uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Hong Li
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Electrical Engineering)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Analysis and Design of Neural Network-Based WLAN Indoor Location System

TITRE DE LA THÈSE / TITLE OF THESIS

Tet Hin Yeap
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

V. Groza

Q-J. Zhang

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Analysis and Design of Neural Network-Based WLAN Indoor Location System

Hong Li

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the degree of

Masters of Applied Science

in Electrical and Computer Engineering

Ottawa-Carleton Institute of Electrical and Computer Engineering

School of Information Technology and Engineering

Faculty of Engineering

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-34083-7
Our file Notre référence
ISBN: 978-0-494-34083-7

Analysis and Design of Neural Network Based WLAN Indoor Location System

By

Hong Li

Abstract

As IEEE 802.11 WLAN networks are widely deployed, WLAN-based indoor location systems have obtained a good chance to develop, because they can be easily achieved on the existing WLAN infrastructure and by the RSSI parameter provided by the standard. Due to the complexity of the indoor RF propagation environment, the traditional triangulation method becomes impractical for indoor location systems and the propagation pattern matching method becomes the major method instead. RSSI values from a set of APs form the unique signatures or the propagation map for locations. The neural network method is one of the major location estimation methods for propagation map matching. Compared with the probabilistic method and the nearest neighbor method, the NN method has advantages in saving computing resources. However, the performance of a NN-based WLAN-based location system is determined by its training process and the resolution of the propagation map. By analyzing the properties of the process of RSSI creation, the underlying relationship between the WLAN RSSI parameter and space or RSSI resolution over space is obtained by applying the log-distance path loss model. Propagation map creation rules or rules for creation of neural network training data set are proposed based on the analysis and simulation results. And finally, guidelines about the design and implementation of an NN-based WLAN location system are proposed.

Acknowledgements

Thanks to Tet Yeap for being my supervisor; giving me this opportunity; and greatly helping me to do my research.

Thanks to all the members of the BARLO Lab. Every discussion with them inspired my thoughts in many ways.

Thanks to many friends in Ottawa who gave me lots of inspiration and brotherly care in my life.

Thanks to Mr. Maoxin Wei who has been a mentor during my days in North America. His advice and guidance is necessary for me to overcome my hard days.

Thanks to my family, especially Ping Gong and Eric Li, who encouraged me to go forward.

Abbreviations and acronyms

AP – Access Point

BSS – Basic Service Set

DSSS – Direct Sequence Spread Spectrum

ETSI – European Telecommunications Standards Institute

FCC – Federal Communications Commission

FHSS – Frequency Hopping Spread Spectrum

GPS – Global Positioning System

HEC – header error check

HIPERLAN – High Performance European Radio LAN

IBSS – Independent Basic Service Set

IEEE – Institute of Electrical and Electronics Engineers

ISM – Industrial, Scientific and Medical bands

LAN – Local Area Network

MAC – Media Access Control

MLP – Multi-layer Perceptron

NIC – Network Interface Card

NN – Neural Networks

PHY – Physical Layer

PLCP – Physical Layer Convergence Protocol

PPDU – PLCP protocol data unit

RSSI – Received Signal Strength Indicator

SFD – Start Frame Delimiter

WAN – Wide area network

WI-FI – Wireless Fidelity, used generically when referring of any type of 802.11 networks, including 802.11b, 802.11a, 802.11g.

WLAN – Wireless Local Area Networks

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Introduction

Location technology has many names, such as location estimation, geolocation, location identification, location determination, localization, or positioning. Most location systems estimate the mobile device's location by using multiple RF or other types of signals transmitted or received by the mobile terminals. The most successful location system currently is the GPS (Global Positioning System), which makes use of RF signals from GPS satellites to determine the location of a GPS receiver. The success of GPS inspired the emergence of many location systems. Some of them are global navigation satellite systems (GNSS), and they are competitors of the GPS system, such as the Russian GLONASS and the European Union's embryonic Galileo positioning system. And some location systems are designed as complementary to GPS system, such as QUALCOMM®'s GpsOne® or A-GPS technology or mobile communication systems. GpsOne uses a combination of GPS satellite signals and signals from cell towers; it allows the user's location to be plotted with greater accuracy than traditional GPS systems in areas where satellite reception is problematic due to buildings or terrain.

Besides a positioning service, location systems can also bring lots of other applications. For the mobile communication industry, which is in the age of significant competition, mobile network operators continuously seek new and innovative ways to create differentiation and increase profits. One of the best ways to accomplish this is through the delivery of highly personalized services. One of the most powerful ways to personalize mobile services is based on location. Location Based Service (LBS) is considered potential profit source in the near future. LBS makes use of Location technology, Geographic Information Systems (GIS), and the Location Management Function to provide users with services, such as location-based information provision, location based billing, emergency services, tracking, etc. In the U.S., the FCC requires that all carriers meet certain criteria for supporting LBS (FCC 94-102). The mandate requires that 95% of handsets resolve within 300 meters for network-based tracking and 150 meters for handset-based tracking. This can be especially useful when dialing an emergency telephone number, such as enhanced 9-1-1 in North America, so that the operator can dispatch emergency services such as Emergency Medical Services, police or firefighters to the correct location.

In recent years, IEEE 802.11 Wireless Local Area Networks (WLANs) have become very popular in office and residential environments. WLAN's advantages in mobility, flexible coverage area, high bandwidth, and easy-setup introduce many new application domains. And since WLANs are usually deployed in indoor environments, they provide ideal platforms for indoor location systems. One important application of WLAN location technology is context-aware computing or ubiquitous computing. Context-aware systems refer to systems that can be aware of their physical and virtual environment or situation and respond intelligently based on such awareness. Ubiquitous computing integrates computation into the environment, rather than having computers which are distinct objects. Location information can provide additional context for location aware mobile stations, and therefore the development of WLAN location technologies is very important for context-aware computing and ubiquitous computing. Besides of that, WLAN locations will also benefit industries and emergency services that require accurate location and tracking of objects and people in indoor environments. Examples of applications include tracking of assets and products within a factory, location of staff and patients in hospitals, and tracking emergency personnel in a burning building.

Compared with an outdoor propagation environment, indoor environments are more complex. Due to severe multipath propagation phenomena inside buildings or complicated urban environments, the relationship between the distance and signal transit time, arrival angle, and signal strength becomes much more complicated than outdoor environment. Geometry methods or triangulation methods cannot achieve satisfied performance under these situations and become impractical. Even though GPS system is very successful in outdoor environments, its application in indoor environments and high dense urban areas is very limited due to its weak signal reception when there are no lines-of-sight from the mobile device to at least three GPS satellites. Thus, signal propagation pattern-based methods, also called fingerprinting methods or location pattern matching methods becomes major methods for indoor location estimation algorithms. Propagation pattern based method is based on the theory that location information can be extracted or estimated by comparing the detected signal characteristic with pre-stored signal characteristics. In a WLAN environment, this technique requires a propagation map which is formed by vectors of received signal strength from a set of access points (APs) for specific locations, and each RSS vector is unique to one specific location. When a mobile device needs to estimate its location, it measures the RSSI from different APs at its current location and searches for the signal pattern with the closest match in the propagation map to extract the corresponding location [15].

Using exiting WLAN and the propagation pattern matching method to achieve an indoor location system has gained lots of researchers' interest. Different location estimation methods for extracting locations from a propagation method have been proposed, such as the Nearest Neighbor method of RADAR system [22], the probabilistic method [16] [21] [23]

[30], and the neural network method [24] [32]. Some researchers [3] [11] [15] [62] [63] compare the above methods by setting up and doing experiments in their own experimental test beds. The test results are variant but similar. All the methods can achieve similar performance - usually the highest accuracy is around 1.5 meters with 25% test data. In all of the above papers, the same research approach is used. The performance of the system is estimated by analyzing the tremendous data collected in the test bed. However, the underlying relationship between system performance and the propagation properties of WLAN networks is not analyzed. Thus the variation of performance with different factors can only be statically estimated. And therefore, the test results and conclusions are usually limited to the specific test environment. Thus, to optimize system performance, it is necessary to understand the relationship between system performance and the WLAN propagation environment.

1.2 Contribution

This thesis researches the location problems on the neural network-based WLAN location system. WLAN-based location systems make use of the RSSI information provided by the standard WLAN NIC card to estimate a mobile device's location. While empirical results and performance studies of different location estimation methods in a WLAN environment have been presented in literature, all of the existing research focuses on data analysis of collected signal information but not on the underlying relationship between WLAN RSSI distribution and system performance. The major contributions of this thesis are:

- The WLAN indoor propagation environment is analyzed. Factors that affect RSSI distribution are analyzed, and measures to make use large-scale fading and eliminate small-scale fading are proposed.
- A concept "RSSI resolution" over space for the IEEE 802.11 WLAN environment is defined and analyzed. The relationship between RSSI distributions over space with the distance between the receiver and AP, the path loss exponent, and the quantification properties of RSSI collection circuit are analyzed. A set of principles of RSSI resolution in a WLAN environment are proposed.
- Given the above principles and the simulation and analysis results, basic rules or guidelines of creating a propagation map in a WLAN environment are proposed.
- A system structure of a propagation map-based WLAN location system using a neural network is proposed.

1.3 Organization

Chapter 2 reviews the indoor location system and technologies and provides the background for this research. Chapter 3 introduces WLAN networks and artificial neural networks. Besides a brief introduction of WLANs, concepts of RSSI, including RSSI in IEEE 802.11 standards, the creation of RSSI, and factors that affect RSSI are introduced. MLP as a major neural network in our application is introduced in detail in Chapter 3. In Chapter 4, WLAN indoor propagation environment is analyzed, and models of large-scale fading and small-scale fading are discussed and verified. Then, based on the large-scale fading models and the properties of RSSI granularity in WLAN, the factors that affect RSSI resolution over space are analyzed, and finally, propagation map creation rules are proposed. The results in Chapter 4 are applied in Chapter 5, and a NN-based WLAN location system structure is proposed. And finally, conclusions are presented in Chapter 6.

2. Literature Review

This chapter reviews the literature and background of a current indoor location system and gives an introduction about existing indoor location technologies. Due to the multipath radio propagation characteristics, indoor location research remains a very challenging task for researchers in this domain. If a radio signal's characteristic, such as signal strength or transit time, has a straightforward relationship with the distance between the transmitter and the receiver, then the location of mobile device can be easily derived from solving quadratic equations. However, in an indoor propagation environment, such kinds of relationships usually do no exist. Successful outdoor geometry location methods such as TDOA, TOA, or AOA cannot gain satisfactory performance in indoor environment. The Signal propagation pattern-based location method is more applicable in this case. A signal propagation pattern is based on the assumption that at each location a related unique signal signature exits, such as a vector of RF received signal strength from different base stations or access points.

In the following sections, a detailed introduction about research in location systems, especially in indoor location systems, is given. Section 2.1 discusses the issues concerning an indoor environment. The following two sections discuss the different physical media and infrastructures a location can use. And Section 4 discusses the location method, including the geometry method and the propagation pattern-based method. Section 5 is focused on presenting the location calculation method, and Section 6 describes existing major indoor location systems.

2.1 Issues Concerning Indoor Location System

GPS has gained remarkable success in outdoor positioning applications. However, the GPS system cannot be used effectively inside buildings and in dense urban areas because the satellite signals can be effectively received. Consequently, infrared, RF, and ultrasound become alternative technologies for indoor location systems. Since indoor environment has dense multipath effects for RF signals and many building material-dependent propagation effects for ultrasound and infrared signals, people usually need to spend more effort but get less reliable performance when they deploy an indoor location system. On the other hand, due to the increasing and numerous deployments of WLANs by many individuals and organizations, WLAN wide coverage provides a great opportunity for location-based service. Location-based service can be achieved on WLAN infrastructure without the need of any additional hardware.

However, indoor location technologies are so new that there are no theoretical and analytical backgrounds. In consequence, it is difficult to effectively evaluate an indoor location system. But there are still many issues to be considered when a location system is designed. And in some measures, how these issues are addressed is one way to evaluate how good an indoor location is. Krishnamurthy [1] pointed out four major issues regarding position location in a mobile environment which are performance, cost and complexity, application requirements, security and privacy. Performance metrics include accuracy, precision, delay, coverage, scalability, capacity, etc. Accuracy is the likelihood that the object is within the estimated area, and precision is the size of the adjacent area that will return the same output. All of these performance metrics depend on the choice of the lower physical media, the system infrastructure, the location method, the characteristics of the application environment, and the performance constraints of its hardware system. Sometimes the relationship between performance and cost and complexity is a tradeoff. Cost and system complexity could increase because of extra infrastructure, additional bandwidth, reliability requirements, installment, and maintenance. The major application requirements for the location information are granularity, performance, and availability. Different applications may have different requirements. Security and privacy refers to location information should be available only to those who have authorization. Different techniques have different strengths in addressing this issue. If the user terminal can derives its own location, just like the GPS receiver in the GPS system, user's privacy can be completely secure. On the other hand, when using location technology such as E-911, which relies on the communication network to cooperate to estimate the location of a mobile device, extra measures must be taken to protect the location information.

2.2 Physical Media

Various technologies can be used to find out the position of a mobile device. The location system can be categorized based on the low-level physical media it use, its location method, and the system infrastructure. The physical media refers to the type of signals used by the physical layer of a location system. And location method refers to the methods and metrics used in higher-level algorithms. System infrastructure refers to the relation and function distribution among each node of the system.

To find out a user's location, most location systems involve multiple nodes. These location systems have to use some kind of wireless signal to correspond between nodes. Based on the physical media they used, location systems inherit certain constraints and features of that type of physical media. The constraints and features mainly refer to propagation properties, such as propagation speed, diffraction, reflection, scattering, effective range, and interference [2]. But sometimes factors such as available bandwidth, regulatory constraints, power constraints, safety, and cost have to be considered when designing a location system. Three types of wireless signals are commonly used in location systems, and they are ultrasound, infrared and RF signals. The following table (Table 2.1) is a brief comparison of these signals [2]:

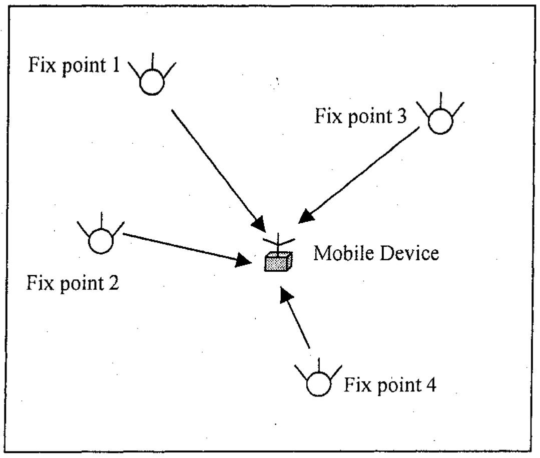
Table 2.1 Comparison of Different Physical Media for a Location System

Signal Type	Ultrasound	Infrared	RF
Propagation Speed	343m/s	$3x10^8$m/s	$3x10^8$m/s
Effective Range	Small (3-10m)	Small (5 -10m)	Large (100m+)
Resolution	Very High (1cm)	High (10-20cm)	Variance (10m – 200m)
Effect of Obstacle	Blocked and reflected by most obstructions	Blocked by most obstacles	Penetrate but attenuated by obstacles
Interferences	Affected by temperature	Other lighting could interfere	Interfere by same band RF signals

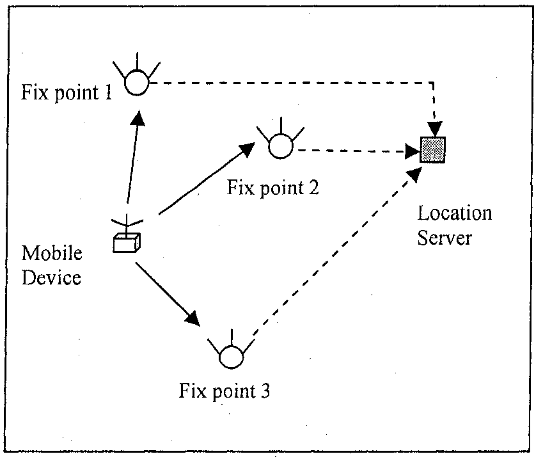
Compared with ultrasound and infrared signals, RF signals have a strong advantage in location systems. RF signals have a fast propagation speed and can penetrate most obstacles, so they can be applied to both indoor and outdoor environments. Ultrasound and infrared signals cannot penetrate walls or most obstacles indoors, so the location systems using ultrasound and infrared techniques have a very limited effective range. Besides that, most location systems can make use of the existing communication infrastructure, especially in indoor environments where WLANs have been deployed. An indoor location system based on the WLAN system does not need additional hardware to be installed. The WLAN adapt card itself is capable of sensing RF signal strength.

2.3 System Infrastructure

Based on the functions each node can perform, location systems have two types of infrastructures: a mobile-based infrastructure and a network-based infrastructure [3]. The following figure shows the basic components of these two types of infrastructures. In the mobile-based infrastructure, the mobile device itself can determine its own location by using radio sources from base stations or access points. In other words, the mobile device actively performs calculations or location functions in this mode. In the network-based infrastructure, the mobile device's location is determined externally from its radio signal by measuring the signal's characteristics at three or more fixed points (base stations or access points). A mobile station can passively acquire its location from an outside source, and the location function is run on a location server or central station, which can acquire the signal characteristics at each fixed point through a network.



(a). mobile-based Infrastructure



(b). network-based Infrastructure

Figure 2.1 Location System Infrastructures

GPS uses the mobile-based infrastructure. Fixed points 1 to 4 in Figure 3.1 (a) represent GPS satellites, and the mobile station represents GPS receivers. GPS is a satellite-based navigation and location system originally developed for military purposes. The system consists of three segments: space segment, control segment, and user segment. The current space segment includes 28 Block II/IIA/IIR satellites [4] occupying six orbital planes inclined at a 55-degree angle with respect to the equator. The GPS satellite broadcast PRN code modulated navigation signals. The control segment includes one master control station, five monitor stations and four antenna upload stations. The main task of the control segment is to maintain the correct operation of satellites. The user segment is basically different types of GPS receivers, which use different line-of-sight GPS satellite signals to determine the user's position, velocity, and time.

Location services in mobile communication systems usually use the network-based infrastructure. In a typical mobile communications system, wide geographical areas are mapped into cells centered by a base station. The base station has base transceiver system units that provide the signal processing capability. Base station controllers in turn manage these units. Groups of base station controllers are managed by the mobile switching centre that functions as the gateway to the traditional public telephone network. Location services in mobile communication systems such as E-9-1-1 are usually based on time of arrival (TOA), time difference of arrival (TDOA), or angle of arrival (AOA) technologies, which require special hardware to be installed in the base stations to measure the signal arrival time or arrival angle from the mobiles. Measurements from different base stations are fed into the location server, where the signal measurements are processed and the user's location is estimated.

A mobile-based infrastructure requires an intelligent mobile device, which has the ability to take the measurement and run a location algorithm. Since the location function is running on the mobile device and there is no extra burden on the outside network, it has a bigger system capacity and a lower processing delay. There is no communication requirement between the fixed points and the mobile device. So the bandwidth requirements and cost for fixed points are lower. Besides that, mobile users can hold their location information, and there is no problem with privacy. However, the location performance is limited by the ability of the mobile device. On the other hand, a network-based approach can use a powerful location centre to provide better performance for a low configured mobile device. Since the measurements are taken at fixed points, the mobile device does not need to be equipped with special antenna or measuring hardware. An Indoor location system based on a WLAN infrastructure usually has powerful mobile devices, such as laptop computers or PDAs. Mobile-based infrastructure is easier to deploy in these systems since it does not need to change the WLAN infrastructure.

2.4 Location Method

Many different technologies and systems to determine the location-related information of mobile devices have been proposed. To determine the position of a user, there are two basic methods:

1. The geometry method or triangulation method, which requires at least three distinct known positions to provide reference distances or two distinct known positions to provide a direction or angle of arrival. (Savarese [5] , Savvides, Park and Srivastava [6], Ji and Zha [7], Hu Evans [8], Niculescu and Nath [9])

2. The signal propagation pattern-based method, which estimates the location based on the unique signal characteristics of locations. This method is also called location fingerprinting. (Hills, Schlegel, Jenkins [10], Krishnan [11])

2.4.1 Geometry Method

The geometry method is the primary location method used in outdoor positioning systems. Given two of three known positions and the distance or angle from the known location to the mobile device, the location of the user's device can be calculated by a triangulation process. Several different location technologies use the geometry method to estimate the location of a mobile device, including GPS technologies, TOA, TDOA, Enhanced Observed Time Difference (E-OTD), AOA technologies, and some Received Signal Strength (RSS)-based location technology.

One good example of using the geometry method is GPS. GPS is currently the most important and successful location system, which uses the distance-based triangulation method to estimate a user's location. The GPS receiver measures the apparent transit time of the satellite signal from the satellite to the user to estimate the distance between them. By using at least four such measurements and knowing the satellite position from the ephemeris data, a receiver can determine its three-dimensional position and its clock bias [12]. GPS can achieve approximately 10 meters accuracy. The following figures illustrate the principle of estimating positions in GPS system. We can see that receiver R is located at the cross point of the three spheres, which are centered by each satellite and have a radius of d_1, d_2 , and d_3 respectively. By measuring the distance of d_1, d_2, and d_3 to three different satellites (s_1, s_2 , and s_3), the GPS receiver can estimate its location if it knows the position of these three satellites. Even the basic principle is clear and simple, but a receiver has to take into account many factors to get an accurate estimation of the position. The distance from the receiver to satellite is calculated as the product of radio speed and the radio travel time between them. So, besides the three satellites above, it needs at least another satellite to estimate the variance in time.

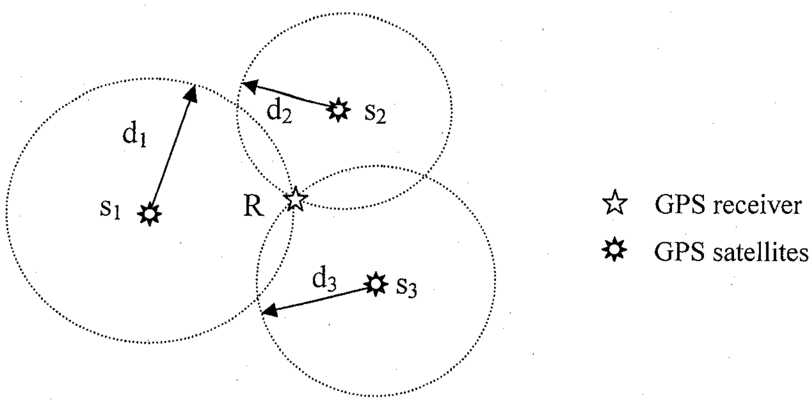


Figure 2.2 Illustration of the Distance-Based Triangulated Method of GPS

TOA, TDOA, E-OTD, and RSS based geometry location technologies have the same working principle as GPS [13]. What differs among them is what they measure. TOA measures the arrival time of a known signal sent from the mobile device to different known location receivers. TDOA measures the difference of arrival time among different receivers of the same signal that is sent out simultaneously from the transmitter. In E-OTD, the mobile mode does the time measurement. And AOA uses a special antenna array at the known locations to measure the angle of arriving signals from the mobile device. RSS-based technology measures the strength of the received signal in order to deduce the possible range the signal has propagated from the transmitter to the receiver. Currently, the reported accuracy of TOA, TDOA, E-OTD, and AOA technologies is lower than 100m. However, to achieve this accuracy, they all need special hardware to be installed on the receiver side, which will increase system cost and complexity. RSS-based technology is applicable only if the transmission power is constant or known in advance. And it is a low accuracy localization method. However, it does not need additional hardware and this makes it easy to be supported by current wireless devices.

Even though the geometry method has been successfully applied in GPS and mobile communication systems, its application in indoor environments and high dense urban area is limited due to its weak signal reception when there are no lines-of-sight from the mobile device to at least three GPS satellites [14]. Due to severe multipath propagation phenomena inside buildings or complicated urban environments, the relationship between distance and signal transit time, arrival angle, and signal strength becomes much more complicated than in outdoor environment. Geometry method cannot achieve satisfactory performance under these situations and becomes impractical.

2.4.2 Signal Propagation Pattern-Based Method

Signal propagation pattern-based methods, also called fingerprinting methods or location pattern matching methods, are based on the theory that location information can be extracted or estimated by comparing the detected signal characteristics with pre-stored signal characteristics. This technique generally requires only measurements of received signal strength or other non-geometric features of different locations to create a location base. When a mobile device needs to estimate its location, it measures the signal characteristic information at its current location and searches for the signal pattern with the closest match in the location base. The assumption to create a location base is that each position or location within the effective range has a unique signal signature [15].

