verifica que se cumplan algunos requisitos para poder hacer las mediciones. Lo primero que verifica es que esté activada la tarjeta inalámbrica. En caso de que no lo esté, el mensaje se muestra en la figura 4.

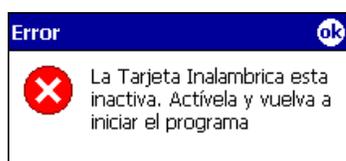


Figura 40. Mensaje mostrado cuando la tarjeta inalámbrica esta inactiva

Cuando la tarjeta está activa se inicia de nuevo la aplicación y ésta verifica que no exista un archivo con el nombre guardado en los datos de configuración. En caso de que exista, la aplicación pregunta si se desea conservarlo o si desea sobrescribirlo. La figura 5 muestra el mensaje mencionado.

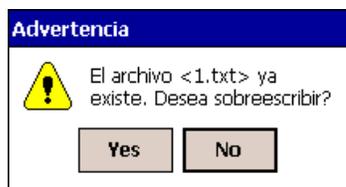


Figura 41. Mensaje para indicar que ya existe el archivo.

Toma de mediciones.

Para tomar las mediciones es necesario arrastrar la imagen y establecerla donde este posicionada físicamente la persona que toma las mediciones. Cuando presione el sitio donde se encuentra aparecerá un mensaje de confirmación para verificar si es en realidad el sitio donde se desea hacer las mediciones. El mensaje es como el que aparece en la figura 42.

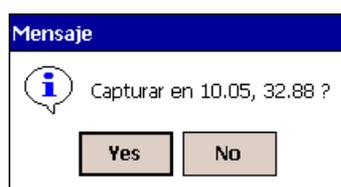


Figura 42. Mensaje de confirmación para iniciar con mediciones

Una vez que se inician las mediciones se deshabilita el botón de arrastre de la imagen en la barra de herramientas para no interrumpir las mediciones. Al finalizar con las mismas, aparece un nuevo mensaje de confirmación para guardar o no las mediciones a disco. El mensaje se muestra en la figura 7.

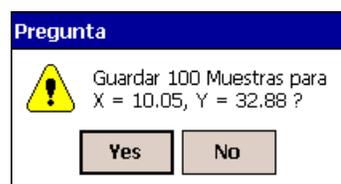


Figura 43. Mensaje de confirmación para guardar mediciones

Apéndice B.

Sistema para estimación de la localización.

Para probar la estimación de la localización se diseñó una aplicación donde se pudiera ejecutar la red neuronal con los pesos sinápticos resultantes del entrenamiento. En la figura 44 se muestra una captura de pantalla de la aplicación.

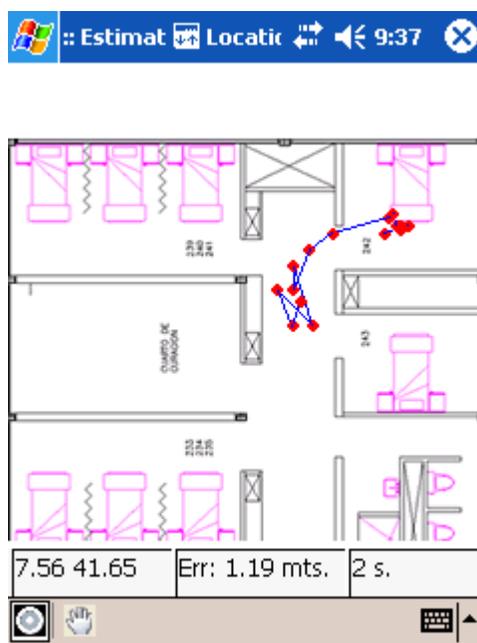


Figura 44. Captura de pantalla de la aplicación para la estimación de la localización.

En la parte inferior, en la barra de estado, se muestran, de izquierda a derecha, las coordenadas de la última estimación, enseguida el error, esto porque esta aplicación se estuvo tomando en cuenta las mediciones que se tomaron en las rutas que se muestran en el apéndice D. El apartado de la derecha indica el tiempo que tomó en hacer la estimación.