The process of using the signal propagation pattern-based method can be divided into two stages. The first stage is the training step. It is the process of building the propagation map. To build the propagation map, large amounts of measurements of signal signatures at multiple locations must be taken manually or automatically. Since there is always some noise or dynamic variance in the collected data, the data has to be calibrated or preprocessed before it goes into the propagation map. After the measurement and preprocessing of the signals are finished at all the interesting locations, the propagation map is built up and the training stage is over. The second stage is the application stage. It is the process of estimating the mobile device's location. Several measurements of the signal propagation are taken at the target location (mobile-based) or fixed positions (network-based) in a network, and these measurements are processed to eliminate the noise and then feed into a location estimation method to extract the estimated location from the propagation map.

As we mentioned above, the RF signal has remarkable advantages in indoor location applications that ultrasound and infrared do not have. And RSS is the simplest and most effective signal signature for the signal patterns method because it is available from any normal WLAN adapt card. Bahl and Padmanabhan [22] found that RSS is more location-dependent than the signal-to-noise (SNR) because the noise component is rather random in nature. However, RSS itself is usually wavy, even at the exact same location, because of the dynamics of the environment. Several approaches can deal with this problem. We will discuss this in the next section.

2.5 Location Calculation Methodology

There are several ways to find out the dependency between measurements and the propagation map and to determine a location for a set of given measurements. It is very clear that the performance of the whole location system depends on the performance of the location calculation method it uses. Location calculation methods can be classified into deterministic and probabilistic. The major deterministic methods include the nearest neighbor method and the neural network method. The location calculation methods can be considered as pattern classifiers as they are just procedures that separate pattern samples into different classes [16]. In this case, each class represents a set of signal signature patterns from one specific location. In the following part, we will discuss the three major location calculation methods: the nearest neighbor, neural network, and probabilistic methods.

2.5.1 Nearest Neighbor Method

The nearest neighbor method was initially introduced by J. G. Skellam, where the ratio of expected and observed mean values of the nearest neighbor distance is used to determine whether a data set is clustered [17]. It is based on comparing the distribution of distances from a data point to its nearest neighbor in a given data set with the randomly distributed data set. In our case, nearest neighbor methods are deterministic because they require only a set of constant propagation pattern vectors. To determine a location, a distance function is used to classify pattern samples into positions. The following is a mathematical description of the nearest neighbor method.

Suppose the propagation map includes a set of m location pattern vectors, denoted by $P = \{P_1, P_2, \ldots, P_m\}$, in which each pattern vector P_i consists of n signal signatures at location

i, (i =1, 2,..., m). n is the number of fix points or access points in the WLAN environment. If the signal signature of access point j at location i is p_{ij}, the pattern vector P_i can be expressed as $P_i = \{p_{i1}, p_{i2}, ..., p_{in}\}$. When a measurement is taken at the application stage, the signal sample vector includes n signal characteristic samples, which can be expressed as $S = \{s_1, s_2, ..., s_n\}$. To search for the nearest neighbor, a distance function $Dist(S, P_i)$ is used, and the nearest neighbor method is to find out the location x that satisfied the following formula. [18]

$$Dist(S, P_x) \leq Dist(S, P_y), \forall x \neq y \tag{2.1}$$

or

$$D = Dist(S, P_x) = \min_{y=1}^{n}\{\frac{1}{n}(\sum_{i=1}^{n}\frac{1}{w_i}\left|s_i - p_{yi}\right|^a)^{1/a} \tag{2.2}$$

Where w_i is a weighting factor, and a is the norm parameter starting from 1. The weighting factor w_i is a bias parameter that denotes or promotes an important signal signature in the pattern map [19]. If all w_i=1 and a =2, then the distance becomes a Euclidean distance, and D is the least mean square error (LMSE).

A problem with the nearest neighbor method is that two or more very different locations could potentially have the same signal signature. And in practice, the actual location is not limited to the locations within a radio map, and there can be more than one closest match. k Nearest Neighbors method is used instead of the nearest neighbor to solve these problems. The final estimated position is an averaging of that k nearest neighbor's coordinates. Bahl and Padmanabhan [22] reported that for small k there is little improvement over the single nearest neighbor method, while for large k the location estimation error performance increased. The nearest neighbor methods are fast to deploy and do not need training or adjusting. However, the complexity increases as the number of elements in the propagation pattern and the number of pattern vectors increase.

2.5.2 Probabilistic Method

The probabilistic method is based on probabilistic models that describe the dependency of observed signal properties on the location of a mobile device. By modeling the propagation map with conditional probability and utilizing the Bayesian inference concept to estimate location, the probabilistic method provides a way to handle uncertainty and errors in signal measurements[20][21][23]. The basic principle of the probabilistic method is discussed below.

The probabilistic approach presumes a priori knowledge of the probability distribution of the user's location. Suppose the user's location can be expressed as $L= \{l_1, l_2, \ldots, l_m\}$, where m is the number of locations. This approach uses models that estimate the probability distribution of the observation variables or measured signal variable s given the value of the location variable l. In other words, for any given location l, a probability distribution of $p(s \mid l)$ that assigns a probability for each measured signal vector s can be obtained. By applying the Bayes rule, a posterior distribution of the location is given:

$$p(l \mid s) = \frac{p\{s \mid l\}p(l)}{p(s)} = \frac{p(s \mid l)p(l)}{\sum_{l' \in L} p(s \mid l'')} \qquad (2.3)$$

$p(\,l\,)$ is the prior probability of being at location l before knowing the value of the measured signal variable, and the summation goes over the rest of possible location values L. The posterior distribution p(l | s) can be used to choose an optimal estimator of the location based on a loss function.

The term p(s | l) is called the likelihood function. Roos et al [23] suggest two methods for estimating the likelihood function: the kernel method and the histogram method. In the kernel method, a probability mass is assigned to a "kernel" around each sample vector in the training set. Thus the resulting density estimate for a signal set s in location l is a mixture of n_l equally weighted density function, where n_l is the number of training vectors in l. Suppose the kernel function is Gaussian. Then the likelihood function can be written as:

$$p(s \mid l) = \frac{1}{n_l} \sum_{i=1}^{n_l} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(s-\rho)^2}{2\sigma^2}\right) \right] \tag{2.4}$$

where σ is an adjustable parameter that determines the width of the kernel. To extend the kernel method to multiple dimensions or access points, Roos et al make an independent assumption and multiply all conditional probabilities together. The histogram method estimates the continuous density functions using the discredited density functions. This histogram is essentially a fixed set of bins that counts the frequency of occurrence of signal samples that fall within a range of each bins. The number and range of bins are adjustable. The larger the number of bins, the better the histogram can approximate the probability.

Probabilistic methods can provide better performance on location estimation than the nearest neighbor method, because they have additional information on the location distribution. The prior distribution p(l) gives a principled way to incorporate background information such as user profiles. In the case where there is no user profile, the priori probability can simply use uniform distribution. Probabilistic methods usually need a large training set to precisely estimate the conditional probability distribution.

2.5.3 Neural Network Method

When the indoor environment has a complex structure, multiform decorations, and lots of furniture and is crowded with people, the multipath effect will be very strong, and it will be impossible or impractical to mathematically analyze the signal propagation pattern. Instead of using methods that are based on theoretical analysis, neural network methods can work as a black box information processing unit. In the next chapter, we will give a detailed introduction about neural networks. Battiti et al [32] introduced a multilayer perceptron (MLP) based neural network method to solve location estimation problems. An MLP has one hidden layer that uses the sigmoid function and an output layer that uses an identity function. The input of the neural network is the RSS of each access point, and the output of the neural network is the location. Battiti et al [24] finally gives an MLP structure that has three inputs, eight hidden units, and two outputs of 2D coordinates for 194 measurement points.

An MLP provides nonlinear mapping between its input and output. It has been proven that, with a sufficient number of hidden nodes, an MLP with as few as two hidden layer neurons is capable of approximating arbitrarily complex mapping within finite support [33]. [34] and [35] also show that an MLP with sufficiently many nonlinear units in a single hidden layer have been established as universal function approximators.

As a flexible model, the MLP neural network does not rely on the priori knowledge of any environment parameters such as the location of access points and building characteristics [32]. The accuracy and precision performance of [32] are slightly better than the nearest neighbor method.

The neural network-based method has a strong advantage over other methods. First, theoretically, a neural network such as an MLP has the capability of mapping any kind of nonlinear function. Usually, an indoor environment is very complicated for RF signal propagations; it is impractical to accurately predict signal properties at a specific location, given the most accurate propagation models nowadays. But the complexity of a propagation environment will not affect the performance of a well-trained neural network.

Second, a neural network requires fewer resources to store. Unlike other methods, which require that all propagation patterns over the target area must be stored, (it includes at least hundreds of vectors of signal signatures), the neural network-based system compresses the propagation patterns into its parameters. The only thing that needs to be saved is the weights and thresholds of neurons. For example, if there are 500 sampling points at the target area and there are 4 APs; to save the propagation pattern, it requires at least 500 samples x (4 AP + 2 location coordinates) = 3000 data. But a MLP with 4 input nodes, 10 hidden nodes, and 2 output nodes, only needs to save 72 pieces of data (60 weights and 12 thresholds). As the location system expands its coverage area and including more APs, the storage size of a propagation pattern will increase dramatically, but there is almost no need to increase the storage space for MLP neural networks.

Third, because the size of a neural network is small, the computing resources for running the neural network-based method are fewer than other methods, especially when the size of the propagation map is big. For other methods, to calculate a location, the location calculation algorithm will traverse the whole propagation map, which will be a huge data set if the target area is big. On the other hand, the neural network method only needs to do much less calculation, and it will not change with the size of propagation map.

However, even though neural network can work without prior knowledge of WLAN deployment, the performance of a neural network really depends on the selection of the training data set, the deployment of the access points, and the indoor propagation environment. Unfortunately, Battiti et al did not analyze the signal propagation environment and its effect on the performance of the neural network-based indoor location system. In this thesis, we will give a detailed analysis of indoor propagation mechanisms and propagation models that describe the signal propagation patterns. Based on the analyses, procedure and measures about creating a training data set and designing system are suggested to improve the performance of the neural network-based indoor location system. Detailed discussion please refers to Chapter 4 and Chapter 5.

2.6 Related Indoor Position System

Many indoor location systems using different technologies have been proposed, such as Active Badge [25], which uses infrared signals; Active Bat [26, 27], Cricket [28], which use a combination of RF and ultrasound signals to estimate distances; PinPoint 3D-iD [29], RADAR [22], Nibble[30] and SpotOn[31], which use RF-based signal characteristics to estimate locations. The major characteristics of these systems are discussed:

Active Badge [25] is one of the earliest indoor location systems. It uses a network-based infrastructure and it is based on diffused infrared technology. Through infrared sensors that are placed in the building, a location server collects IR signals sent out by a badge and uses these signals to determine the badge's location. Although Active Badge can provide accurate location information, its effect range is limited, and its performance is not stable because of the nature of infrared.

The Active Bat [26][27] location system is developed by AT&T researchers. It has a similar infrastructure to Active Badge. A central controller works as location server. The difference between Active Bat and Active Badge is that Active Bat utilizes both radio and ultrasound signals instead of infrared. The controller sends an RF request packet to the Active Bat tag and an RF reset packet to ceiling-mounted receivers or ultrasound sensors at the same time. The tag responds with an ultrasound signal. The receivers calculate the distance based on the time difference between receiving an RF reset packet and detecting ultrasound signals. The accuracy and precision are very high; it can reach to 9cm for 95% of locations. However, the use of ultrasound time-of-flight requires a large fixed-sensor infrastructure throughout the ceiling. So the effect range and scalability of Active Bat is limited.

The Cricket location support system [28] also utilizes both ultrasound and RF signals. But Cricket is based on a mobile-based infrastructure. A small device is attached to a mobile user, the listener, to estimate the distance to the nearest beacon. The mobile device is responsible for calculating its own location. The mobile device measures the ultrasound signal to calculate the distance with TDOA technologies and uses RF signals for synchronization, identifying the beacon and activation of the time measurement. Because it uses a mobile-based infrastructure, Cricket has the benefits of privacy, decentralized administration, and low cost. And it can achieve room-sized granularity of 4x4 feet.

3D-iD [29] is an RF-based commercial location system developed by PinPoint Corp to determine the 3D location of items inside buildings. 3D-iD uses a network-based infrastructure. It divides the space into cells. And each cell consists of a cell controller and multiple antennas. These antennas continuously broadcast a signal. Upon receiving the signal, a tag attached to a mobile device will immediately retransmit the message on another frequency and encode it with its own ID. The cell controller measures the distances between the tag and antennas using RF round-trip time and then uses the geometry method to determine the tag's location. A cell's effective range is 30m and can achieve 1m to 3m accuracy. However, the cell is working at 2.4G Hz, and the tag transmits at 5.78GHz, which will have a radio spectrum collision with WLANs. Besides, the system requires several cells to be installed to cover a building, and this will greatly increase the cost because of the expensive hardware it uses.

RADAR [22] is developed by Microsoft researchers and it is based on WLAN technology. RADAR estimates the distance between access points and mobile devices by measuring the RF signal strength. A propagation pattern-based method and k nearest neighbor algorithm are used to compute the mobile device's 2D location. Two approaches are provided in [22] to create the location map; one is called the empirical method, which is based on field measurements, and the other is a propagation model-based approach. The result shows that the empirical approach has better performance than the propagation model-based approach. But the propagation model-based approach makes deployment much easier.

SpotON [31] is an ad hoc location system. And it derives distance from signal strength attenuation instead of time-of-flight. The SpotON system is built by using RFID badges and AIRID base stations. It is similar to RADAR and 3D-iD in developing a fine-grained tagging technology based on RF signal strength. However, laboratory experiments show that SpotON can achieve better resolution and accuracy than RADAR and with a much lower cost than 3D-iD.

In conclusion, we can see that different technologies have been applied to indoor location systems. As with RADAR, most existing indoor location systems are using the geometry location method, because the underlying hardware system they are using, an infrared system, an ultrasound system or special antennas, can provide an accurate estimation of distance even in an indoor environment. Ultrasound-based and infrared-based systems can achieve very high resolution and accuracy, but their disadvantages in effective range, scalability and cost make their application very limited. WLAN-based indoor location systems, on the contrary, have the advantage of a large cover range, and low cost, but their performance is much lower too. Besides, because of the complexity of indoor environment, the performance an indoor location system also depends on how the system is deployed. As analyzed in Chapter 4, the performance of a location system is severely affected by AP configuration, WLAN NIC card properties, the creation of the propagation map, and the effectiveness of data pre-processing.

2.7 Summary

In this chapter, we gave a detailed introduction of different indoor location technologies. Different WLAN location systems are categorized based on the difference in underlying physical media, system infrastructure, and location estimation methods. There are two major location estimation methods: the geometry method (triangulation method) and the propagation pattern-based method. Due to the complexity of the indoor propagation environment and the fact that RSSI is the only signal signature an IEEE 802.11 WLAN standard can provide, the propagation pattern-based method becomes a major method in indoor WLAN location systems. Three types of propagation pattern classification methods are discussed: the nearest neighbour method, probabilistic method, and neural network-based method. And finally, related indoor position systems are briefly introduced.

3. WLAN and ANN

Wireless location technology refers to the equipment and algorithms that apply to wireless communication systems to determine geographic position, and in some cases, the velocity and direction of traveling wireless devices. Recent developments in such areas as cellular technology, wireless modems and services, and GPS have made the real-time location of mobile devices feasible and cost-effective. In the near future there should be a surge of brand new services.

In a typical mobile communications system, each cell has a radio tower that communicates with mobile devices, including wireless-equipped Pocket PCs and cell phones passing through that cell. There are two primary methods of locating a mobile device. The mobile device itself can determine its own location by using GPS. Or the mobile device's location can be determined externally from its radio signal by measuring the signal's time of arrival at three or more base stations. When the base stations calculate a mobile device's location, the signal propagation pattern-matching method is used. In this method, the mobile device's signal is received at various cell sites, and it is transferred to a mobile switch working as a location server to determine the wireless caller's location by measuring the distinct radio frequency patterns and multipath characteristics of the signal from a single call as it arrives at different cell sites at different times. This unique radio frequency signature of the signal is matched to similar patterns stored in a central database to determine the caller's latitude and longitude.

Unlike the mobile communication system, IEEE 802.11 WLAN usually does not have a function note as a mobile switch, and the standards do not provide the functions or method to accurately measure the arrival time of a wireless signal. Without the help of special devices, the only characteristic of RF signals that can be used to estimate the location is the radio signal strength (RSS). Fortunately, the WLAN standard does provide a way for higher-level applications to access lower level signal strength parameter – the received signal strength indicator or RSSI. RSSI is one of the lower-level parameters available from a wireless network interface card (NIC), and it is the only way for a higher-level application to obtain signal strength information solely based on the IEEE 802.11 standard configurations. Each access point broadcasts its beacon message periodically. So a mobile station can receive the RSSI information from all the APs it can reach, no matter whether it is been associated with or not. In this chapter, we will discuss the IEEE 802.11 Wireless LAN system, and the RSSI - the parameter that the IEEE 802.11 standard provided to measure wireless signal strength.

As we introduced in the previous chapter, the neural network method [24] [32] is one major method used to estimate the mobile station's location from pre-stored radio propagation patterns. In this thesis, a neural network- based scheme will be used to extract location information from the RSSI patterns of a given WLAN. Neural networks have many benefits that can be used in signal analysis. In the following parts of this chapter, after the discussion about WLAN, we will discuss neural networks and their application in the field of signal analysis.

3.1 Wireless Local Area Networks

Wireless local area network (WLAN) refers to wireless networks between computers within one building or a group of buildings. WLANs can be achieved by infrared or radio techniques. The growth of WLAN began in the mid-1980s and was triggered by the US Federal Communications Commission (FCC) decision to authorize the public use of the Industrial, Scientific and Medical (ISM) bands [36]. Since the bands are open to individuals and companies, there has been substantial growth in the area of WLANs, and several kinds of WLAN standards have arisen. The most well-known representatives of WLANs are based on the standards IEEE 802.11[38], HIPERLAN [37], and their variants. The IEEE published a series of 802.11 standards; including 802.11, 802.11b, 802.11g, and 802.11a. HIPERLAN was another WLAN standard developed by group RES10 of the ETSI as a pan-European standard for high-speed WLANs. Just like 802.11, the HIPERLAN/1 standard covers the physical and MAC layers offering data rates from 2 to 25 Mbps in the 5.2 GHz band.

3.2 802.11 Standards Family

Since 1997 IEEE has developed its 802.11 standard for wireless LAN. The 801.11 standards family has had several members. The major standards include 802.11, 802.11a, 802.11b, and 802.11g [39]. IEEE 802.11b works in the 2.4GHz frequency band, like 802.11g, while the IEEE 802.11a solution works in the 5 GHz band. The 802.11 standard defines the media access control (MAC) and physical (PHY) layers for accessing point-based networks and ad hoc networks. In doing so, the initial standard supports three physical layers: infrared, frequency hopping spread spectrum (FHSS), and direct sequence spread spectrum (DSSS). The RF methods in 802.11 operate in the 2.4GHz ISM band. The first extension to

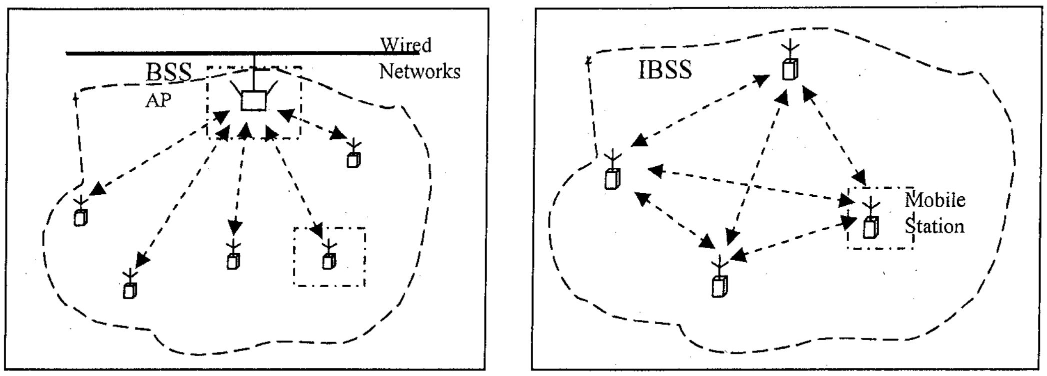
802.11 standards, the 802.11b specification, continues operation in the 2.4GHz band with data rates of 1, 2, 5.5, and 11Mbps. The last two bit rates are achieved through complementary code keying (CCK). Extensions 802.11a and g are for a high bate rate orthogonal frequency division multiplexing (OFDM) PHY standard providing bit rates in the range of 6 to 54Mbps in the 5GHz and 2.4GHz band respectively. Although the physical layer differs for each signaling mechanism, the above standards use the same method of media access control, and the frame formats supported by the MAC layer are relevant for each physical layer. It was supported by the basic 802.11 standard as well as each of the extensions to the standard. The following table compares these standards in the working frequency band and maximum transmitting power in different regions.

Table 3.1 Worldwide Regulatory Frequency and Maximum Transmitting Power for 802.11 WLANs

<table>
<tr><th colspan="2">Location</th><th>Regulatory Band (GHz)</th><th>Maximum Output Power (mW)</th></tr>
<tr><td rowspan="4">North America</td><td>802.11b,g</td><td>2.400 – 2.4835</td><td>1000</td></tr>
<tr><td rowspan="3">802.11a</td><td>5.150-5.250</td><td>2.5 /MHz (max. 50mW)</td></tr>
<tr><td>5.250-5.350</td><td>12.5/MHz(max 250mW)</td></tr>
<tr><td>5.725-5.825</td><td>50 /MHz (max 1000mW)</td></tr>
<tr><td rowspan="3">Europe</td><td>802.11b,g</td><td>2.400 – 2.4835</td><td>10 /MHz (max 100mW)</td></tr>
<tr><td rowspan="2">802.11a</td><td>5.150-5.250</td><td>200</td></tr>
<tr><td>5.470-5.725</td><td>1000</td></tr>
<tr><td rowspan="4">Japan</td><td>802.11b,g</td><td>2.400 – 2.497</td><td>10 /MHz (max 100mW)</td></tr>
<tr><td rowspan="3">802.11a</td><td>5.150-5.250</td><td rowspan="3">Indoor 200</td></tr>
<tr><td>4.900-5.000 (until 2007)</td></tr>
<tr><td>5.030-5.091 (from 2007)</td></tr>
</table>

3.3 802.11 WLAN Working Modes and Topology

802.11 WLANs support two kinds of basic working modes – the infrastructure mode and the ad hoc mode. As shown in Figure 3.1 (a), the infrastructure mode involves the use of an access point, either by itself or connected to a wired LAN. The Infrastructure mode WLAN makes use of the higher-speed wired or wireless backbone. In such a topology, mobile nodes access the wireless channel under the coordination of the AP. The AP and its client mobile stations form a basic service set (BSS). In an infrastructure BSS, all mobile stations communicate the AP. The AP provides both the connection to the wired LAN, if there is any, and the local relay function for the BSS [40]. Thus, if a mobile station in the BSS must communicate with another mobile station in the BSS, it must send the communication to the AP first, and then the AP forwards it to the destination. Though this kind of operation increases the consumption of bandwidth, this method gives AP the capability of buffering the traffic of a mobile station when that station is operating in a very low-power state. In the infrastructure mode, the wireless network adapter cannot directly communicate with other computers. Since it only can receive traffic from APs, it only can collect the RSSI information of APs. Hence, our location system cannot make use of signal-decaying information between mobile stations in infrastructure mode. However, in the WLAN using infrastructure mode, it can detect not only the RSSI of the AP that the computer is associated with, but also the RSSIs of all the APs that have the effect range covering the user's mobile device.



(a). Infrastructure mode **(b). Ad-hoc mode**

Figure 3.1 802.11 WLAN Working Modes

"An ad hoc network is a collection of wireless mobile hosts forming a temporary network without the aid of any centralized administration or standard support services regularly available on the wide area network to which the hosts may normally be connected" [41]. An ad hoc WLAN is a peer-to-peer (computer-to-computer) network that is set up in order to serve a temporary need. When two or more stations are within close proximity and they can communicate in a peer-to-peer manner, the area within which communication occurs is referred to as an independent basic service set (IBSS). Figure 3.1 (b) illustrates a group of stations communicating with one another on a peer-to-peer ad hoc basis that forms an IBSS. An IBSS is a BSS that is formed completely by mobile stations, and there is no connection to a wired network. The IBSS is the entire network, and only those stations communicating with each other in the IBSS are part of the LAN. Unlike other ad hoc networks where information packets are transmitted in a store-and-forward method from the source to the destination, there is no relay function in an IBSS of the 802.11 WLAN standards. If one mobile station wants to communicate with another one, they must locate in each other's communication range. In an ad hoc mode WLAN, the mobile stations can receive signals from each other, so a location system can collect RSSI information of all the links. In an IBSS, an AP could be a member of an ad hoc LAN if it runs in ad hoc mode without connecting to outside wired networks. Meanwhile a mobile station in an ad hoc WLAN can work as a gateway between ad hoc LAN and outside WAN, and all the other mobile stations can access to outside network through it.

3.4 Received Signal Strength Indicator

3.4.1 RSSI in 802.11 Standards

RSSI is one of the parameters available from a wireless network interface card (NIC). There are two sublayers – the higher Physical Layer Convergence Protocol (PLCP) sublayer and the lower Physical Media Dependent (PMD) sublayer in the physical layer of an IEEE 802.11 system. RSSI is a measure in the PMD sublayer of the energy observed at the antenna used to receive the current PLCP protocol data unit (PPDU). RSSI is intended to be used in a relative manner, and it is a monotonically increasing function of the received power. The statement about RSSI in the 802.11[42] standard is

"The received signal strength indicator (RSSI) is an optional parameter that has a value of 0 through RSSI_Max. This parameter is a measure by the PHY sublayer of the energy observed at the antenna used to receive the current PPDU. RSSI shall be measured between the beginning of the start frame delimiter (SFD) and the end of the PLCP header error check (HEC). RSSI is intended to be used in a relative manner. Absolute accuracy of the RSSI reading is not specified."

Figure 3.2 illustrates the PLCP frame format in FHSS mode [42]. RSSI shall be measured during the reception of the PLCP preamble, between SFD and HEC (in FHSS of 802.11). Absolute accuracy of the RSSI reading is not specified in the standards. There is nothing in the 802.11 standard that stipulates a relationship between the RSSI value and any particular energy level, as would be measured in mW or dBm.

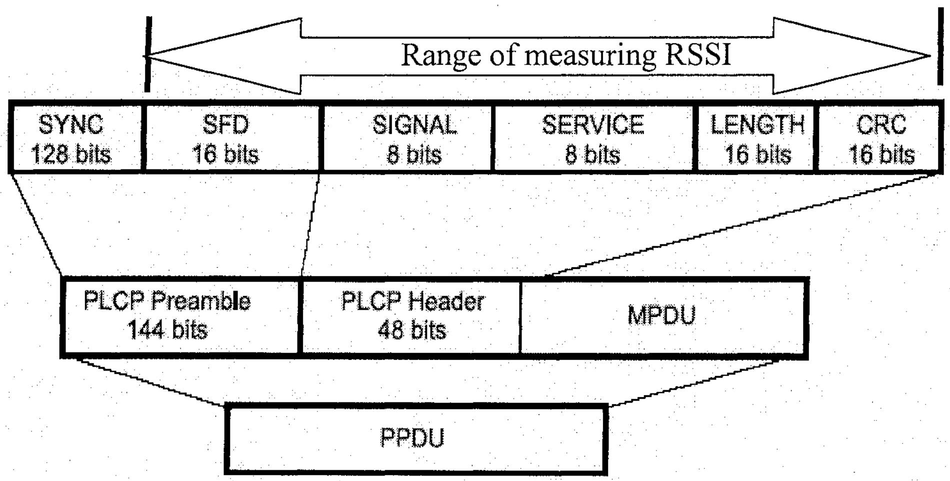


Figure 3.2 Long PLCP PPDU Format

The PMD sublayer sends out a parameter to indicate the magnitude of the energy observed at the selected antenna. This reported value is used to generate the RSSI term in the PLCP sublayer and might also be used by any diversity function. If networks use FHSS, the parameter ranges from 0 to 15. Since RSSI is only used in a relative manner by the MAC sublayer, this PMD parameter is defined to have no more than 16 values, ranging from 0 through RSSI_Max which is parameter set by user. The value zero is the weakest signal strength, while RSSI_Max is the strongest signal strength. If networks use DSSS or OFDM, the RSSI shall be a measure of the RF energy received by the DSSS or OFDM physical layer. RSSI is represented in 8 bits, and up to 256 levels are supported. The value 0 usually refers to the minimum level of available RF energy to the receiver circuit in an 802.11 adapter card. The minimum level is also called "receive sensitivity", and it is an NIC specification measured in dBm. For example, some NIC have a receive sensitivity equal to -96dBm. The

RSSI reported by this card will be 0 whenever the received RF energy is less than -96dBm or 0.2511E-10mW.

RSSI measurements will vary from 0 to 255 depending on the vendor. It consists of a one byte integer value. A value of 1 will indicate the minimum signal strength detectable by the wireless card, while 0 indicates no signal. The value RSSI_Max is different for different vendors. For example, Cisco Systems cards will return an RSSI of 0 to 100. In this case, the RSSI_Max is 100. The Cisco card can report 101 distinct power levels. Another popular WLAN chipset is made by Atheros. An Atheros®-based card will return an RSSI value of 0 to 60. Figure 3.3 illustrates a circuit diagram of generating RSSI [43]. The precision of RSSI is decided by the analog-digital-converter (ADC). Besides that, since the RSSI value is an integer, it must increase or decrease in integer steps. If RSSI changes by 1, it means the power level changed by some proportion in the measured power range.

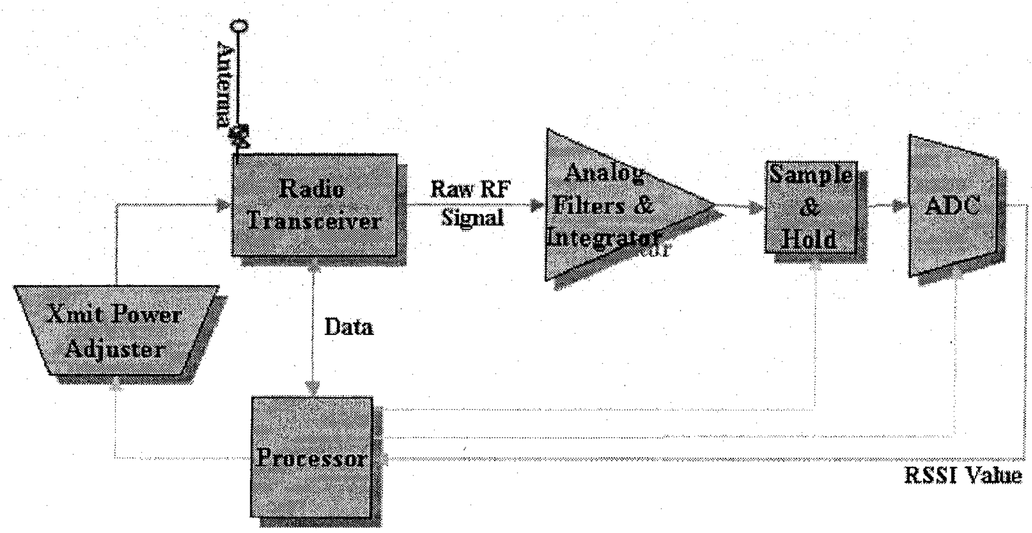


Figure 3.3 Architecture of a Circuit to Measure RSSI

3.4.2 Factors that affect RSSI

The received signal strength and the working range of a wireless station or AP are affected by many factors [44]. The factors include transmit power (equivalent isotropically radiated power - EIRP at the antenna), the receive sensitivity of the wireless adapter, antenna gain, RF propagation pattern of the stations or APs, and obstructions between the source and reception.

Received signal strength is usually proportional with the transmit power of the source station. Take propagation in free space as an example; free space is an ideal propagation medium. Consider an isotropic point source fed by a transmitter of P_t watts. At an arbitrary, long distance d from the source, the radiated power is uniformly distributed over the surface area of a sphere of radius. Thus the received signal power at distance d is given by

$$P_r = \frac{G_e G_t P_t}{(\frac{4\pi d}{\lambda})^2} \tag{3.1}$$

where G_e is the receiving antenna gain, G_t is the transmitting antenna gain, and λ is the wavelength of the electromagnetic wave. It is obvious that the received signal strength will increase when the transmitting power increases. In addition to frequency and bandwidth allocation, transmitting power is also a key parameter that is regulated worldwide. The maximum allowable radiated emissions for the 802.11 DSSS PHY vary from region to region, as shown in Table 3.1. Many of the IEEE 802.11 DSSS PHY wireless products on the market today have selected 100mW as the nominal RF transmits power level. In 802.11 standards about OFDM PHY, three levels of transmit RF power levels are specified: 40mW, 200mW, and 800mW. These three power levels are assigned to different frequency bands in

802.11g. 800mW is for 5.725-5.825GHz, and the power level is suitable for bridging applications. 200mW and 40mW are assigned to 5.250-5.350GHz and 5.150-5.250GHz, and they are suitable for short-range indoor home and small office environments [40].

In the U.S., the FCC defines power limitations for wireless LANs in FCC Part 15.247. Part 15.247 provides details on the limitations of EIRP. EIRP represents the total effective transmit power of the radio, including gains that the antenna provides and losses from the antenna cable.

As one of the most important parts of an RF system, the antenna is used to both transmit a signal as well as shape and focus a received signal. The gain of an antenna represents how well it increases effective signal power in a particular direction, with dBi (decibels relative to an isotropic radiator) as the unit of measure. dBi represents the gain of an antenna as compared to an isotropic radiator, which transmits RF signals in all directions equally. More precisely, dBi equals 10 times the logarithm (base 10) of the electromagnet field intensity of the antenna's favored direction divided by the electromagnetic field intensity of an isotropic antenna with measurements taken at the same distance. All the antennas have a radiation pattern that indicates the power radiated in any direction relative to the direction of maximum radiation. Figure 3.4 illustrates an example of the radiation pattern of an antenna [61].

An isotropic antenna is defined as "a hypothetical lossless antenna having equal radiation in all directions." Clearly, an isotropic antenna is an ideal entity, since even the simplest antenna has some degree of directivity. Although hypothetical and not physically realizable, an isotropic radiator is taken as a reference for expressing the directional properties of actual antennas. A directional antenna is one "having the property of radiating or receiving electromagnetic waves more effectively in some directions than in others." The term is usually applied to an antenna whose maximum directivity is significantly greater than that of a linear dipole antenna. The power pattern of a half-wave linear dipole is shown below.

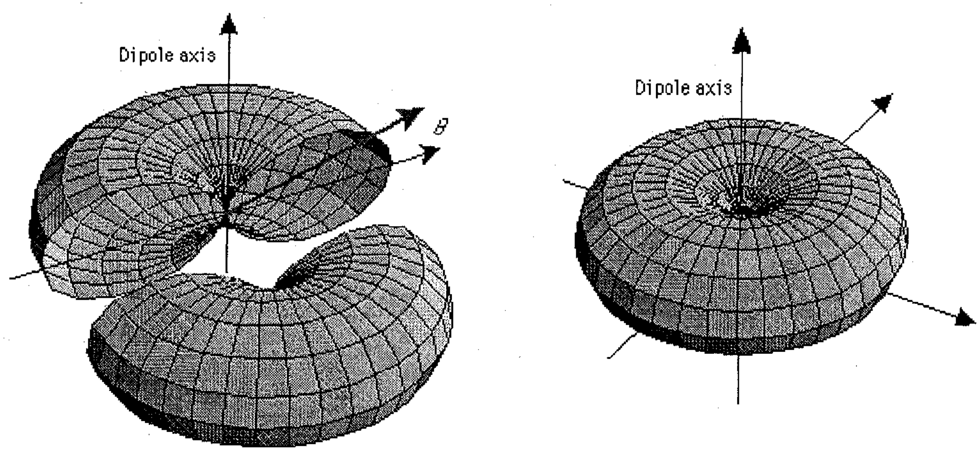


Figure 3.4 Example of an Omni-Directional Antenna Radiation Pattern

The linear dipole is an example of an omni-directional antenna -- i.e. an antenna having a radiation pattern that is nondirectional in a plane. As the figure above indicates, a linear dipole has uniform power flow in any plane perpendicular to the axis of the dipole, and the maximum power flow is in the equatorial plane. Usually, a directional antenna can transmit a higher level of signal power as well as have the ability to receive a lower level of received signal power. However, omni-directional antennas are relatively inexpensive, and most WLAN adapters have omni-directional antennas, which are usually embedded into the edge of WLAN adapter cards.

Besides the antenna, the most critical item in a WLAN system is the receiver. In particular, it is important to make clear the receiver sensitivity. The receive sensitivity is the lowest-level signal that can be decoded by the receiver. The lower the receive sensitivity is, the longer the coverage range is. When a signal is received in the PMD sublayer of a wireless card, the RSSI of the received signal must be compared with a threshold value, which is specified in the standard as ED_THRESHOLD. If the RSSI is smaller than the threshold, the received signal will be discarded. This threshold decides a receiver's sensitivity.

As we mentioned earlier, the received signal strength at the destination is greatly affected by the propagation environment between the source and destination. The RF communication media for home, enterprise, and manufacturing environments are very different, and no environments are the same. Multipath and path loss are issues to be considered when designing an IEEE 802.11 WLAN system. The surface of furniture, elevator shafts, walls, factory machinery, and metal-constructed buildings all contribute to the amount of delay spread that caused by multipath. Path loss is the signal power lost as the distance between source and destination increase. For indoor applications beyond 20 feet, propagation loss increases at about 30dB per 100 feet. This occurs because of the combination of attenuation by walls, ceilings, and furniture. Each wall constructed with sheet rock and wood typically attenuates the signal by 6dB, and walls constructed with cement block attenuate the signal by 4dB. Multipath fading is another key contribution to path loss. Multipath fading occurs when the reflected signal paths refract off walls, windows, furniture, and people and scatter the transmitted signal. Sometimes just moving the receiver a few inches farther from the transmitter can produce an additional multipath fading loss of signal power in the order of 20dB or more. In Chapter 4 we will analyze indoor propagation environment and its effect on RSSI reception in more detailed.

3.4.3 Application of RSSI

In the IEEE 802.11 standard, RSSI is an arbitrary integer, and it was originally designed for internal usage by the microcode on the adapter and by the device driver. RSSI can be used in an adapter to determine when the amount of radio energy in the channel is below a certain threshold, at which point the network card is clear to send (CTS).

The media access control layer specification for 802.11 has similarities to the 802.3 Ethernet wired line standard. The protocol for 802.11 is a protocol scheme know as carrier-sense multiple access, collision avoidance (CSMA/CA). This protocol avoids collisions instead of detecting a collision by sensing channels before transmitting signals. The MAC layer operates together with the physical layer by sampling the transmitted energy over the medium transmitting the data. The physical layer uses a clear channel assessment (CCA) algorithm to determine whether the channel is clear. This is accomplished by measuring the RF energy or RSSI at the antenna and determining the strength of the received signal. If the RSSI is below a specified threshold, the "Clear Channel Threshold", the channel is declared clear and the MAC layer is given the clear channel status for data transmission. A "Roaming Threshold" in 802.11 is also based on the measurement of RSSI to determine the point when a client should be handover to a new AP. As discussed above, the PMD PHY sublayer measures the RSSI during the reception of the PLCP preamble, and compares the RSSI with predefined threshold. If the RSSI is bigger than the threshold, the PMD sublayer will inform the upper layer of the arrival of data.

Since RSSI is a parameter for the channel quality between the source station and the destination station and is related to the distance and channel properties, RSSI is useful in location management, routing, and RF energy distribution modeling [45][46]. For example, SpotON [31] is a system that is used to analyze ad hoc location sensing. SpotON tags use received radio signal strength information as an inter-tag distance estimator. The obtained locations might not be exactly the physical location of the nodes or stations, but the distance is described in link qualities. The location information is useful for routing. It can be used to form the routing table. RSSI is also useful to estimate channel properties and model a specified channel. Together with other signal quality parameters, RSSI can also provide a base for channel assignment [47] or intelligent channel management.

3.5 Artificial Neural Networks (ANN)

Half a century has passed since the first model neural network was developed by McCulloch and Pitts in 1943. The research and development of artificial neural networks have gone through a rough road. And ANN has increased more and more over the last few decades, and is being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology, and physics. Neural networks are widely used to solve problems of prediction, classification, or control because of the following advantages of ANN. First advantage is its power. Neural networks are very sophisticated modeling techniques capable of modeling extremely complex functions. In particular, neural networks are nonlinear. Neural networks have advantages in dealing with complex nonlinear modeling problems. The second advantage is ease of use. Neural networks learn by example. The neural network user gathers representative data and then invokes training algorithms to automatically learn the structure of the data. Although the user does need to have some heuristic knowledge of how to select and prepare data and to select an appropriate neural network, the level of user knowledge needed to successfully apply neural networks is much lower than would be the case using some more traditional nonlinear statistical methods.

3.5.1 Fundamentals of Artificial Neural Networks

An ANN is a computing paradigm that is loosely modeled after cortical structures of the brain. It consists of interconnected processing elements called neurons that work together to produce an output function. The key element of this paradigm is the novel structure of the information processing system. The output of a neural network relies on the cooperation of the individual neurons within the network to operate. Processing of information by neural networks is often done in parallel rather than in series or sequentially. Usually, an ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Similar to biological systems, learning in ANN systems involves adjustments to the synaptic connections that exist between the neurons.

3.5.1.1 Neurons

In biological neural systems, transmission of information between neurons is performed by the exciting signal, which raises the receiver's electric potential, and the suppressing signal, which lowers the receiver's electronic potential. The receiver neuron decides whether to send the information or not based on the electronic potential and the threshold of the neuron: when the electric potential exceeds the threshold, the information is transmitted, otherwise it is not.

To capture the essence of biological neural systems, an artificial neuron can be defined as follows:

- It receives a number of inputs either from the original data or from the output of other neurons in the neural network. Each input comes via a connection that has a strength or weight; these weights correspond to the synaptic efficacy in a biological neuron.

Each neuron also has a single threshold value. The weighted sum of the input is formed and the threshold subtracted to compose the activation of the neuron, also known as the post-synaptic potential, or PSP, of the neuron.

- The activation signal is passed through an output function also known as a transfer function to produce the output of the neuron.

If the step output function is used, i.e., the neuron's output is 0 if the input is less than zero, and 1 if the input is greater than or equal to 0, then the neuron acts just like the biological neuron described earlier by subtracting the threshold from the weighted sum and comparing with zero. And this is equivalent to comparing the weighted sum to the threshold. Actually, the step function is rarely used in artificial neural networks. And weights can be negative, which implies that the synapse has an inhibitory rather than excitatory effect on the neuron.

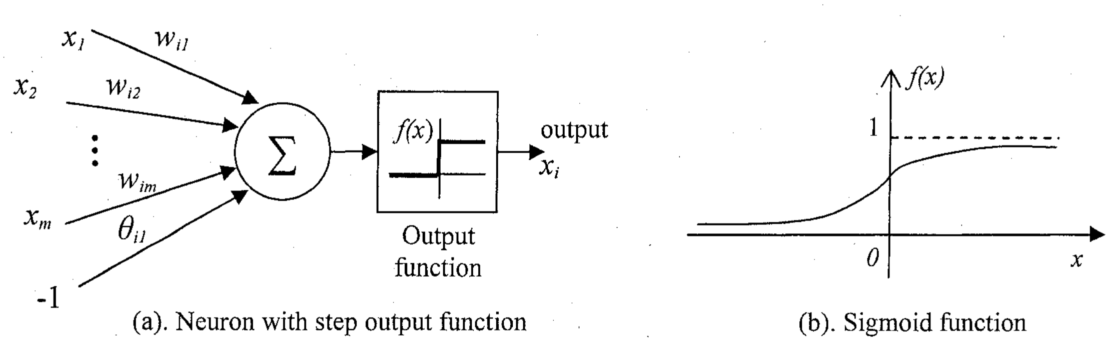


Figure 3.5 Artificial Neuron Model

The definition of an artificial neuron can be illustrated using the above figure and the following formula:

$$x_i(n+1) = f(\sum_{j=1}^{m} w_{ij}x_j(n) - \theta_i) \qquad (3.2)$$

Where $x_i(n)$ is the ith neuron output at the time n; w_{ij} is the synapse weight from the jth neuron to the ith neuron, the positive w_{ij} generates an exciting signal, and the negative w_{ij} generates a suppressing signal; θ_i is the threshold of the ith neuron; and $f(x)$ is the output function of the neuron. There are many different output functions; some of the most commonly used are the step function and the sigmoid function. (a) uses the step function, which can be defined as:

$$f(x)=\begin{cases}1, x>0\\0, x\leq 0\end{cases} \tag{3.3}$$

The step function is used in discrete neural network models. For continuous neural network models, to avoid mathematical difficulty in handling discontinuity, the sigmoid function is widely used. A sigmoid function is defined as:

$$f(x)=\frac{1}{1+e^{\frac{-x}{T}}} \tag{3.4}$$

where T is a positive parameter to control the slope of the function. When T is smaller, the slope of the function becomes steeper and more approaches to the step function.

3.5.1.2 ANN Structure

ANN can be divided into feed-forward and recurrent classes according to their connectivity. An ANN is feed forward if such a method exists, which numbers all the nodes in the network such that there is no connection from a node with a large number to a node with a smaller number. All the connections are from nodes with small numbers to nodes with large numbers. An ANN is recurrent if such a numbering method does not exist. In other words, feed-forward networks allow signals to travel one way only, from input to output. There is no feedback or loops, so that the output of any layer does not affect that same layer or higher layers. Feed-forward ANNs tend to be straight-forward networks that associate input with output. Examples of feed-forward ANNs include the multilayer perceptron (MLP), radial based functions (RBF). Recurrent ANNs can have signals traveling in both directions by introducing loops in the network. Recurrent networks are very powerful and can get extremely complicated. They are dynamic, and their state changes continuously until they reach an equilibrium point. The networks remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Examples of recurrent ANNs include the Hopfield neural network and the Boltzmann machine.

The most common type of artificial neural network consists of three groups or layers of units: a layer of input units is connected to a layer of hidden units, which is connected to a layer of output units. The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units. This simple type of network is interesting because the hidden units are free to construct their own representations of the input. The weights between the input and hidden units determine when each hidden unit is active, and so by modifying these weights, a hidden unit can choose what it represents. Based on the number of layers, ANNs can be divided into single-layer or multi-layer organizations. In a single-layer ANN, all units or neurons are connected to one another. In multi-layer networks, units are often numbered by layer instead of following a global numbering.

3.5.1.3 Learning in ANNs

Learning in ANNs is typically accomplished by using examples or training samples. The learning procedure is called training because the learning is achieved by adjusting the connection weights in an ANN iteratively so that trained ANNs can perform certain tasks. Learning in ANN can be divided roughly into supervised, unsupervised, and reinforcement learning.

An important part of supervised learning is an external teacher. The teacher has full knowledge of the problem and knows the desired output given by each input signal. Supervised learning is based on direct comparison between the desired output given by the teacher and the actual output of an ANN. The training procedure is the process of minimizing the error function, such as total mean square error, between the actual output and the desired output summed over all training data. Usually, gradient decent-based optimization algorithms such as backward propagation can be used to adjust weights and thresholds in an ANN. This adjustment or training procedure is carried out iteratively in steps until it reaches the training goal.

Reinforcement learning is a special case of supervised learning. The deference between reinforcement learning and common supervised learning is that the exact desired output is not given in reinforcement learning. It is based only on the information of whether or not the actual output is correct. Reinforcement learning is often called learning with a critic as opposed to learning with a teacher. In reinforcement learning, it is common to think explicitly of a network functioning in an environment. The environment supplies inputs to the network, receives output, and then provides a reinforcement signal.

Unsupervised learning does not use external teachers and is only based on local information. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects data's emergent collective properties. Unsupervised learning is solely based on the correlations among input data. Paradigms of unsupervised learning include Hebbian learning and competitive learning.

3.5.2 Multilayer Perceptron (MLP)

Multilayer perceptron is by far is the most well-known and most popular neural network among all of the existing neural network paradigms. MLP is one kind of feed-forward neural networks. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes, if any, and to the output nodes. There are no cycles or loops in the network. In an MLP, the basic unit or neuron is called perceptron.

The perceptron is a type of ANN invented in 1957 at the Cornell Aeronautical Laboratory by Frank Rosenblatt. It can be seen as the simplest kind of feed-forward neural network: a linear classifier. Similar to the neuron model introduced above, in the perceptron model, a single neuron with a linear weighted net function and a threshold activation function is employed. The input to this neuron is a feature vector x = $\{x_j, j = 1 \dots n\}$ in an n-dimensional feature space. The net function is the sum of the weighted inputs vectors.

$$net(x) = \sum_{j=1}^{n} w_j x_j + w_0 \tag{3.5}$$

$$y(x) = \begin{cases} 1, net(x) \geq 0 \\ 0, net(x) \prec 0 \end{cases} \quad \text{or} \quad y(x) = f(net(x)) \tag{3.6}$$

where $w = \{w_j, j=1..n\}$ is a vector of real-valued, w_0 and is the 'bias', a constant term that does not depend on any input value. The sign of $y(x)$ is used to classify x as either a positive or a negative instance, in the case of a binary classification problem. The bias can be thought of as offsetting the activation function or giving the output neuron a "base" level of activity. If w_0 is negative, then the weighted combination of input must produce a positive value greater than $-w_0$ in order to push the classifier neuron over the 0 threshold. Spatially, the bias alters the position (though not the orientation) of the decision boundary. Since the inputs are fed directly to the output unit via the weighted connections, the perceptron can be considered the simplest kind of feed-forward neural network.

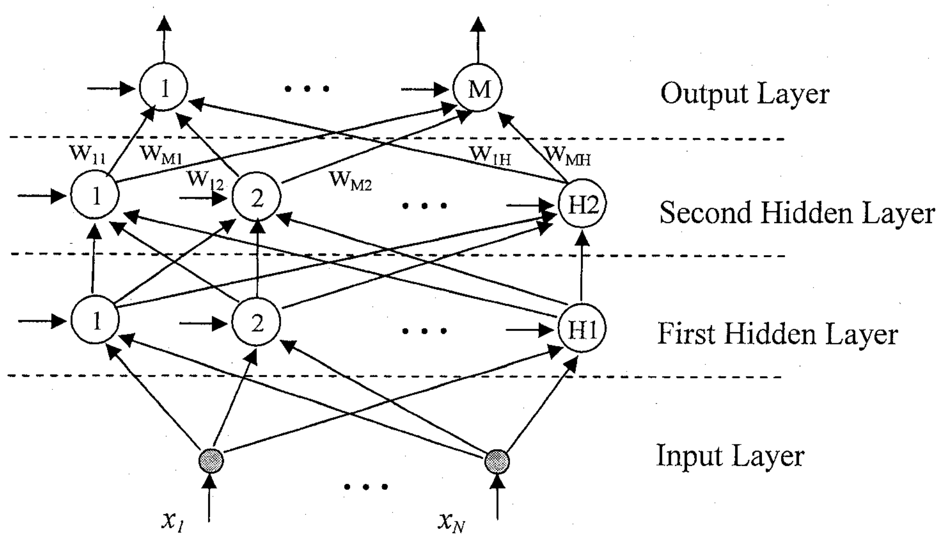


Figure 3.6 a Four-Layer MLP Configuration

The upper figure shows a typical MLP. Each big circle in the graph represents a neuron, and these neurons are organized into layers. There are four layers in the above MLP: input layer, the first hidden layer, the second hidden layer, and the output layer. The gray spots represent nodes in the input layer. Input nodes are not neurons, they just represent the input vector. An MLP has to consist of at least one layer of hidden nodes. The name "hidden layer" refers to the fact that the output of neurons will be fed into the upper layer neurons and therefore is hidden from the user who can only observe the output neurons in the output layer. In an MLP, two consecutive layers are completely connected by weights, w = $\{w_{ij}, i = 1 \ldots P, j = 1 \ldots Q\}$ where P is the number of neurons in the higher layer; Q is the number of neurons in the lower layer; w_{ij} represents the weight between ith neuron of the higher layer and jth neuron of the lower layer. There are no weights that connect two non-consecutive layers, and no weights that connect neurons in the same layer, and no weights that feed back signals from the output layer to the input layer. Figure 3.6 shows an MLP with two hidden layers.

3.5.2.1 Nonlinear Mapping of an MLP

An MLP provides nonlinear mapping between its input and output. In other words, an MLP is a static neural model in the sense that its input-output relationship can be described by a nonlinear mapping function. It has been proven that, with a sufficient number of hidden nodes, an MLP with as few as two hidden layer neurons is capable of approximating an arbitrarily complex mapping with finite support [47]. [34] and [35] also show that MLPs with a sufficient number of nonlinear units in a single hidden layer have been established as universal function approximators. MLPs have several significant advantages over conventional approximations [33]. First, MLP basic functions – hidden unit output- change adaptively during training, making it unnecessary for the user to choose them beforehand. Second, the number of free parameters in the MLP can be unambiguously increased in small increments by simply increasing the number of hidden neurons. And third, MLP basic functions are bounded, making round-off and overflow errors unlikely [4]. However, since the complexity of the problem varies from one application to another, the complexity of the applied MLP varies according to the complexity of the application. In general, the complexity of the MLP depends on its topology, which consists of the number of hidden neurons and the number of hidden layers. An MLP with optimal topology should be able to provide the best approximation accuracy to the unknown model using the most appropriate number of hidden neurons and hidden layers. In general, the common principle indicates that the most appropriate number of hidden neurons is application-dependent and can only be decided empirically during the early stages of topology design. However, theoretically, an MLP is a universal approximator, and it can be applied to approximate the unknown model to any degree of accuracy, and one or two hidden layers are all it takes to reach this arbitrary mapping capability.

3.5.2.2 Back Propagation Learning Algorithm

Multi-layer networks use a variety of learning techniques, the most popular being back propagation. Back propagation is a supervised learning technique. It was first described by Paul Werbos in 1974 and further developed by David E. Rumelhart, Geoffrey E. Hinton , and Ronald J. Williams in 1986.

Back propagation requires that the transfer function or net function used by the artificial neurons be differentiable. By presenting the training data of the desired output to the MLP, the actual output values are compared with the desired output to compute the value of the predefined error-function. The error is then fed back through the network from the output layer to the input layer. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. Gradient descent techniques are usually used to adjust weights to achieve non-linear optimization. The derivative of the error function with respect to the network weights is calculated, and the weights are then changed such that the error can be reduced. After repeating this process for a sufficiently large number of training cycles, the network will usually converge to some state where the error of the calculations is small. In this case, one says that the network has learned a certain target function. The process of back propagation can be illustrated by the following steps:

1. Apply input vectors to input nodes and calculate the actual output values.

The output of neuron i in the MLP y_i is

$$y_i = f_i(\sum_{j \in A_i} w_{ij} y_j) = f_i(net_i) \tag{3.7}$$

where f_i is the activation function of node i and $A_i = \{j : \exists w_{ij}\}$ is the set of nodes anterior to node i. y_j is the output of anterior node j. If the anterior nodes are input nodes, then y_j represents the input vector $x_p = (x_{p1}, x_{p2}, \ldots, x_{pN})$, where $p = 1, \ldots P$ represents the patterns of the training samples. N is the number of input nodes. The activity of the input nodes is determined by the network's external input **x**. For all other nodes, the activity is propagated forward according to Formula 3.7. And finally the output of the output nodes can be calculated. Note that, before the activity of node i can be calculated, the activity of all of its anterior nodes must be known.

2. Calculate the error function.

Given the desired output to specified input vector, the error function is:

$$E = \frac{1}{2P}\sum_{p=1}^{P}\sum_{m=1}^{M}(y_{pm} - o_{pm})^2 = \frac{1}{2P}\sum_{p=1}^{P}\sum_{m=1}^{M}(f(net_{pm}) - o_{pm})^2 \tag{3.8}$$

The desired output of the MLP for pattern p is $o_p = (o_{p1}, o_{p2}, \ldots, o_{pM})$, where M is the number of output nodes. In the above function, y_{pm} is the actual output value of the output node m for input pattern p, and o_{pm} is the desired output value of the output node m for input pattern p. The error is summed over the neurons in the output layer and over all the training patterns.

3. Calculate how much the weight should be adjusted.

The error E is minimized in most cases by using gradient descent, whereby weights are adjusted in such a manner that E is always made smaller. Then the gradient of weight w_{ij} can be written as:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \tag{3.9}$$

Besides gradient descent, there are several other methods, such as conjugate gradients, which involve a search along selected directions in weight space, as opposed to simple gradient descent. Such methods can sometimes improve the training speed, but they also add to system complexity.

By applying the chain, we can get

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial net_i}\frac{\partial net_i}{\partial w_{ij}} = \frac{\partial E}{\partial net_i}\frac{\partial}{\partial w_{ij}}\sum_{k\in A_i} w_{ik}y_k = \frac{\partial E}{\partial net_i}y_j = \delta_i y_j \tag{3.10}$$

where $\delta_i = \frac{\partial E}{\partial net_i}$. From the above formula, we can see that back propagation can be simplified as calculating the derivative of the error with respect to the weights.

4. Error back propagation

For hidden units, the error can be propagated from the output nodes. Again using the chain rule, the error of a hidden node can be expanded in terms of its posterior nodes

$$\delta_j = -\sum_{i\in P_j}\frac{\partial E}{\partial net_i}\frac{\partial net_i}{\partial y_j}\frac{\partial y_j}{\partial net_j} = f_j'(net_j)\sum_{i\in P_j}\delta_i w_{ij} \tag{3.11}$$

where $P_j = \{i : \exists w_{ij}\}$ is the set of nodes posterior to node j. Note that, in order to calculate the error for node j, the error of all its posterior nodes is be given. As long as there are no cycles in the network, there is an ordering of nodes from the output nodes back to the input that respects this condition.

5. Update the weights and biases. And repeat the above steps to keep updating weights and biases until the error E is small enough and no modification of weights is made. Then the training is over.

4. WLAN Indoor Propagation Environment

Although the neural network-based location calculation method does not rely on priori knowledge of any environment parameters such as location of access points and building characteristics, it is necessary to analyze the propagation pattern of WLAN RF signal strength. The properties of the propagation pattern, such as resolution, accuracy, and precision, determine the performance of all propagation map-based WLAN location systems. By analyzing the RF signal propagation model in an indoor environment, we can find out the rules for RSSI distribution, which is a dominant factor that affects the signal signature in a propagation map. After we understand these basic rules, a better way to create the propagation map or the setup training data set for neural networks can be achieved. In this chapter, we will discuss the basic mechanisms of radiowave propagation first. And then two types of RF signal fading -- large-scale fading and small-scale fading -- are discussed, and several fading models to describe the fading are also presented. Analysis of field-measured data has proven the effectiveness of those models. Based on the indoor propagation models of path loss, we will discuss issues of how to create the propagation map in the third section. In the first part of the third section, we will discuss the effect of RSSI quantization on the RSSI propagation pattern. Then a concept, "RSSI resolution" over space, is introduced, and the relationship of RSSI resolution with the distance to the transmitter, the path loss exponent, and RSSI granularity is analyzed. And in the end, based on the analysis and simulation results, basic rules of creating a propagation map are proposed.

4.1 Propagation Physics Mechanism

Radio wave propagation generally means the movement of electromagnetic waves from a transmitting antenna to a receiving antenna in the presence of one or more types of ground, troposphere, ionosphere, terrain, man-made obstacles, etc. As we discussed in Chapter 3, the exact mechanism of propagation depends on the frequency used, the gains of the antenna, the proximity of the antennas to the ground, the electrical characteristics of the atmosphere, terrain features, building structures and materials, etc. Maxwell's equations govern the behavior of electromagnetic waves in space and time. Plane waves are the solution to Maxwell's equations in the Cartesian coordinate system, where the position vector is specified by its distance (x, y, z) along the coordinate axes. The following are Maxwell's equations [50].

$$\vec{e}(x,y,z;t) = \mathrm{Re}[\vec{E}(x,y,z)e^{j\omega t}] \quad (4.1)$$

$$\vec{h}(x,y,z;t) = \mathrm{Re}[\vec{H}(x,y,z)e^{j\omega t}] \quad (4.2)$$

$$\nabla \times \vec{E} = -j\omega\mu\vec{H} - \vec{M} \quad (4.3)$$

$$\nabla \times \vec{H} = j\omega\varepsilon\vec{E} + \sigma\vec{E} + \vec{J} \quad (4.4)$$

$\vec{e}(x,y,z;t)$ and $\vec{h}(x,y,z;t)$ are the electric and magnetic field intensities of a plane wave, and t here is the time variable. $\omega = 2\pi f$ is the radian frequency of the oscillating wave. (4.3) and (4.4) are simplified notions when the propagation media has constitutive parameters, including permittivity ε, permeability μ, and conductivity σ. $\vec{M}$ and $\vec{J}$ denote the phasor magnetic and electric current densities respectively. ε, μ, and σ can, in general, change with position as well as frequency.

When a radio wave propagates in an environment with obstacles, based on the difference of the obstacle's surface roughness, materials, shape, size, and quantities, the radio wave could be absorbed, reflected, or diffracted. The radio wave's propagation behavior or mechanisms can generally be categorized into three basic propagation mechanisms: reflection, diffraction, and scattering [51]. An important factor that affects the radio wave's propagation is the obstacle's size. When the obstacle's dimension is much bigger than the wave length of the radio wave, reflections could occur when the radio wave impinges upon the surface of the obstacle. Diffraction is also called "shadowing". It happens when a bigger-than-wavelength-size impenetrable obstacle blocks the line-of-sight (LOS) between the transmitter and receiver. When multiple obstacles in the radio propagation channel have sizes on the order of wavelengths or less, scattering could happen. The radio wave diffracts among the obstacles and is finally radiated in many different directions. Figure 4.1 illustrates the above three propagation mechanisms.

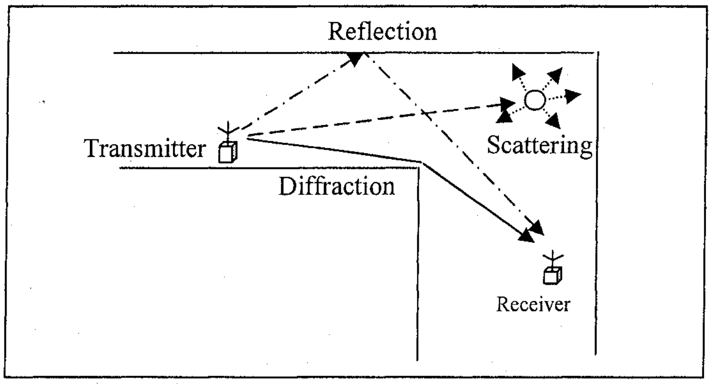


Figure 4.1 Three Radiowave Propagation Mechanisms

The effects of diffraction and scattering have been studied for many years. Scattering, which follows the same physical principles as diffraction, is proven to be the most difficult of the three propagation mechanisms to predict. It is well known that objects whose dimensions are small compared to a wavelength act as weak scattering centers, with effective scattering cross-section proportional to the fourth power of the impinging wavelength. Since the wavelength of a WLAN is 12cm in the 2.4 to 2.5GHz ISM band and around 5.5cm in the 5.2 and 5.8 GHz UNII band, the dominant scattering obstacles in an indoor WLAN environment usually have sizes from 2 to 3 cm to human dimensions of 2 to 3 meters. So most scatterings that happen in indoor environments fall into the regime of Mie scattering [52], which is a complete analytical solution to Maxwell's equations for the scattering of electromagnetic radiation by spherical particles whose dimension is one to ten times bigger than wavelengths.

All of the three propagation mechanisms contribute to the instantaneous received signal strength at the receiver. However, in different propagation environments, each propagation mechanism contributes differently. For example, when the receiver is located in a wide-open area with a clear LOS to the transmitter, the effect of diffusion and scattering can usually be ignored. In the indoor environment without an open area, where there is usually no LOS reception, diffraction and scattering are most likely to dominate the propagation.

Reflection, diffraction, and scattering result in multiple paths for a signal transmitted from the transmitter to the receiver. Thus, at the receiver, the received signal should be the sum of all of these multipath signals. Because the paths traversed by these signals are different, these signals interact with each other. The effects of multipath include constructive and destructive interference and phase shifting of the signal. And this causes multipath fading, such as Rayleigh fading or Rician fading. Multipath fading is often referred to as small-scale fading, which is used to describe local effects on the rapid fluctuations of the instantaneously received power. Besides of this local effective small-scale fading, the large-scale effects of path losses cause the received power to vary gradually relative to the distance between the transmitter and receiver. In the following section, we will discuss analytical models that can describe large-scale and small-scale fading.

4.2 Indoor Propagation Modeling

Propagation models are usually aimed to predict the average received signal strength at a given distance from the transmitter, as well as the variability of the signal strength in close spatial proximity to a particular location [53]. Propagation models can be divided into two categories based on the scale they are describing. Large-scale propagation models predict the mean signal strength for an arbitrary transmitter-receiver (T-R) separation distance. They usually describe the signal characteristics over large T-R distances and determine a power level averaged over an area of tens or hundreds of meters, therefore called the "area-mean" power. On the other hand, small-scale propagation models describe the rapid fluctuations of received signal strength over very short travel distances, usually less than a few wavelengths, or short time durations in the order of seconds.

4.2.1 Large-Scale Propagation Models -- Path Loss

As we can see, modeling radio wave propagation is a difficult task, especially when the propagation environment is complicated with unknown variance in ε, μ, and σ. To account for the effects of various environmental factors, people always compare the actual received signal to the signal one would receive under free-space conditions, as the Friis transmission equation (equation 3.1) describes.

$$P_r = G_e G_t P_t \left(\frac{\lambda}{4\pi d}\right)^2 \qquad (3.1)$$

In a non-free-space environment, the effects of various mechanisms can be lumped into a single parameter, and the Friis transmission equation is modified as

$$P_r = G_e G_t P_t (\frac{F^2}{PL_0})^2 \tag{4.5}$$

where F is known as the propagation factor, and $PL_0 = (4\pi d / \lambda)^2$ is the free-space path loss defined as the ratio of P_t to P_r when the antennas have unit gains. Many propagation problems are aimed to determine F for a given environment.

In both indoor and outdoor environments, the average large-scale path loss for an arbitrary T-R separation distance is expressed as a function of distance. Theoretical analysis and field measurements have proven that the average received signal power decreases logarithmically with distance. The path loss at distance d to the transmitter can be expressed as the following equation, which is also referred as log-distance path loss model [53].

$$\overline{PL}(dB) = \overline{PL}(d_0) + 10n\log(d / d_0) \tag{4.6}$$

where n is the path loss exponent, d_0 is the reference distance, and $\overline{PL}(d_0)$ is the average received signal power at reference distance d_0. This value n is dependent on the specific propagation environment, i.e. type of construction material, architecture, and location within a building. The smaller n is the lower the signal path loss. The value of n ranges from 1.2 to 8 [54]. In free space, n is equal to 2. If the shadowing fading is included, the above equation can be written as:

$$\overline{PL}(dB) = \overline{PL}(d_0) + 10n\log(d / d_0) + X_\sigma \tag{4.7}$$

where X_σ is a zero-mean Gaussian distributed random variable in dB with a standard deviation of σ. The random variable indicates the shadowing effect. Shadowing, also referred as log-normal shadowing, represents the effect of different travel paths on the different locations that have the same distance to the transmitter.

Many empirical outdoor path-loss models have a similar form to the log-distance path loss model. These models vary widely in their approach, complexity, and accuracy. And most of them are based on a systematic interpretation of measurement data obtained in several areas. However, applications of these models are not restricted to the environments in which the model was made. Many wireless systems use these models as a basis for performance analysis. Commonly used outdoor path-loss models include the Okumura model, the Hata Model, the Walfisch and Bertoni Model, etc.

The log-distance path loss model can also apply in indoor environments. Many researchers have shown that indoor path loss obeys the following equation:

$$\overline{PL}(dB) = \overline{PL}(d_0) + 10n\log(d/d_0) + Y_\sigma \qquad (4.8)$$

The above equation has the same form to the log-normal shadowing model of equation 4.7. Y_σ is a normal random variable in dB having a standard deviation of σ dB. The pass value of σ and loss exponent n depends on the surroundings and building type. Table 4.1 shows some typical values for various indoor environments [51]. In the table, hard partitions refer to the fixed walls or aisles in office buildings. And soft partitions refer to the moveable obstructions such as wall panels or furniture panels that are lower than the ceiling. The value of σ implies how accurate a path loss model predicts the actual loss in a building. The smaller σ is, the more accurate the model is.

Table 4.1 Path Loss Exponent and Standard Deviation in Different Indoor Environments for Log-Distance Path Loss Model

Environment	Frequency (MHz)	n	σ (dB)
Retail store	914	2.2	8.7
Office, hard partition	1500	3.0	7.0
Office, soft partition	1900	2.6	14.1
Chemical factory (obstructed)	4000	2.1	9.7
Chemical factory (LOS)	4000	2.1	7.0
Suburban home	900	3.0	7.0

4.2.2 Attenuation Factor Model

The attenuation factor model is a more accurate model than the log-distance model. The log-distance model can be applied to most kinds of indoor environments by selecting a different path loss exponent n and standard deviation σ. But standard deviation between measured and predicted path loss by the log-distance model can reach 13 dB. The attenuation factor model [54] increases this accuracy by introducing site-specific factors. It has been proven that the attenuation factor model can reduce the standard deviation between measured and predicted values to 4 dB. The attenuation factor model can be expressed as

$$\overline{PL}(dB) = \overline{PL}(d_0) + 10n_f \log(d/d_0) + FAF[dB] + \sum PAF[dB] \qquad (4.9)$$

where n_f is the path loss exponent for one floor, and FAF is a floor attenuation factor that represents the losses between the floors of a building, and PAF is a partition attenuation factor that refers to the loss caused by the obstruction located on the path from the transmitter to the receiver. The total path loss is the summary of the log-distance factor, the path loss caused by different floors, and all the obstructions in the propagation path. It has been shown that this model can yield excellent accuracy while offering big computational efficiency [55]. n_f can be selected from Table 4.1 according to the type of the environment. Tables 4.2 and 4.3 give some typical values of FAF and PAF [54] [53].

Table 4.2 Average Floor Attenuation Factors in dB in Two Buildings

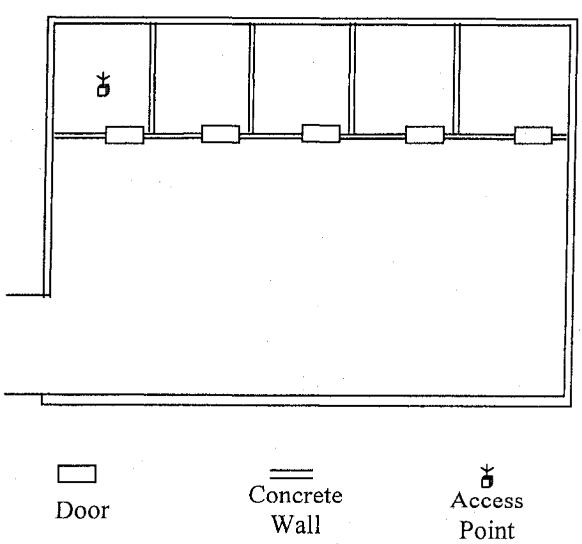
Buildings (Office Building)	FAF(dB) Building 1	FAF(dB) Building 2
Through one floor	16.2	12.9
Through two floor	27.5	18.7
Through three floor	31.6	24.4

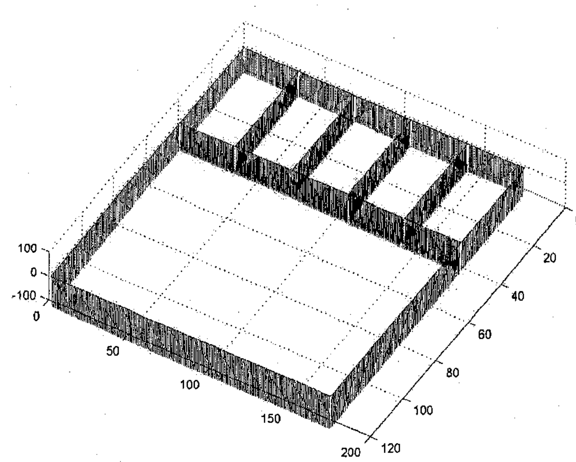
Table 4.3 Average Signal Loss for Different Obstructions

Obstruction (Materials Type)	PAF (dB)	Frequency(MHz)
All metal	26	815
Concrete block wall	13-20	1300
Metal catwalk/stairs	5	1300
0.6m square reinforced concrete pillar	12-14	1300
Concrete floor	10	1300
Dry plywood (3/4 inch)- 1 sheet	1	9600
Dry plywood (3/4 inch)- 2 sheet	4	9600
Wet plywood (3/4 inch) - 1 sheet	19	9600
Wet plywood (3/4 inch) - 2 sheet	39	9600
Aluminum (1/8 inch) – 1 sheet	47	9600
Light textile	3-5	1300
Heavy textile	8-11	1300
General Machine (10-20 sq ft)	5-10	1300
Metal blanket -12 sq ft	4-7	1300

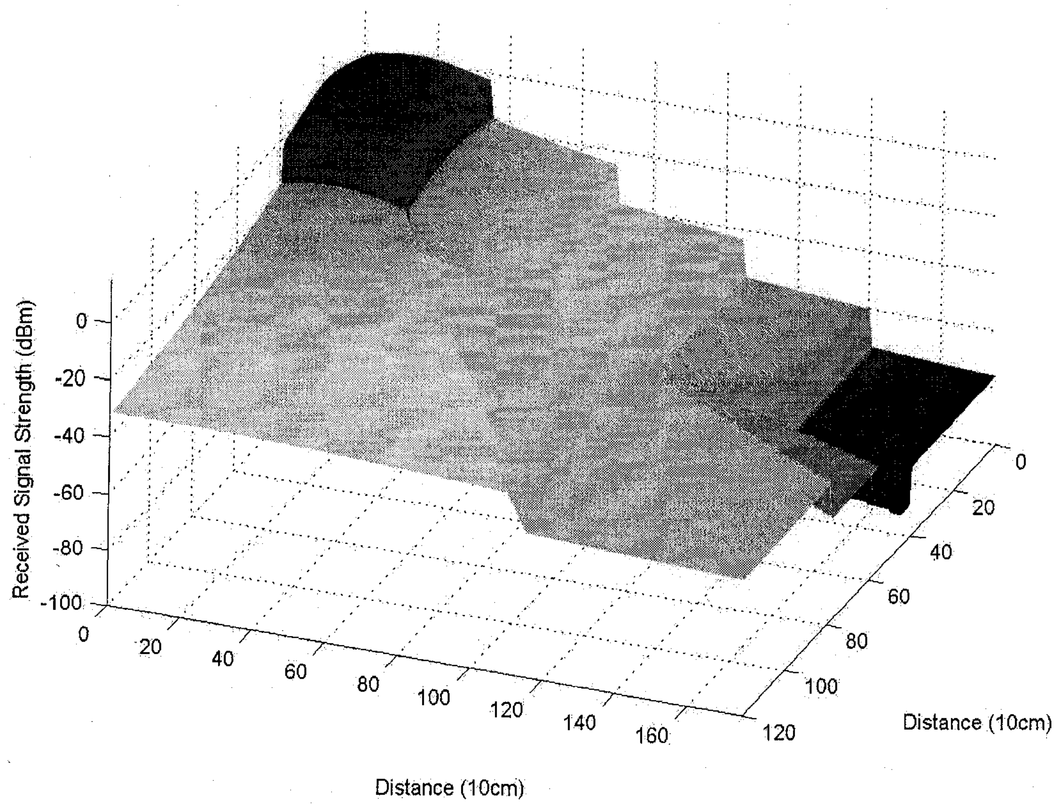
Based on the attenuation factor model, we can predict the received signal strength at a given location by predicting the path loss from the transmitter to the receiver in a specific propagation environment. In the application of indoor location technology, a signal strength map can be generated based on these predictions. The predicted signal strength map is as accurate as the field-measured one, but it is more convenient and efficient in time and cost. Even though the predicted received signal strength might not be able to be used to train a neural network, we still can use propagation models to qualitatively and quantitatively analyze the signal propagation pattern, which will help us pick up a training data set in return.

Figure 4.2 illustrates a predicted RF signal propagation pattern based on the attenuation factor model. The selected location is lab office 5026 (12m x 17.5m) on the 5th floor of the SITE building at the university of Ottawa. Figure 4.2 (a) shows the structure of the lab. There are five small enclosed offices located in a row in the lab, and the rest area is a big open area. An access point is located in the first small office, and its transmitting power is 100mw or 20dBm. Since the small offices have concrete walls, according to table 4.3, the PAF is set to 13dB. And according to Table 4.2, the path loss exponent is 3.0. The Matlab® simulation result is shown in Figure 4.2 (b).





(a) Office Structure of SITE 5026 in University of Ottawa



(b) Predicted RF Propagation Pattern

Figure 4.2 Propagation Pattern Prediction Based on the Attenuation Factor Model

Empirical measurements of path loss usually include the effect of path loss, shadowing, and multipath. In order to remove multipath effects, empirical measurements for path loss typically average their received power measurements and the corresponding path loss at a given distance over several wavelengths. This average path loss is called the local mean attenuation (LMA) at distance *d*. The benefits of using LMA are the elimination of uncertainty in the measurement and making the propagation models more applicable to various environments. However, to utilize the received signal strength in the location system, techniques must be taken to deal with the multipath effects in the real-time field measurements. The following sections we will give a detailed analysis of multipath effects or small-scale fading in indoor environments.

4.2.3 Small-Scale Fading – Multipath

Small-scale fading is a characteristic of radio propagation at the receiver. It is the result of the presence of reflectors and scatters that cause multiple versions of the transmitted signal to arrive at the receiver. Each version has a distorted amplitude, phase, and angle of arrival. All of these multiple versions of signal waves, called multipath waves, combine at the receiver antenna and result in a signal with random variation in amplitude, phase, and even in frequency. Because of the random characteristics of the multiple waves, sometimes these waves add constructively, producing a received signal with a large amplitude or RSS, while at other times they add destructively, resulting in a very low RSS. The range of RSS variation can be up to 60 to 70 dB. Figure 4.3 shows a field measurement of small-scale fading in lab office 5026 on the 5th floor of SITE building on November 29, 2005. The figure shows 2500 seconds of signal strength variation at a fixed position.

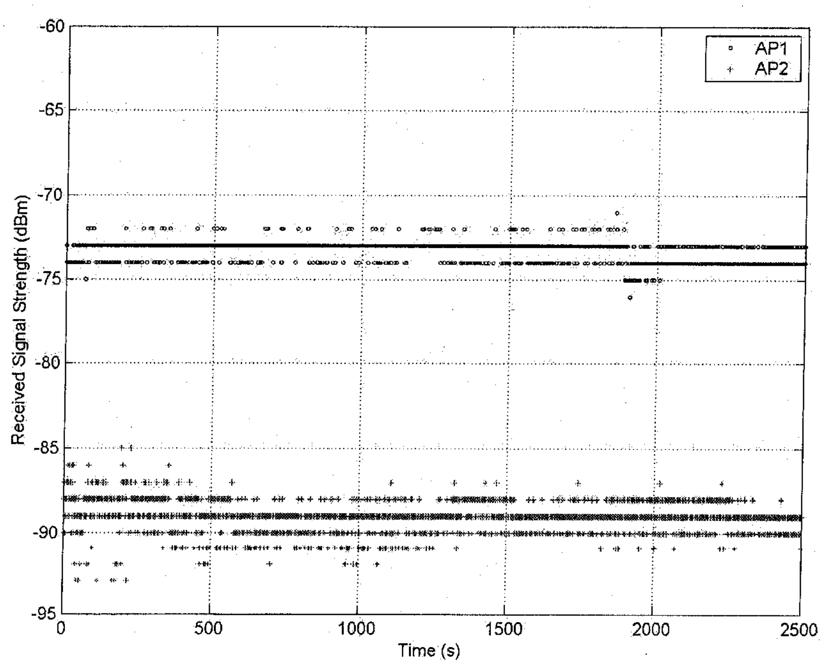


Figure 4.3 Small-Scale Fading on Received Signal Strength at Fixed Location

Small-scale fading has the following three major effects. First, it causes rapid change in RSS over a small travel distance or time interval. This effect is very important to indoor location systems. Because of small-scale fading, the real-time RSS detected by the receiver varies greatly. Methods have to be taken to eliminate the variation in the RSS samples before the application of location methods, or the performance of location system will be compromised. The second and third effects are random frequency modulation due to varying Doppler shifts and time dispersions due to multipath propagation delays. Since frequency shifts and signal delay measurements are not supported in the WLAN standard, information about frequency and signal delay cannot be utilized in our WLAN location systems. In the following sections, we will focus on the small-scale fading on signal strength.

There are many physical factors in the radio propagation channel that can influence small-scale fading. The most important factors include the speed of the mobile device, the speed of the surrounding objects, the transmission bandwidth of the signal, and the multipath propagation [53]. The movement of the mobile device and the surrounding objects in the propagation channel results in random Doppler shifts on multipath waves. If the objects in the propagation channel are static and the motion is considered to be due only to that mobile device, then fading is purely a spatial phenomenon. However, as Figure 4.3 showed, due to the dynamic nature of the indoor environment, small-scale fading in indoor environments usually varies with time. The bandwidth of the channel can be quantified by the coherence bandwidth, which is related to the specific multipath structure of the channel. If the transmitted signal bandwidth signal is greater than the coherence bandwidth of the multipath channel, the small-scale fading will not be significant even the received signal will be distorted. On the other hand, if the transmitted signal has a narrow bandwidth compared to the channel, the amplitude of the signal will change rapidly, but the signal will not be distorted in time. Multipath propagation is very normal in indoor environments or dense urban areas because these propagation environments are full of variance objects, obstacles in various materials and motions. Multipath propagation introduces small-scale fading, signal distortion, or both. As a result, the statistics of small-scale signal strength over small distances is related to the specific amplitudes, the delays of the multipath channel, and the bandwidth of the transmitted signal. In Figure 4.3, we can see that, even at the same location, the small-scale effect from different access points varies. In this case, the RSSI detected from the AP2 experience more small-scale effects than AP1.

Many researchers have worked in the area of studying small-scale fading in different propagation channels [53] [56] [57] [58]. One type of method to evaluate multipath propagation in a channel is using special hardware to measure the parameters of the channel, such as time dispersion parameters, coherence bandwidth, Doppler spread, and coherence time. A number of wideband channel sounding techniques have been developed for this purpose, including direct pulse measurements, spread spectrum sliding correlator measurements, and swept frequency measurements. The other type of method is using theoretical propagation modeling to analyse the propagation channel. Due to its random unpredictable nature, small-scale fading is always studied as a stochastic process. Numerous researchers have focused on models of measuring and analyzing the first-order statistics of these processes, which involves the characterization of small-scale fading with a probability density function (PDF), or the autocorrelation statistics of fading processes. The second order of statistics such as power spectral density (PSD), level-crossing rate, and average fade duration, have also been studied. However, in our case, what we can access in the WLAN infrastructure and what the location system needs are only the small-scale effects on RSS. In the next section, we will discuss in detail two major multipath fadings, Rayleigh fading and

Rician fading, which have been proven effectively describe the distribution of RSS in the multipath propagation channel.

4.2.4 Rayleigh and Rician Fading

Rician distributions and Rayleigh distributions are used to model small-scale fading introduced by scattered or multipath signals at the receiver. Depending on the density of the scatter, the signal will display different fading characteristics. Rayleigh distributions are used to model dense scatters without an LOS component, while Rician distributions model small-scale fading with a stronger LOS component. As shown in Figure 4.4, the small-scale fading at receiver 1 follows Rician distribution, and small-scale fading at receiver 2 follows Rayleigh distribution.

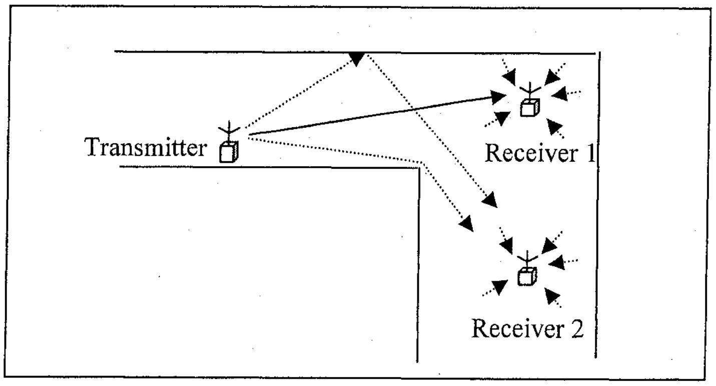


Figure 4.4 Scattering With and Without LOS

Rayleigh fading is a reasonable model when there are many objects in the environment that scatter the radio signal before it arrives at the receiver. It is well known that the envelope of the sum of two quadrature Gaussian noise signals follows a Rayleigh distribution. When multipath waves arrive at the receiver's antenna, the in-phase and the quadrature-phase component of the total received signals can be treated as two random variables. According to the central limit theorem, if there is sufficient scatter, the channel impulse response will be well-modeled as a Gaussian process, irrespective of the distribution of the individual components. This makes the in-phase and quadrature-phase component independent Gaussian random process. If there is no dominant component to the scatter, then such a process will have zero mean and with phase evenly distributed between 0 and 2π radians. The envelope of the channel response will therefore be Rayleigh distributed. Supposing that the envelope variable is r, the Rayleigh probability density function is given by

$$p(r) = \begin{cases} \dfrac{r}{\sigma^2}\exp\left(-\dfrac{r^2}{2\sigma^2}\right) & (0 \le r \le \infty) \\ 0 \quad (r < 0) \end{cases} \tag{4.10}$$

where σ is the rms value of the received voltage signal, and σ^2 is the time average power of the received signal. The mean value r_{mean} and variance of the Rayleigh distribution σ_r^2 which represents the AC power in the signal envelope, and the median valued r_{median} is given by

$$r_{mean} = E[4] = \sigma\sqrt{\frac{\pi}{2}} = 1.2533\sigma \tag{4.11}$$

$$\sigma_r^2 = E[r^2] - E^2[r] = \sigma^2(2 - \pi/2) = 0.4292\sigma \tag{4.12}$$

$$r_{median} = 1.177\sigma \tag{4.13}$$

The instantaneously received power P, with $P = r^2/2$, which is averaged over one RF-cycle, has the following exponential probability density function:

$$p_P(P) = p(\sqrt{2r})\left|\frac{dP}{dr}\right| = \frac{1}{\sigma^2}\exp\left(-\frac{P}{\sigma^2}\right) \tag{4.14}$$

Simulations have shown that the Rayleigh probability density function appropriately describes the fading of the amplitude if the number of multipath waves is larger than six.

Measurements over non-line-of-sight paths at UHF frequencies in urban environments confirm the accuracy of the Rayleigh probability density function.

Rician fading distribution is used to describe the situation that there is a dominant stationary signal component present besides the random multipath waves at the receiver. In such a case, the random multipath components are superimposed on the dominant signal. As a result, a dc component is added to the random multipath signal envelope. The Rician distribution is indicated

$$p(r)=\begin{cases}\dfrac{r}{\sigma^2}\exp\left(-\dfrac{r^2+A^2}{2\sigma^2}\right)I_o\left(\dfrac{Ar}{\sigma^2}\right) & (A\geq 0, r\geq 0)\\ 0 \quad (r<0) & \end{cases} \qquad (4.15)$$

where A is the peak amplitude of the dominant signal and $I_o(\ )$ is the modified Bessel function of the first kind and zero-order. A modified form of Rician distribution with a parameter K which is defined as the ratio between the deterministic signal power and the variance of the multipath is more commonly used. K is called the Rice factor, and is indicated by $K = A^2/2\sigma^2$ or the form of dB:

$$K(dB) = 20\log A - 20\log\sigma - 10\log 2 \qquad (4.16)$$

The Rice factor K in the Rician distribution is used to characterize the ratio of the LOS components and the NLOS components. When the Rice factor K is zero, the Rician distribution reduces to Rayleigh distribution. In other words, Rayleigh fading distribution is a special case of Rician fading distribution. Expressed in terms of the local-mean power and the Rice factor K, the probability density function of the signal power becomes

$$p_P(P)=\frac{(1+K)e^{-K}}{\overline{P}}\exp\left(-\frac{(1+K)P^2}{\overline{P}}\right)I_0\left(\sqrt{\frac{4K(1+K)}{\overline{P}}P}\right) \qquad (4.17)$$

where $\overline{P}$ is the total local-mean power, which is the sum of the power in the LOS and the local-mean scattered power and $\overline{P} = (K+1)\sigma^2$. Figures 4.4 and 4.5 show the probability density function of a signal envelope in Rayleigh fading and the probability density function of the received power in Rician fading with variant K.

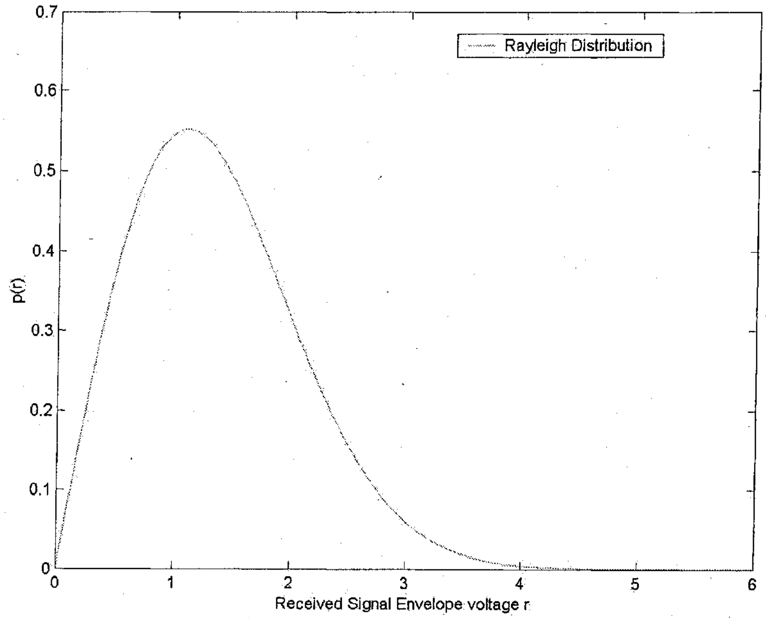


Figure 4.5 Rayleigh Probability Density Function (with σ =1)

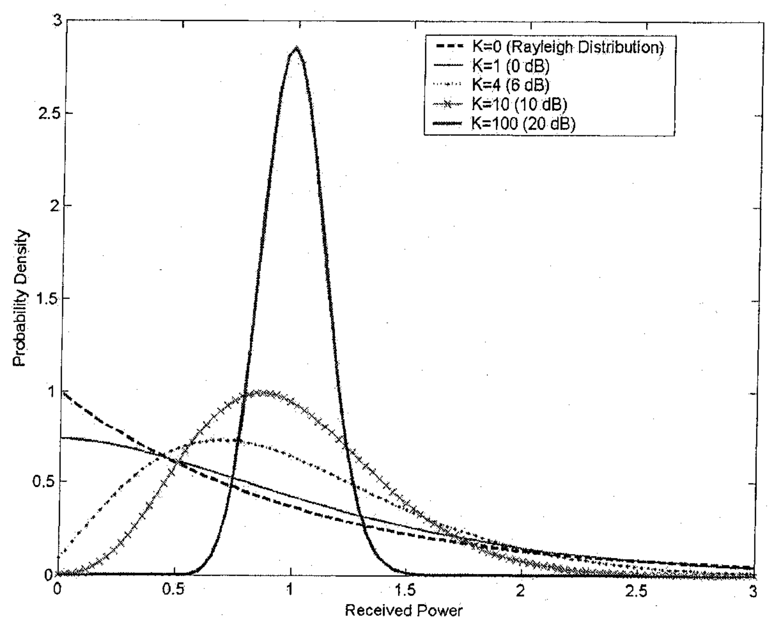


Figure 4.6 Rician Probability Density Function for Received Signal Power (with σ =1)

From Figure 4.6, we can see that the RSS probability density function of Rician fading is significantly affected by Rice factor K. When K is small, it means the LOS component at the receiver is not significant. As a result, Rician fading is approaching to Rayleigh fading (where $K = 0$), and the probability density of RSS decreases exponentially. As K becomes larger, the effect of the LOS component increases, and when K=20dB, the LOS component at the receiver becomes the dominant factor for RSS distribution. For the WLAN indoor propagation channel, it is very common that there is no LOS path between the AP and a mobile station, and in this case Rayleigh fading is appropriate. However, since the distance between them is very short, a dominant component could exist besides other random multipath waves. In this case, Rician fading is a better way to describe it.

One example is the two sets of RSSI shown in Figure 4.3. The signal received from AP1 is Rician fading with a larger K, and signal received from AP2 is close to Rayleigh fading or Rician fading with smaller K. The distribution of RSSI for the two access points are illustrated in Figure 4.7. As we can see, the RSSI distribution of AP1 is more concentrated to the mean median value than that of AP2. The standard deviation of AP1 is 0.43 dB and 4.20dB for AP2. The reason for the difference is that the propagation channel between AP1 and the receiver has an LOS path, which contributes significantly to the RSSI distribution of AP1. On the contrary, the propagation channel between AP2 and the receiver is blocked by several concentrated walls; the distribution of detected RSSI disperses.

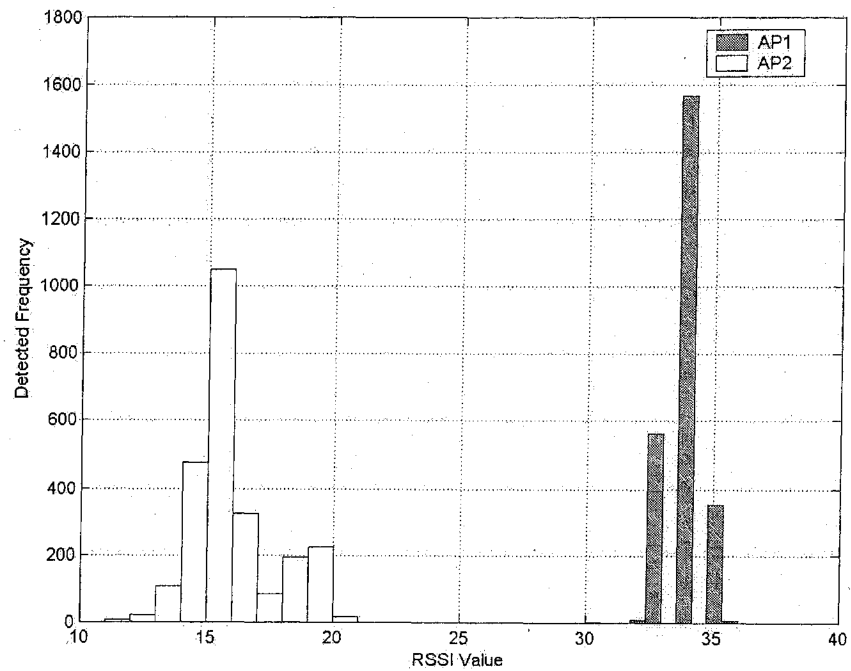


Figure 4.7 RSSI Distributions of 2 APs (2500 seconds)

Large-scale fading and small-scale fading are both important fading models, and both types of fading must be considered when analyzing the indoor propagation channels. Figure 4.7 illustrates the detected RSSI from an access point by a laptop computer while moving office SITE 5026. 4.7(b) shows the detected RSSI at eight different locations along the line in 4.7(a), and each location is 1.5 m apart. From 4.7(b), we can see that, because of the large-scale fading, while the mobile device moves away from the access point, the detected RSSI gradually decreases. Meanwhile, at each sensing location, the RSSI varies greatly because of the small-scale fading. In the application of the location system, large-scale fading analysis can help to deploy access points to maximize the effective range and build the signal propagation map. Analysis of small-scale fading helps us preprocess or calibrate the training data set and the real-time measurements. In the next chapter, a preprocess algorithm based on the analysis of small-scale fading to remove random noise is proposed.

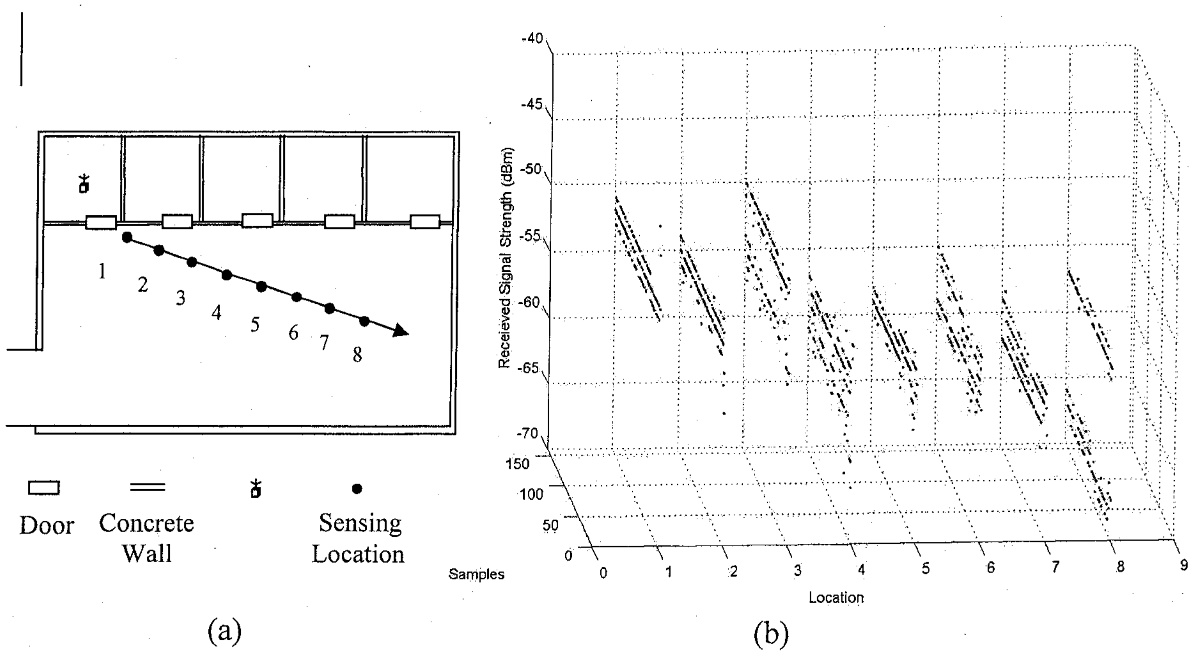


Figure 4.8 Detected RSSI along a Direct Course

4.3 Propagation Map Creation in a WLAN Environment

The propagation pattern0matching method determines a mobile device location by comparing the detected signal signature, such as RSSI in WLAN, with a pre-defined propagation pattern or propagation map. The creation of a propagation pattern is critical for WLAN indoor location systems. The performance of a WLAN indoor location system depends on the accuracy and precision of the propagation pattern and the accuracy of RSSI measurements. So it is important to understand the underlying working principle of RSSI measurements and the properties of RSSI distribution in indoor environments.

In the previous chapter, we discussed RSSI in WLAN standards, how RSSI is obtained and possible physical factors in WLAN environments that have a direct effect on RSSI. In this section, we will analysis how RSSI is distributed in an indoor WLAN environment based on the indoor RF propagation models discussed above. But before that, let us review how RSSI is generated. As we mentioned in Chapter 3, the RF energy measured by the circuitry on an IEEE 802.11 WLAN card is quantified to numeric values from 0 to RSSI_Max. And different vendors use different RSSI_Max values in their chipsets. Because RSSI is intended to be used in a relative manner and absolute accuracy is not required in the standard, each vendor made their own definition about the relationship between the RSSI value and the RF energy level detected by the circuits. As a result, individual vendors have chosen to provide their own levels of accuracy, granularity and range of actual power, and range of RSSI values.

4.3.1 Quantified RSS – RSSI Granularity

Since the RSSI value is an integer, it must increase or decrease in integer steps, and granularity in RSSI measurements must be considered when matching RSSI values to the RF energy levels. RSSI granularity depends on two factors. One is the range of energy the circuit actually measures. Second is how the detected energy levels converge to RSSI values and finally transfer to the dBm values reported to the upper-level applications. One thing to note here is that not all possible energy levels can be represented by the integer set of RSSI values.

According to the relationship between dBm and mW, when the energy level is bigger than 5mW, its dBm value changes little as the mW change greatly. For example, 200mW equal 23dBm while 100mW is 20dBm; 100mW difference means only 3dB changes in dBm value. So WLAN chip manufacturers do not measure signal strength bigger than 5mW. A typical maximum measured energy value for RSSI is -10dBm, which is 0.1mW. Any received energy bigger than 0.1mW will be treated as -10dBm, and RSSI will be the RSSI_Max value. On the other hand, the minimum level of RF energy RSSI can represent is determined by the physical limitation of the circuit on the WLAN card. The received signal energy must on a higher level than the background noise so that the circuit can extract a bit-stream. The minimum signal energy a WLAN card can detect is receive sensitivity and is a specification for each WLAN card. For example, the Intel® Pro/Wireless 2011 LAN PC Card has a receive sensitivity of -87dBm at 1Mbps. When the received signal energy is less than -87dBm, the card is not able to differentiate signal and noise. And as a result, RSSI is set to zero in such case. In general, the range of received signal energy that a WLAN card can measure is from its receive sensitivity (i.e. -87dBm) to about -10dBm.

To map the RSSI values with detected signal strength (dBm) values, different methods are taken by different vendors. Due to the logarithmic nature of the dBm measurement and the quantification of RSSI values, many WLAN chipset vendors, such as Cisco®, map the RSSI value to the dBm using tables. These mapping tables allow for adjustments to accommodate the logarithmic nature of the curve. Because the RSSI_Max value is different for each vendor, different chipset manufacturers have different mapping tables. One thing to be noted is that, because of the limited number of RSSI values, the values in the tables do not always increase in a linear manner. Some chipset manufacturers use formulas to map RSSI with dBm values. For example, Atheros® uses the following formula

$$RSS(dBm) = -(RSSI - 95) \qquad (4.18)$$

where RSSI is an integer number from 0 to RSSI_Max. Since RSSI_Max is set to 60 for Atheros chips, the possible detected dBm range for an Atheros® WLAN card is from -35dBm to -95dBm. Symbol® uses a very simple table to convert between RSSI and dBm. The following is the converting table of Symbol® [66].

Table 4.4 RSSI and dBm Converting Table of Symbol® WLAN Chip

RSSI Value	dBm Value
0-4	-100
5-8	-90
9-14	-80
15-20	-70
21-26	-60
27-60	-50

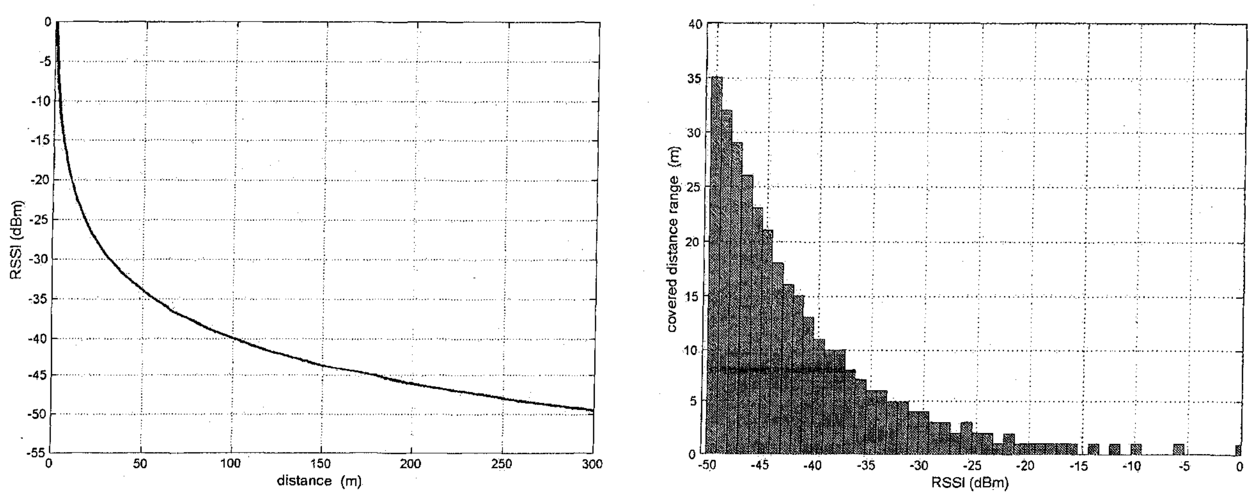
The Symbol® dBm value has a range of -50dBm to -100dBm within 10 dBm steps and only six values are given. Compared with the above simple relationship between RSSI and dBm, Cisco® provides a more accurate and complex table for its product. For a detailed converting table please refer to the appendix. Cisco® provides a dBm range from -10 to -113. But the Cisco® card usually has a receive sensitivity of -96dBm at its lowest, which is corresponds to RSSI = 16, so it is not possible to detect an RSSI value lower than 16 or -96dBm. And according to the table, RSSI values greater than 93 are assigned -10dBm. So the actual detected range for a Cisco® card is -10dBm to -96dBm.

Comparing the mapping tables between Symbol® and Cisco® we can see that the RSSI (dBm) value from Cisco has less granularity and a higher level of accuracy. The bigger value space it provides, the more accurate RSSI is. Since the performance of the location system is based only on the RSSI values, the RSSI granularity is very important for it. To increase the system's performance, higher RSSI granularity should be selected. Besides that, the selection of a training set and creation of a propagation map are also strongly affected by the RSSI granularity. A detailed analysis about the effect of RSSI granularity on the RSSI pattern is given in the following sections.

4.3.2 RSSI Pattern Analysis

4.3.2.1 Free Space Propagation Channel

Small-scale fading is random signal variance in a small space or time difference. Because of its random nature, it can hardly provide any reliable signal signatures to build the propagation map. Usually, only large-scale fading is used to build a propagation pattern map because it is much more stable with distance and time. If we ignore the small-scale fading and shadowing, according to the attenuation factor model, the distribution of RSS should be continuous over space, except in the place where an obstacle is located, which introduces a sudden drop. However, as we discussed above, because the RSSI cannot present all of the RSS values and RSSI (dBm) is converted from integer values, the RSSI propagation pattern is cut into "steps" even without any obstacles in the propagation path. For higher-level applications such as the location system here, the RSS information converted from RSSI has been quantified, and the relationship between distance and detected signal strength is not continuous.



(a) (b)

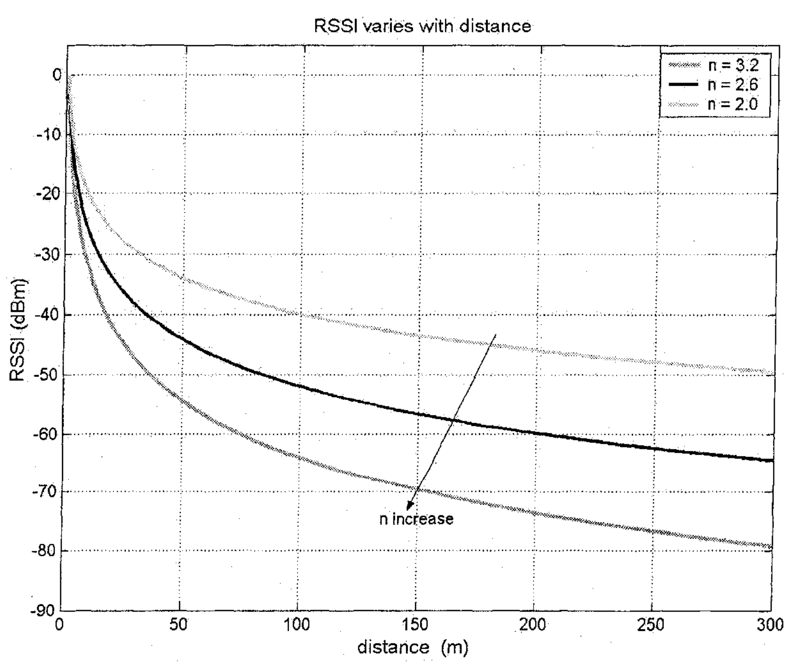
Figure 4.9 Relationship Between RSS with Distance

Figure 4.9 (a) illustrates the relationship between the predicted RSS value (dBm) and distance (meter) based on the log-distance path loss model. In the example, the reference distance is $d_0 = 1m$, the average received signal power at reference distance d_0 is $\overline{PL}(d_0) = 0dBm$, and the path loss exponent is n=2.0, which means free space propagation. The relationship between distance and RSS is not linear; the RSS changes logarithms with distance. In other words, in a free space propagation channel, each dBm change in RSS represents a different distance variance. Figure 4.9 (b) shows the relationship with each RSS dBm value changing with its corresponding distance variance. We define the RSSI resolution as the required T-R separation distance variation for RSSI (dBm) to change (increase or decrease) one dBm. The RSSI resolution represents the effect of distance on RSSI distributions. From Figure 4.9 (b), it is clear to see that, when the RSS is strong, the RSSI resolution is high, and RSSI is very sensitive to the distance changes. When the RSS is bigger than -20dBm or the distance between transmitter and receiver is less than eight meters, as the distance increases one meter, the RSS will decease more than one dBm. However, when the RSS is weak, the RSSI resolution increases, and the change in distance has less effect on the change of RSS. When the RSS is less than -40dBm or the distance is bigger than 100 meters, the RSS will not change 1 dBm until the distance changes more than 10 meters, and when the RSS is less than -50dBm, the RSS will barely change within a 35-meters change in distance. This means that if the indoor propagation channel is a free propagation channel and the RSS is less than -40 dBm, then the RSS will only have slight changes everywhere in a typical office-size room, which is a disadvantage for WLAN indoor location systems, because the RSS propagation pattern will not have significant differences in identifying locations, and the system performance will decrease. From the above discussion, we can conclude the following principle for WLAN indoor location systems.

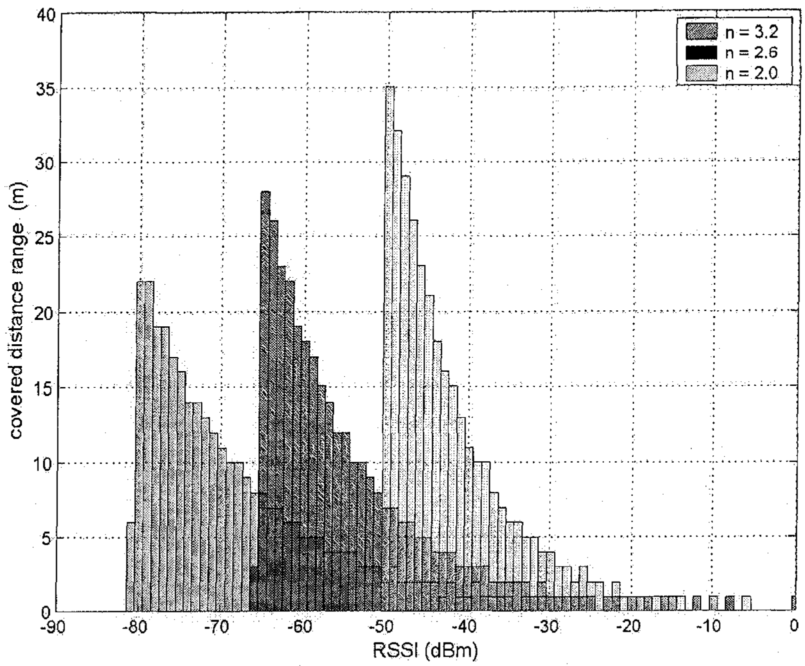
Principle a: If the effect of obstacles in a WLAN propagation channel is not considered, the precision and accuracy of RSS to represent location difference decrease as the RSS decreases.

4.3.2.2 The Effect of Path Loss Exponent *n*

Fortunately, almost all indoor propagation channels are not free-space propagation channels and signals fade much more quickly in complex indoor environments. In the log-distance path loss model, the effect of indoor structures and obstacles is expressed by different path loss exponent *n*. As *n* increases, signals fade more quickly. The following figures show the effect of path loss exponent *n* on relationship between RSSI and distance. As *n* increase, the distance needed to change the RSSI decreases from 35m/dBm to 22m/dBm.



(a)



(b)

Figure 4.10 RSSS vs. Distance with Different Path Loss Exponent *n*

The effect of path loss exponent *n* on the relationship between RSS and distance can be concluded as:

Principle b: For the same T-R separation, as the path loss exponent *n* in the WLAN propagation channel increases, the precision and accuracy of RSS to represent location differences increase.

4.3.2.3 The Effect of RSSI Granularity

One important thing need to be considered when analyzing RSSI distribution and RSSI resolution is RSSI granularity. As we discussed above, RSSI granularity depends on the received signal energy range that a card circuit can detect and the mapping relationship between RSSI (integer number) and the RSSI dBm values. Different WLAN chip vendors have different ways to achieve RSSI granularity. As a result, the detected RSSI pattern at the receivers will be different if the receivers have WLAN cards using different chipsets, even on the same WLAN network and connecting to the same AP. Understanding the effect of RSSI granularity on the RSSI propagation pattern is important for system hardware design and the creation of the RSSI propagation pattern.

In previous sections, we discussed three types of WLAN chips: Atheros®, Cisco®, and Symbol®. In this section, we will illustrate how the RSSI patterns differ from each other for these three and how RSSI granularity affects the RSSI pattern. Figure 4.11 illustrates the relationship between RSSI and the distance for WLAN cards from Cisco®, Atheros®, and Symbol® in the same propagation channel where path loss exponent equals 3.2 and there are no other obstacles.

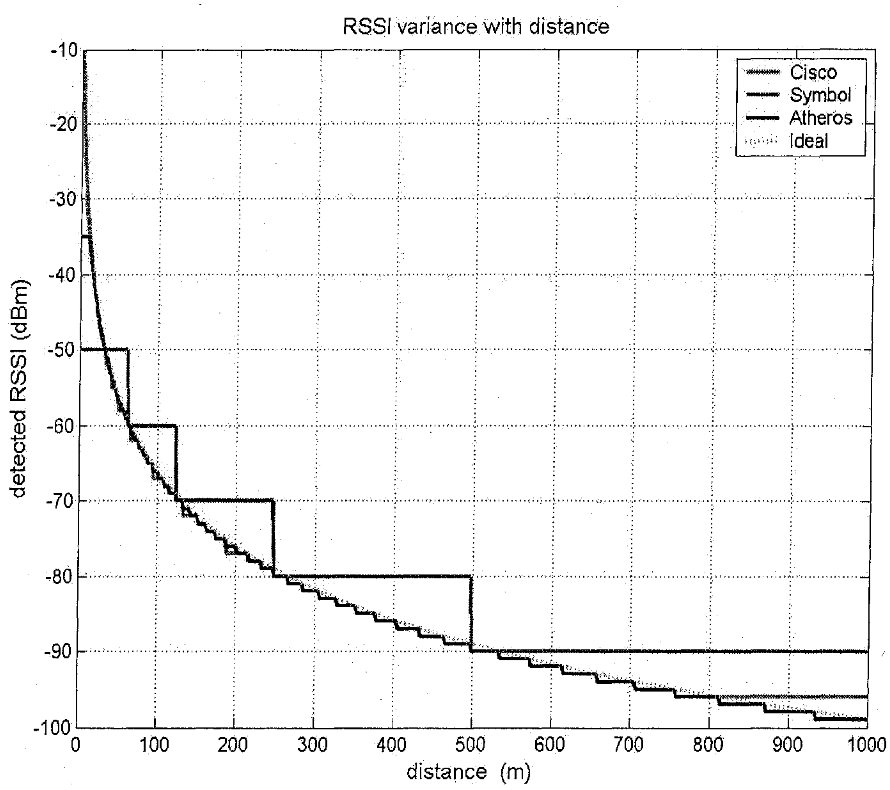


Figure 4.11 Effect of RSSI Granularity on RSSI Distribution

Because the definition of RSSI_Max is different, the RSSI granularities for the three types of cards vary greatly, especially for the Symbol® chipset. Since the Symbol® chipset uses an RSSI_Max value of 31 and a rough RSSI conversion table, it can provide only an RSSI with 6 values. Based on the log-distance path loss model, in a clear channel with n = 3.2, the minimum distance for the detected RSSI to change in a Symbol card is around 60 meters, as shown in Figure 4.12. It is obvious that coarse RSSI granularity will damage the performance of a WLAN location system.

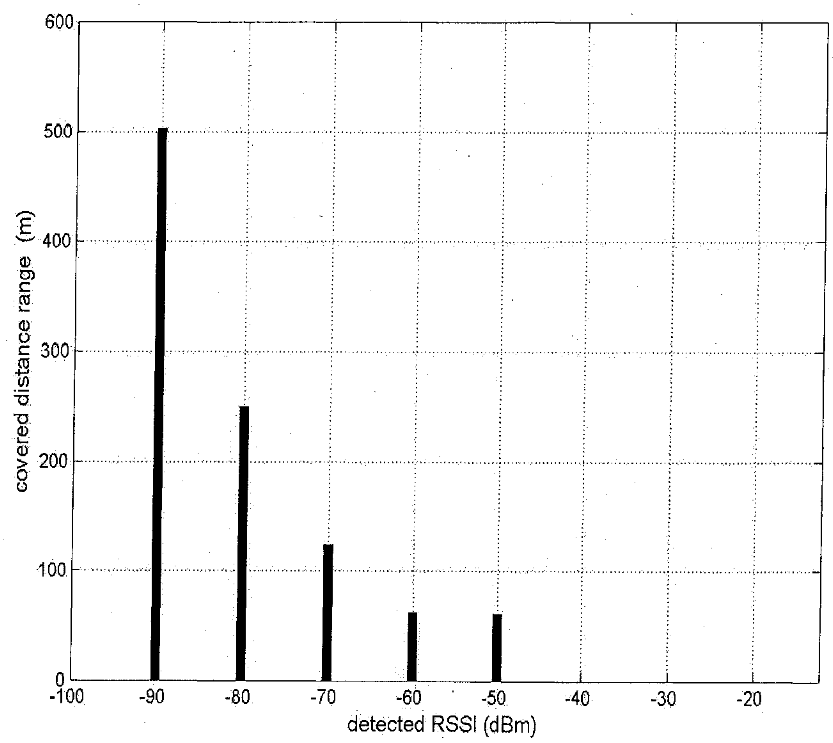
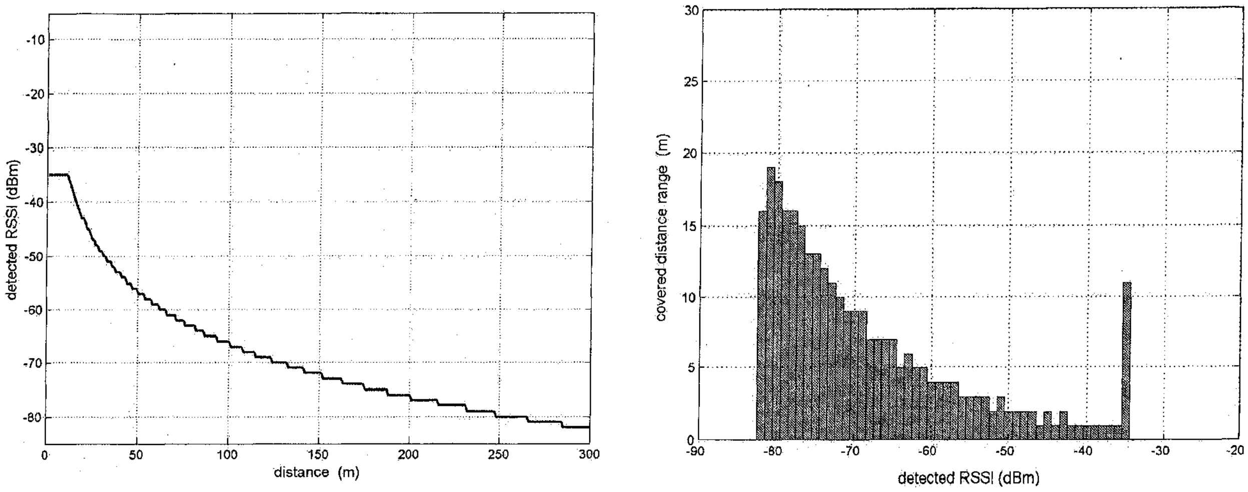
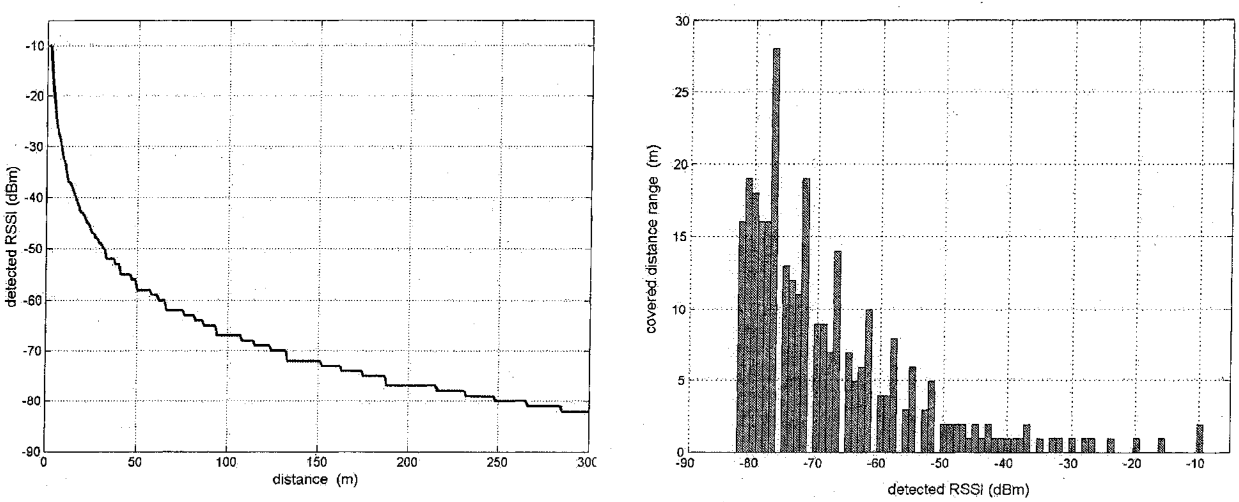


Figure 4.12 Distance Changes vs. RSSI Variation (Symbol ® WLAN chipset)

In contrast, WLAN cards with larger RSSI_Max values, such as Cisco and Atheros® chipsets, can provide much more refined RSSI granularity, and in result the variation of RSSI is much more sensitive to changes in distance between the transmitter and the receiver. Figure 4.13 illustrates the relationship between RSSI and the distance for Cisco® and Atheros® WLAN cards. The relationship is estimated based on the log-distance path loss model, given n = 3.2 and not considering the effect of obstacles located between the transmitter and the receiver.



(a) RSSI vs. Distance for Atheros® WLAN Cards



(b) RSSI vs. Distance for Cisco® WLAN Cards

Figure 4.13 Distance Changes vs. RSSI Variation (Cisco ® and Atheros® Cards)

Notice that the effect of RSSI granularity on RSSI distribution is not only from the resolution of RSSI values, but also from the measuring range of the received signal energy level, especially the highest detectable RSSI value. For cards that have a larger RSSI_Max value, such as Cisco's RSSI_Max = 100, it can measure a higher level of energy. The resolution of RSSI varies greatly around values for the Cisco® card because these RSSI dBm values are not represented in its conversion table. Compared with Cisco®, some vendors'

RSSI_Max is smaller and their chips' measuring ability in a high-energy region is reduced to keep the measuring ability at a lower-energy region. For example, Atheros® has an RSSI_Max = 60 to make sure there is a high resolution of RSSI measurement at the lower level; it cannot measure signal strength over -35 dBm. As a result, for WLAN location systems that using WLAN card with Atheros® chipsets, it will be difficult to achieve high performance in closing AP areas. According to our simulation results, in a propagation channel where $n = 3.2$, the detected RSSI for the Atheros® card will remain -35dBm if the receiver is less than about 10 meters away form the AP. Compared with high-receiving energy regions, the effect of measuring an energy range is not significant in a lower receiving energy region. Both Cisco and Atheros cards' measuring abilities are limited by their hardware. They are not able to detect signal energy lower than their receive sensitivity.

The relationship between RSSI granularity and RSSI distribution can be concluded as:

Principle c: For the same WLAN propagation channel, the precision and accuracy of RSS to represent location differences increase as RSSI granularity becomes more refined.

The above discussion about RSSI distribution does not take consideration the effect of obstacles in propagation channels, such as walls and furniture, which are very common in an indoor environment. However, by replacing the log-distance path loss model with the attenuation factor model, the effect of obstacles can be evaluated. The above three principles hold even obstacles are taken into account, because according to the attenuation factor model, the obstacles in the propagation channel only add constant factors to the path loss; they do not change the logarithm relationship between RSS and distance and have no effect on the detecting mechanism of a WLAN card.

4.3.3 RSSI Propagation Map Creation Rules

The performance of a WLAN location system, such as accuracy, effective range, and precision, is theoretically limited by the accuracy and precision of the signal propagation map. The signal signature associated with a location can be identified correctly if it is adequately different from other signatures. For a WLAN-based location system, the signal signature is RSSI-detected by WLAN cards. In order to obtain a signal propagation map, we need to estimate the expected signal strengths as measured by the mobile device to be localized for the various locations in a target area. There are two types of approaches to create such a map. One type of approach is using propagation prediction or cell-planning tools to estimate the signal strength over a target area with a detailed floor plan. The other type of approach is empirical approach. In such approaches, signal strength distribution is estimated based on the measured data, or training data, collected at different locations in the target area. Experimental studies suggest that the first type of approaches are not competitive against empirical models in terms of accuracy due to the insufficient precision of RF propagation models [16][22].

One critical step for the empirical method in creating a propagation map is to decide the way to collect training data or select sampling locations where training data sets are collected. There are ways to select sampling locations. Roos al. [23] used a grid method, dividing the target area into cells of the same size, e.g. 1m x 1m or 2m x 2m, and collected a sufficient number of RSSI samples at each cell. Another way to collect data is measuring while walking. In this way, only one data set is collected at each location, and lots of locations are involved. To form the propagation map, those locations are grouped into clusters [21]. Each cluster consists of a sufficient number of locations to supply a good estimation of average RSSI over its area. The experiment results in [16] [23] [21] have shown that the above methods can give acceptable outcomes. However, the above methods failed to take into account the effect of RSSI distribution in the process of creating of the map, and all the locations in the target area are treated equally. But in fact, as we discussed in previous sections, the RSSI distributions varies at different locations. And the variation in RSSI distribution determines the variation of accuracy in the propagation map. So it is necessary to analyze the relationship between RSSI distribution and accuracy and the precision of the propagation map.

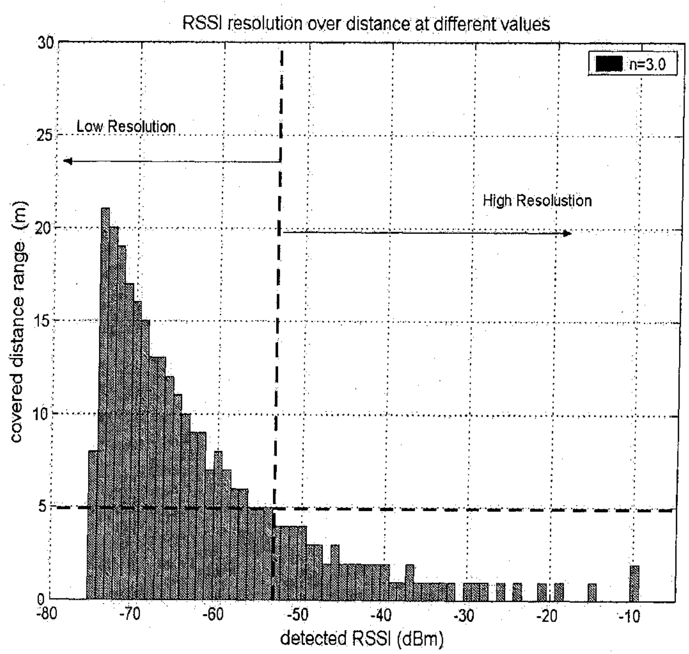


Figure 4.14 RSSI Resolution Variation (n=3.0)

Theoretically, the resolution of RSSI detected by a WLAN card over space is determined by its distance to the receiver, the local propagation environment, and the RSSI granularity of the WLAN card. For a given target area and location system, the local propagation environment and RSSI granularity are fixed. The RSSI resolution over space is a function of T-R separation distance. According to Table 4.1, a typical office environment with hard partitions has a path loss exponential equal to 3.0. If the WLAN card has the most refined RSSI granularity, such as the Cisco® or Athenros® card, then the RSSI resolution over space can be predicted by the log-distance path loss model, and the result is illustrated in Figure 4.14.

The simulation result in Figure 4.14 shows that the RSSI resolution over its detecting range varies greatly. At a high-energy level, for example over -40dBm, the resolution is about 1 to 2 meter. On the other hand, when RSSI is lower than -70dBm, the resolution is bigger than 15 meters. Obviously, there is no need to increase the data collection density to one meter over these areas, because there are no significant signal variations among the adjacent sampling points to create location signatures in the propagation map. Meanwhile, it is appropriate to increase the data collection density at areas with a high RSSI resolution. To simplify, we can divide the RSSI distribution area into two categories: a high resolution area and a low resolution area. The separation of these two areas is based on resolution threshold m meters. For example, we can select m=5 meters. According to the attenuation factor model and the simulation results in Figure 4.14, in a propagation channel with n=3.0, higher resolution areas are areas with RSSI larger than -53dBm. When the RSSI is less than -53dBm, the corresponding location falls into low resolution category. Then the coverage area of an AP can be represented by the circles and circle bands in Figure 4.15 (a). The inner circles around AP form a high resolution area, and the outer circle band forms a low resolution area. These two areas are separated by the circle of RSSI = -53dBm. Note, that -53dBm here is just a predicted RSSI threshold value for a typical office environment (with n=3.0) based on the attenuation factor model and the assumption that other random features in RSSI, such as small-scale fading or shadowing, have been removed from the signals. The number is the threshold by which the RSSI resolution over space is bigger than 5 meters. It varies as the requirements or the environment change.

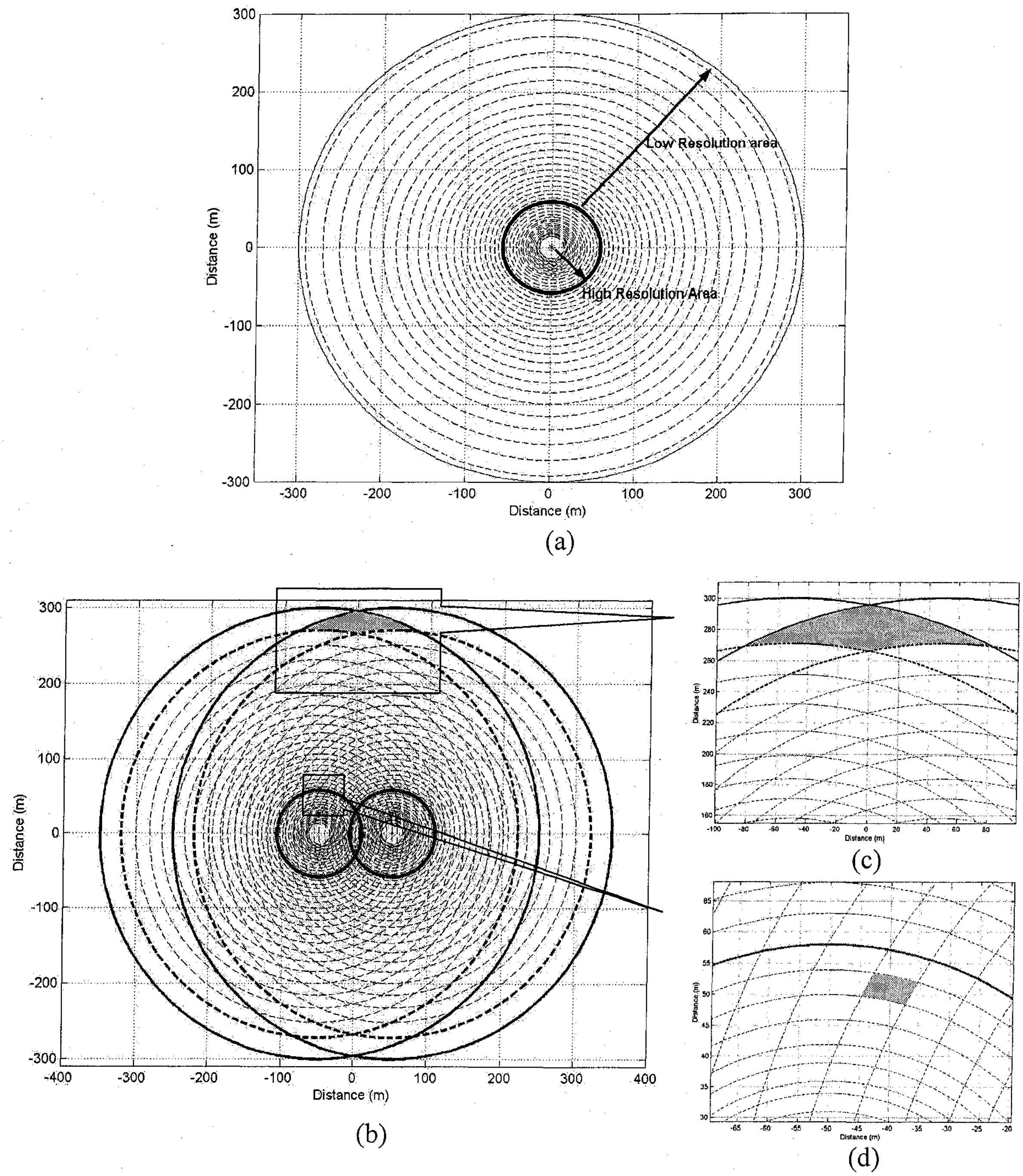


Figure 4.15 RSSI Distributions of One and Two APs

The resolution of a propagation map is determined by the resolution of RSSI distribution over the target area. For areas with low RSSI resolution, the signal signatures for adjacent locations in the propagation map will have very small distances, and it will be difficult for any location calculation method to identify the desired location from its adjacent locations, thus reducing the performance of the whole location system. We define the areas with the same signal signature in the propagation map as "uncertain areas". Clearly, the resolution of a propagation area at a specific location is determined to be the size of the uncertain area around it. Usually, in the WLAN location system, a signal signature in a propagation map is a vector composed by the RSSI values from several APs. The resolution of this signature is determined by the highest RSSI resolution of those APs at the specific location.

Figure 4.15 (b) illustrates the propagation map over an open area with $n=3.0$. There are two APs in these areas, and they are 100 meters away from each other, and the RSS at 1 meter away from each AP is 0dBm. As shown in the figure, the coverage area of these two APs overlaps. The area where the receiver can detect both APs' RSSI forms the effective coverage area of the propagation map. And the resolution of the propagation map varies greatly at different locations. The gray area in Figure 4.15 (c) is an uncertain area with a RSSI signature {-74dBm, -74dBm}, which means that the detected RSSI from both APs within the gray area will be {-74dBm, -74dBm}. The uncertain area is covered by the low RSSI resolution area of both APs, and the RSSI resolution of both APs over there is bigger than 25 meters/dBm. As a result, the uncertain area becomes a rhombus area over 100 meters long and 20 meters wide. In comparison with this, Figure 4.15(d) illustrates an uncertain area within the high resolution area of AP1, and the signal signature over this area is {-52dBm, -60dBm}. Even though it is located in the low resolution area of AP2, the high RSSI resolution from AP2 makes the uncertain area only a 4m x 8m rectangular area. In other words, the map resolution around this area is about 4 meters, because there is enough signal signature difference to identify different locations for every 4 meters distance.

Obviously, the resolution of the above propagation map can hardly satisfy the requirements of an indoor location system. The most efficient way to increase the resolution is to introduce more signal signatures into the map. In other words, the best way to increase the performance of a WLAN location system is to add more APs to the target area. The RSSI distribution of the new AP will cut the above uncertain area into small parts and thus dramatically increase the resolution of whole propagation map. Notice that, just as it is difficult to locate a position in GPS system with only two satellites, the coverage area of two APs is usually symmetric along the direct line between them. A WLAN location system will not be able to judge which side a location belongs to based solely on the symmetric signal signatures. Adding one additional AP (not on the extension line of those two APs) will theoretically eliminate the symmetry in the propagation map, thus increasing the resolution to at least twice as high as before.

According to Figure 4.14, given path loss exponent n, the RSSI resolution decreases with the RSSI value. Because of the logarithm relationship between RSSI value and distance, we can judge the RSSI resolution over a place by measuring its RSSI value. And since the dominant factor for resolution of the propagation map is the highest RSSI resolution of the signal signature over a location, we can determine the propagation map resolution around the specific location by measuring the highest RSSI values from different APs. And the sampling locations and density can be determined in the same way.

The variation of path loss exponent n could also affect the RSSI resolution. As shown in Figure 4.10 (b), the RSSI resolution increases as n increases. Typical values of n are estimated by averaging measured data over all the locations in the target area. However, in a real indoor environment, path loss exponent n varies with different environments or locations. For example, in areas with LOS to the transmitter in a building, the typical value of n is 1.6 to 1.8; on the other hand, for areas with complex obstructions, the typical values of n change from 4 to 6. Given the same RSSI value, the RSSI resolution and propagation map resolution in an open space with LOS is smaller than those of in obstructed office without LOS. When we select the sampling location and create the propagation map, the effect of path loss exponent n or complexity of environment must be taken into consideration.

The above discussion is based on the assumption that small-scale fading and shadowing would not affect the creation of signal signatures and the propagation map. However, in the real world with complex indoor environments, the predicted RSSI resolution may vary in some ways from the field-measured one. The spatial small-scale fading introduced by reflection and diffraction from walls and furniture and the random shadowing caused by radio range irregularity can bring RSSI variations into an uncertain area. It is good for location systems; because the spatial small-scale fading usually is stable with time, but they varies with space, and thus they increase the density of signal signatures over space, thus increasing the resolution of the propagation map.

In conclusion, the performance of a WLAN indoor location system is theoretically limited by the resolution of its signal propagation map, which varies over the target area and is determined by the RSSI resolution. The resolution of propagation is determined by the RSSI resolution and the manner of selecting sampling points. Research in [62] illustrates that the size of the grid will affect the accuracy of the estimated locations, and the higher density of sampling points or smaller grid sizes do not mean better performance. The density of sampling points should vary with the RSSI resolution. High density is in higher resolution areas and low density is in lower resolution areas. As we discussed, the RSSI resolution changes with RSSI values, path loss exponent n, and the complexity of the propagation environment. Based on our analysis above, to increase the resolution of the propagation map and meanwhile decrease the efforts to create it, we propose the following map creation rules.

- Based on the RSSI value, the target area can be divided into a high RSSI resolution area with a higher RSSI value and a low RSSI resolution area with a lower RSSI value by comparing the detected RSSI values with a RSSI threshold. A theoretical threshold value for typical office environment is -53dBm.
- The resolution of overall RSSI or of a propagation map is mainly determined by the RSSI resolution with the highest RSSI value in the signal signature.
- Compared with equally grid method, the density of sampling points should increase in high RSSI resolution areas to increase propagation map resolution; in the meantime, the density of sampling points in low RSSI resolution areas should decrease to save costs.
- Considering the effect of a complex environment on the path loss exponent, the density of the sampling rate should increase in obstructed offices, and should decrease in open areas with LOS.
- Considering the RSSI variation introduced by spatial small-scale fading and shadowing, the density of the sampling rate should increase in areas with complex structures and furniture and should decrease in open area with fewer obstacles.
- Adding more AP will effectively increase the resolution of the propagation map.

The above rules apply to all the methods for RSSI sampling when creating the propagation map. For the grid method, the size of the grid should change with the resolution of the propagation pattern. The size of the grids should increase as the resolution decreases. For the clustering method, the measuring density should be higher in a high resolution area than in a lower one.

4.4 Summary

For indoor location systems using the propagation pattern-based location estimation method, the properties of the propagation pattern, such as resolution, accuracy and precision, determine the performance of all of the systems. By analyzing the properties of the process of RSSI creation, the underlying relationship between the WLAN RSSI parameter and space or the RSSI resolution over space is obtained by applying the log-distance path loss model. We find out the rules about RSSI distribution, which is a dominant factor that affects the signal signature in a propagation map. Based on these basic rules, a better way to create the propagation map or to setup the training data set for neural networks can be found.

In this chapter, we analyze the two basic types of indoor RF signal fading: large-scale fading and small-scale fading. Large-scale fading models: the log-distance path loss model and the attenuation factor model are introduced; small-scale fading models: Rayleigh fading and Rician fading are discussed; and field-measured WLAN RSSI data also shows that severe small-scale fading exists in indoor office environments, and thus measures such as data smoothing or filtering must be applied in the process of training or location computing to eliminate the effect of small-scale fading. Besides small-scale fading, the distribution of RSSI is affected by many factors. RSSI granularity and propagation environment determines the resolution of the propagation map. Based on the indoor propagation models, we analyze the effect of RSSI granularity and the path loss exponent on RSSI distribution. The concept of "RSSI resolution" is introduced, and the relation of RSSI resolution with distance, the path loss exponent and RSSI granularity is analyzed. Based on the analysis and simulation results, basic rules of creating a propagation map are proposed.

5: Neural Network-Based WLAN Indoor Location System Design

In previous chapters, we have discussed the propagation map-based location calculation method and the neural network method. In this chapter, we propose a system structure for the propagation map-based WLAN location system using neural networks. The first section will introduce issues of design for a NN-based WLAN location system. And in the second part, we present a system structure for a NN-based WLAN location system. And the third part of this chapter is focused on introducing the detailed system design and development of each part of a WLAN location system.

5.1 Issues of Neural Network-Based WLAN Location System Designs

Several issues must be considered before designing and implementing a neural network-based WLAN location system. As we discussed in Chapter 2, there are four general issues for an indoor location system: performance, cost, application, security, and privacy, as Krishnamurthy introduced in [1]. All the above issues exist in our NN-based WLAN location system. Performance is the most important aspect to be addressed in WLAN location system design. Performance metrics include accuracy, precision, delay, coverage or effective area, scalability, and capacity. In the following part of this section, we will discuss in detail the issues of design for a WLAN-based location system.

Accuracy: The accuracy of the location system is determined by several factors. The accuracy and resolution of the propagation map put a limitation on the accuracy of the location system. Obviously, the accuracy of the location system cannot be higher than the accuracy of its propagation map. The resolution of the propagation map is determined by the resolution of the RSSI distribution and the method of creating the map.

There are several ways to increase the RSSI distribution resolution. The most effective way is to increase the APs' density or add more AP into the system. Increasing the APs' density will decrease the size of an uncertain area of RSSI distribution. However, it will also decrease the coverage range of the system. Theoretically, adding one AP will divide the uncertain area into two smaller ones and increase the accuracy. However, since the highest RSSI resolution from a WLAN card is only around one meter, taking into consideration the small-scale fading and randomness contribution of the environment, the system accuracy will hardly increase in the high-resolution area by adding more APs when the accuracy already reaches around one to two meters. It can increase the accuracy only in a low-resolution area. [63] and [62] show that the highest accuracy a location system can achieve is around 1.5 meters for 25 percent of data. Experiments in [23] also proved that systems using 10 APs have the highest accuracy, of around two meters with 90% of location estimation error, while system using three APs has the highest accuracy, of three meters in the same condition. So the performance does not increase linearly with the number of APs, even when the overall accuracy increases. Besides that, increasing more APs will increase the cost and also increase communication interference, which will decrease the performance of WLAN.

Coverage Area: Coverage area refers to the effective area of a WLAN location system where it is covered by several APs and effective location estimation can be given. The size of the coverage area is different than the coverage area of a WLAN network, because an effective area of a location system should be covered by at least three APs, while WLAN only needs one. Similar to accuracy, coverage area is the direct result of AP density and the number of APs. Note that the accuracy within a coverage area usually is not the same; the edge of a coverage area has low accuracy, and an area with more AP coverage has high accuracy.

Security and Privacy: The security requirement is one important factor in determining the infrastructure of a location system. For a network-based system in which the location estimation is performed in a location server, security methods such as authentication, authorization and encryption for communication, etc., need to be considered to protect the user's credentials. For mobile-based systems in which the location estimation is done by the mobile device itself, privacy is easier to be kept. However, considering the higher-layer application in which the location information of more than one mobile user is needed, security and privacy is still an issue to be addressed. A detailed discussion of security and privacy methods in a location system is out of the range of this paper.

Cost: A WLAN-based location system does not need special hardware to be implemented. The existing infrastructure is enough to build a system. However, to achieve the specific performance requirements, some systems may need more investment in hardware, such as increasing the number of APs to increase accuracy or the coverage area. Cost is also an important factor in deciding the system infrastructure. Networked-based systems require an additional location server, which will add additional costs. In contrast, mobile-based systems can be created using solely software on a laptop, a PDA, a pocket PC, or another intelligent mobile device. However, mobile-based systems have higher requirements for the computing ability of the mobile device.

Besides the above issues, a higher-layer application is also a factor that will affect system design. Different applications have different requirements for performance, update frequency, stability, and system infrastructure. For example, some applications may need the user to share their location information, and then a network-based system will be desired. To address all of the above issues, a system designer must consider the following problems.

- Number of APs: Increasing the number of APs can increase the effective range and accuracy of the location system. However, it causes more interference in WLAN communications, and the cost increases.
- Location of APs: The principle for selecting a location for APs is to provide the biggest coverage area and meet the accuracy requirements in target areas. The best AP location for a location system is different from the best AP location for a WLAN system, because the target area should be covered by two APs' high-resolution area, and WLAN coverage only needs one.
- Sampling location and density: Given the propagation pattern, using inappropriate sampling locations or density can dramatically decrease the accuracy of the propagation map. Especially for a neural network-based system, the propagation map is compressed into the neural network, and its performance is totally determined by its training data set. Inappropriate sampling will lead to overtraining (too many samples) or not converging (not enough samples).
- Hardware selection: Different WLAN cards will have different RSSI granularity, which result in different performance of the location system.
- System infrastructure: Mobile-based systems have an advantage in security and privacy; however, network-based systems have the advantage in system management and maintenance.

Using propagation models to analyze the propagation environment of the target area is an effective and efficient way to solve the above problems. Based on the floor print, the attenuation factor model, and RSSI granularity, we can estimate the RSSI distribution over a target area for a given configuration of APs. And the sampling density and locations can also be determined by applying the propagation map creation rules. Even though the predicted RSSI distribution and resolution over the target area may not be accurate due to the inaccuracy of the propagation model, analyzing the environment using propagation models can still help decide the configurations of APs and is important for the selection of sampling locations.

5.2 System Structure

NN-based WLAN location systems include two major subsystems: the propagation map creation system and the location estimation system. The function of the former system is collecting sampling data from sampling points, which are selected based on the map creation rules introduced in Chapter 4. The function of the location estimation system is estimating the user's location according to real-time measured RSSI values and trained neural networks. The propagation map must have been created before the location estimation system can output correct locations. And the propagation map creation is usually an off-line process while the location estimation process is online.

Figure 5.1 illustrates the system structure of an NN-based WLAN location system. The two subsystems have a similar structure. The propagation map creation subsystem has three basic modules: RSSI collection, the pre-process model, and the NN training model.

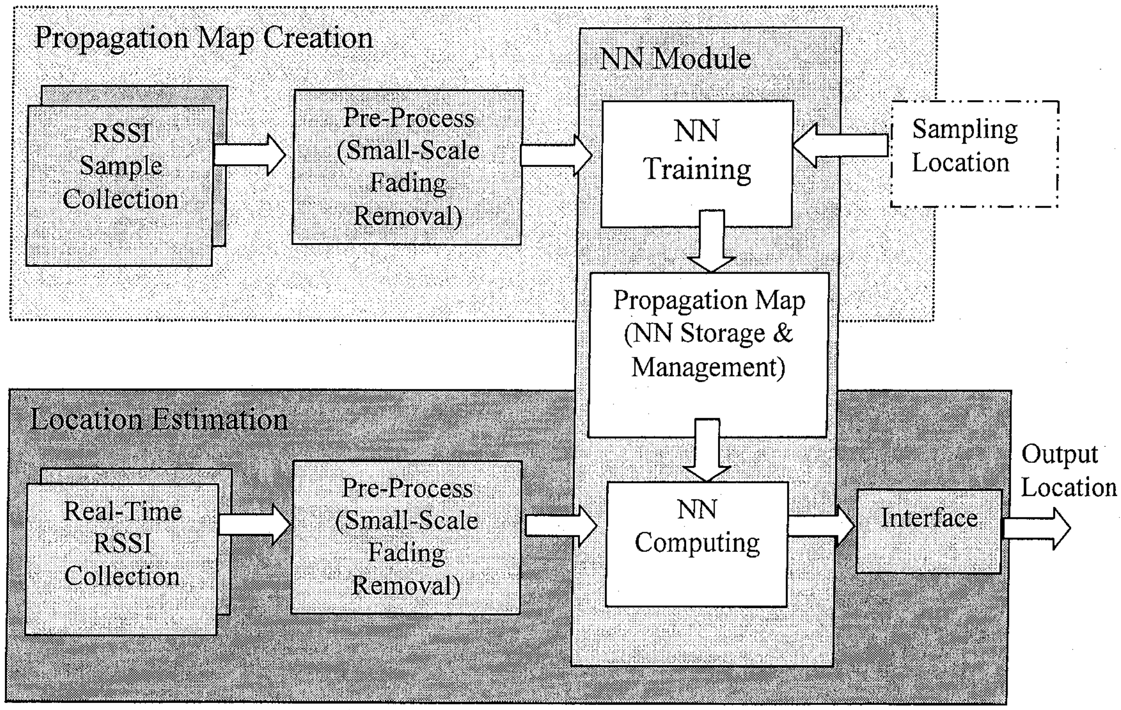


Figure 5.1 System Structure of an NN-based WLAN Location System

The RSSI collection module collects RSSI values from the WLAN card. Depending on the system infrastructure and application requirements, a location system can collect RSSI from a local machine or from other machines via a communication link. The pre-process model is necessary to remove the effect of small-scale fading on the RSSI information. By eliminating the randomness in RSSI information, the large-scale fading effect will dominate in RSSI information. Then this RSSI information is fed into an NN training model to train the neural networks. The location estimation subsystem has the same RSSI collection and pre-process modules. The difference is that the online location estimation subsystem has an NN computing module and interface instead of NN training module. NN computing computes the location, given the pre-processed RSSI and trained neural networks. The interface module transfers the output of neural network computing into location information that the upper layer application needs. Trained neural networks contain propagation maps, and the propagation map module is responsible for neural network storage and management.

For network-based systems, this module runs on the location server, and it contains measures to protect privacy and security.

5.3 System Design and Implementation

5.3.1 RSSI Sample Collection

The process of RSSI sample collection includes three basic steps: floor print or propagation environment analysis, sampling point selection, and RSSI collection. The first step analyzes the propagation environment using propagation models to predict the RF propagation pattern over the target area. The second step is selecting locations for collecting RSSI sample values based on the analysis results of the floor print and the propagation map creation rules proposed in the previous chapter. And finally, the last step is to collect RSSI values at each sampling point. Floor print analysis and selecting sampling points are not necessary for RSSI sample collection, but they are important for optimizing system performance and saving costs.

The main purpose of a floor print analysis is to predict the RSSI distribution over a target area for specific RF configurations. As discussed in the previous chapter, the number and location of APs can dramatically affect the performance of a WLAN location system. Sampling points are also defined based on the predicted propagation pattern. The key for RSSI sample collection is the selection of sampling points. For neural network-based WLAN location systems, the sample locations determine the training data set of the Neural Network. Appropriate RSSI sample collection is the guarantee of achieving the best performance of a WLAN location system. The density of sampling points is determined using the properties of the specific location, such as the highest predicted RSSI, complexity of the surrounding environment, or LOS or NLOS. Figure 5.3 shows an example of the selection of sampling points in the SITE 5026 of University of Ottawa. As shown in Figure 4.2(b), the RSSI propagation pattern for the AP is predicted based on the attenuation factor model for a typical office environment. Two types of sampling density are provided, 3m x 3m for low-resolution areas, and 1m x 1m for high-resolution areas, and the definition of the high and low resolution area is based on the propagation map creation rules. Figure 5.3(a) illustrates the high-resolution area and low-resolution area given the threshold as -53dBm.

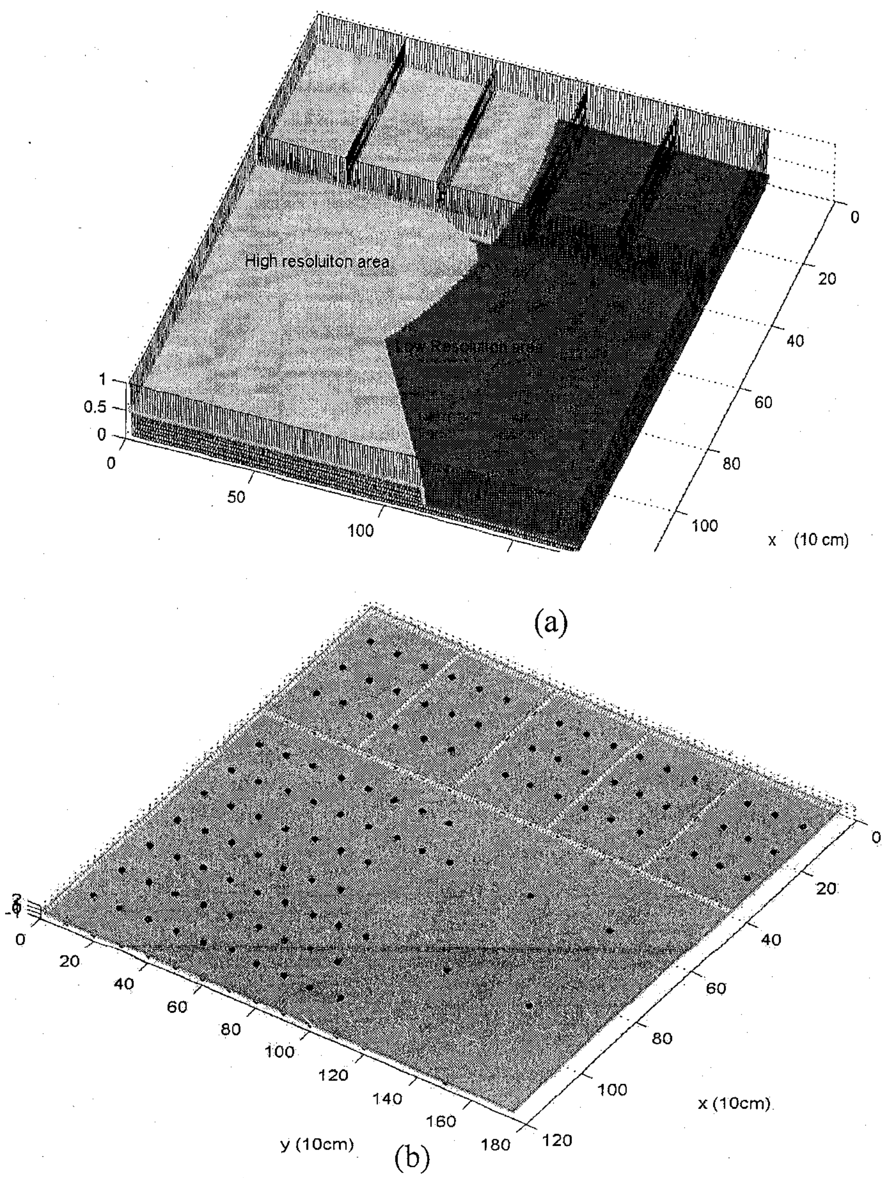


Figure 5.2 Selection of Sampling Points in SITE 5026 for One AP

Figure 5.2(b) shows the selection of sampling points based on the separation of the resolution area and map creation rules. Note that, even in some small offices, the RSSI resolution should be low, but the sampling density is still high because these small offices usually have more complex decorations and hence more complicated propagation channels. The above example is the sampling point selection for one AP. When there are multiple APs, the high-resolution area of the location system is the sum of the high-resolution area of all of the APs. And clearly, the high-density area of sampling points in the target area is the sum of the high-density area of each AP too.

5.3.2 Pre-Process Module Design

The pre-process module is the function of removing randomness in detected RSSI information. It is important because randomness in RSSI can dramatically decrease the performance of the whole location system. The pre-process module consists of filters which cancel out the small-scale fading or shadowing effects. There are two important criteria for pre-process filter design: the ability to remove randomness and the speed of doing it (or the number of required samples to remove randomness) or filter delay. These two criteria are not independent; they are tradeoffs. For the same digital filter, to increase the performance of eliminating random noise, it needs more sample windows, which will decrease the response time or the processing speed.

One of the most important parts of random effects in detected RSSI is caused by small-scale fading, which is either Rician or Rayleigh distributed, depending on the LOS conditions. Signal smoothing techniques are used to remove small-scale fading effects in the detected RSSI. The basic idea of signal smoothing is get the median value of the detected

RSSI signals over a small area and period or average the data to get the mean value. This is accomplished by adding all the signals in mW and dividing the sum by the number of samples to get the median value or by sorting all of the samples and taking the value of the sample in the middle as the median value. [64][65][60] and [48] introduced three small-scale fading eliminating methods: the mean value methods, the median value method, and the optimized method.

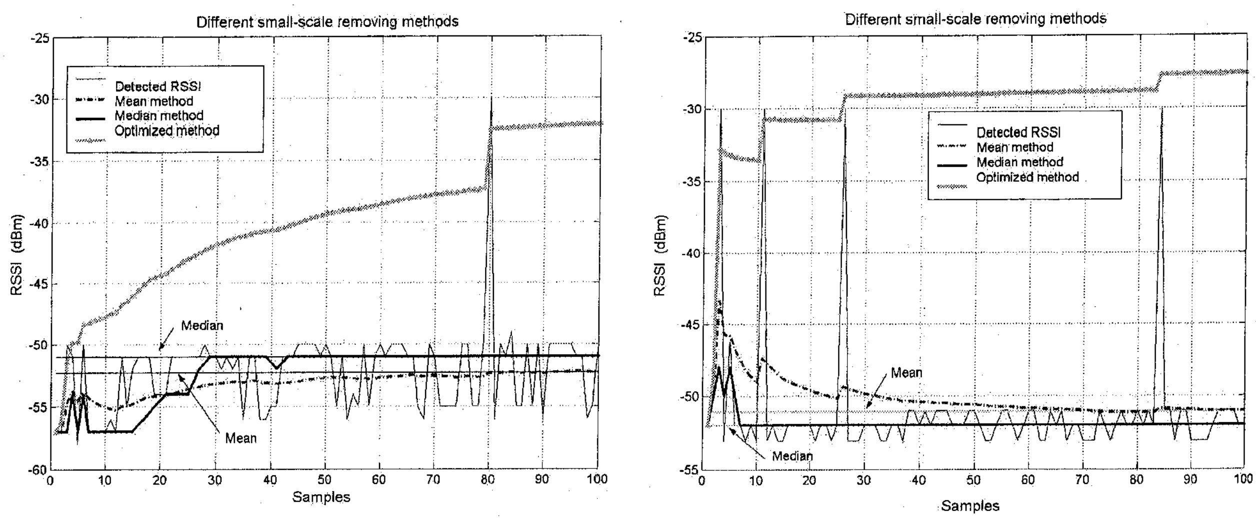
- Mean method [65] (over mW): $RSSI_{mean} = -10^{\frac{1}{N}\sum_{j=1}^{N}\log(RSSI_j)}$ (5.1)

  Mean method (over dBm): $RSSI_{mean} = mean[RSSI_j]_{j=1}^{N}$ (5.2)

- Median method [48]: $RSSI_{median} = median[RSSI_j]_{j=1}^{N}$ (5.3)

- Optimized method [60]: $RSSI_{opt} = 10\left[\log(\sum_{j=1}^{N} 10^{RSSI_j/10}) - \frac{1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{N-1}}{\ln(10)}\right]$ (5.4)

In the above formulas, $RSSI_j$ is the detected RSSI value and N is the number of samples or the size of sampling windows for the filter to produce an effective output value. Usually, the bigger N is, the closer the output is to the desired value, and the more time it needs to produce an output. It has been [65] proven that mean over the power in mW has the same effect as mean over the decibel values, and thus 5.1 and 5.2 has the same effect. The optimized method was proposed in [60] by Wong and Cox, and it has been proven that this method has a faster convergence speed than other methods in a Rayleigh fading environment. To compare the effect of small-scale removal of the above methods, a group of field-measured RSSI data was filtered by each method. The RSSI data was measured at 10 fixed locations in the open area of an office environment. All of the filtered data shows similar results. Two typical results are shown in Figure 5.3.



(a) (b)

Figure 5.3 Small-Scale Fading Removal of Detected RSSIs

Figures 5.3 (a) and (b) are the measured RSSI of two fixed locations in the open areas of SITE 5026, and these data are processed by the mean, median, and optimized methods. The optimized method has the advantages of faster converges in a Rayleigh fading environment. However, our experiment shows that it does not work well in our typical office WLAN environment, and it does not converge as the number of RSSI samples increases, probably because the fading environment is Rician. On the contrary, both the mean method and the median method can converge to the desired values, but the speed of converge or the number of samples needed by each method to reach to the desired values is different. In Figure 5.2(a), it takes 30 samples for the median method to reach the total median value of the data set, and 65 samples for the mean method to reach an overall mean value. In Figure 5.2 (b), the median method needs 10 samples, and the mean method needs 60 samples. It is clear that the median method is faster at outputting the right value than the mean method in both cases. It is because the RF signal fading channel has a deep fading structure, and mean values are more easily to be affected by those transient peak values. As we illustrated in the above figures, peak values, such as -30dBm in the detected RSSI are very common in our test environment. These peak values make the output of the mean method unstable even at fixed locations. On the other hand, the median method can provide faster converges and a more accurate output, given the same number of samples in our office propagation environment. Thus, compared with other methods, the median method has an advantage in the ability and speed of small-scale fading removal. Our field tests also find out that the number of samples for a median method to converge in our tested environment is between 10 and 30. For a WLAN location system in a different environment, a similar simulation can be performed to evaluate the effectiveness of each method and decide the sampling window size N for the desired method.

5.3.3 Neural Network Module

There are two major functions of the neural network module: one is the storage of the propagation map and the other is the estimation of locations. The process of neural network training is also the process of creating the propagation map. And the process of neural network computing is the process of extracting locations from the propagation map. As introduced in Chapter 3, the most commonly used artificial neural network is MLP and the training algorithm for MLP is backward propagation. MLP is powerful enough for any WLAN location system.

There are several issues that need to be considered when designing or implementing a WLAN location MLP. The most important one is determining the architecture of the MLP. Battiti [24][32] and Wassi [62] proved that a three-layer MLP can provide good performance for location estimation. More hidden layers lead to more complexity and a bigger computing burden. The input nodes and output nodes of an MLP are determined by the applications. The primary input of the MLP is the RSSIs of each AP at the target location. However, different applications have different requirements about the operation procedure and performance metric, which results in different metrics for the input and output nodes of an MLP. For example, as Figure 5.4 shows, in a WLAN location system in a multi-floor environment, the input to the MLP could include the floor number or other information that is not from the AP directly but can greatly improve the performance of the MLP.

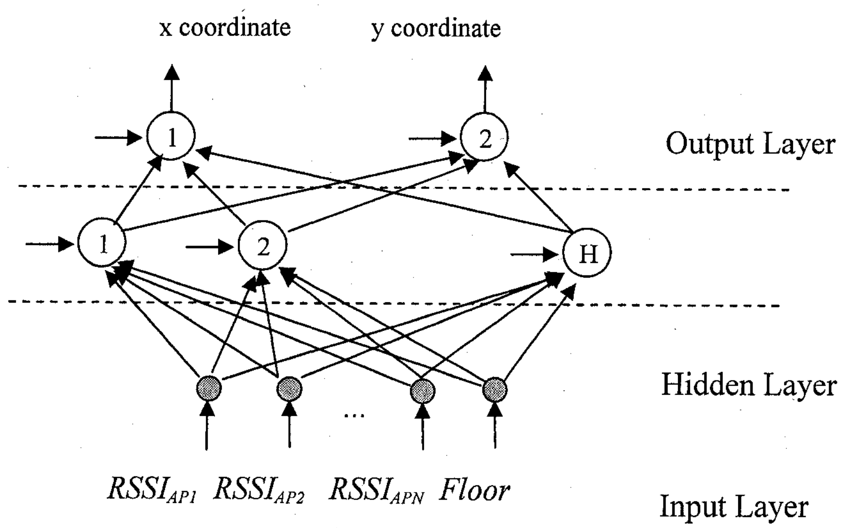


Figure 5.4 MLP Architecture for a WLAN Location System

The most intuitionistic outputs are the coordinates of the target location, but if the WLAN location system needs only to identify the room in which the device is located, the output of the MLP can be simplified to just room numbers, and thus one output node is enough in this case. Compared with input and output nodes, the number of hidden nodes is more difficult to determine. The experiment in [24] and [32] tests the performance of MLPs for different number of hidden nodes, and the architecture with 16 hidden nodes achieves the best performance. However, this architecture is limited to this case, where only three APs are applied and the target area is their specific experimental area. For other applications, where the number of APs, the propagation environment, and the performance requirements are dramatically different from the above experiment, the optimum number of hidden nodes will change. The best way to determine the number of hidden nodes is still trial-and-error. Since this process can be done online, and the architecture is determined except for a number of hidden nodes, even the trial-and-error method is very practical.

Another issue for MLP computing is the availability of RSSI information for input nodes. In an indoor WLAN propagation environment, the mobile device can not always detect all of the RSSI information from all of the APs, because in some areas, the RSS of some APs is lower than the receive sensitivity of the WLAN card. And this results in an inconstant number of inputs for the MLP and thus introduces more complexity in MLP training and computing. Since the missing RSSI is caused by low RSS, it is reasonable to set receive sensitivity as the default input value for the input node. When the input RSSI from one AP is not available, the input value is set to receive sensitivity (i.e. -87 dBm for the Intel® Pro/Wireless 2011 LAN PC Card).

5.4 Summary

In this chapter, we discuss issues of design for an NN-based WLAN location system, including accuracy, coverage area, cost, security, and privacy. Based on the analysis and simulation results and the propagation map creation rules, the effect of AP configuration, the number of APs, the creation of training data set, hardware selection, and system infrastructure selection for the above issues are discussed. We proposed a system structure of a propagation map-based WLAN location system using a neural network. Then we proposed the design of major modules in the structure: the sampling collection or creation of a training data set, the pro-process module, and the architecture of MLP. The sampling point selection is determined by the propagation map resolution. A propagation resolution-based dynamic grid method for sampling point selections is proposed. To eliminate the small-scale effects on the detected RSSI information, a different filtering method is presented. And field tests show that the median method can provide better performance. And finally, the way to determine the architecture of an MLP for location estimation is also proposed.

6 Conclusions

This thesis researches the location problems of a neural network-based WLAN location system. WLAN-based location systems make use of the RSSI information provided by the standard WLAN NIC card to estimate a mobile device's location. While empirical results and performance studies of different location estimation methods in WLAN environments have been presented in literature, all of the existing research focuses on data analysis of collected signal information, but not on the underlying relationship between WLAN RSSI distribution and system performance. In this thesis, we defined a concept called "RSSI resolution" over space for an IEEE 802.11 WLAN environment. And based on the log-distance path loss model and the WLAN RSSI granularity properties over space, we found that RSSI distribution over space is affected by the distance between the receiver and the AP, the path loss exponent, and the quantification properties of the RSSI collection circuit. The following principles of RSSI resolution are proposed:

Principle a: If the effect of the obstacles in a WLAN propagation channel is not considered, the precision and accuracy of RSS to represent location differences decrease as the RSS decreases.

Principle b: For the same T-R separation, as path loss exponent n in the WLAN propagation channel increases, the precision and accuracy of RSS to represent location differences increases.

Principle c: For the same WLAN propagation channel, the precision and accuracy of RSS to represent location differences increase as RSSI granularity becomes more refined.

Given the above principles and the simulation and analysis results, the following basic rules of creating a propagation map in a WLAN environment is proposed:

- Based on the RSSI value, the target area can be divided into a high RSSI resolution area with a higher RSSI value and a low RSSI resolution area with a lower RSSI value by comparing the detected RSSI values with a RSSI threshold. A theoretical threshold value for typical office environment is -53dBm.
- The resolution of overall RSSI or of a propagation map is mainly determined by the RSSI resolution with the highest RSSI value in the signal signature.
- Compared with equally grid method, the density of sampling points should increase in high RSSI resolution areas to increase propagation map resolution; in the meantime, the density of sampling points in low RSSI resolution areas should decrease to save costs.
- Considering the effect of a complex environment on the path loss exponent, the density of the sampling rate should increase in obstructed offices, and should decrease in open areas with LOS.
- Considering the RSSI variation introduced by spatial small-scale fading and shadowing, the density of the sampling rate should increase in areas with complex structures and furniture and should decrease in open area with fewer obstacles.
- Adding more AP will effectively increase the resolution of the propagation map.

Given the above analysis, we proposed a system structure of a propagation map-based WLAN location system using a neural network. The structure includes two phases: the off-line training phase and the online location estimating phase. There are three major modules in the structure, the RSSI collection module, the pro-process module, and the neural network module. For the RSSI sampling point selection, a propagation resolution-based dynamic grid method for sampling point selection is proposed. To eliminate the small-scale effects in the detected RSSI information, a different filtering method is presented. And field tests shows that the median method can provide better performance. And finally, the way to determine the architecture of an MLP for location estimation is also proposed.

Appendix A. RSSI conversion Table of the Cisco® WLAN card

RSSI	dBm Value	RSSI	dBm Value	RSSI	dBm Value
0	-113	34	-78	68	-41
1	-112	35	-77	69	-40
2	-111	36	-75	70	-39
3	-110	37	-74	71	-38
4	-109	38	-73	72	-37
5	-108	39	-72	73	-35
6	-107	40	-70	74	-34
7	-106	41	-69	75	-33
8	-105	42	-68	76	-32
9	-104	43	-67	77	-30
10	-103	44	-65	78	-29
11	-102	45	-64	79	-28
12	-101	46	-63	80	-27
13	-99	47	-62	81	-25
14	-98	48	-60	82	-24
15	-97	49	-59	83	-23
16	-96	50	-58	84	-22
17	-95	51	-56	85	-20
18	-94	52	-55	86	-19
19	-93	53	-53	87	-18
20	-92	54	-52	88	-17
21	-91	55	-50	89	-16
22	-90	56	-50	90	-15
23	-89	57	-49	91	-14
24	-88	58	-48	92	-13
25	-87	59	-48	93	-12
26	-86	60	-47	94	-10
27	-85	61	-46	95	-10
28	-84	62	-45	96	-10
29	-83	63	-44	97	-10
30	-82	64	-44	98	-10
31	-81	65	-43	99	-10
32	-80	66	-42	100	-10
33	-79	67	-42		

BIBLIOGRAPHY

1. P. Krishnamurthy, "Position location in mobile environments", *In Proc. NSF workshop on Context-Aware Mobile Database Management (CAMM)*, Providence, RI, Jan. 2002

2. J. A. Tauber, "Location systems for pervasive computing", *Area Exam Report,* Massachusetts Institute of Technology, Aug. 2002

3. Yiming Ji, "Location estimation in wireless networks", *Ph.D* Dissertation, Auburn University, 2005

4. www.navcen.uscg.gov/Ftp/gps/status.txt

5. Chris Savarese, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks", *Master's thesis*, Berkeley, 2002

6. A. Savvides, H. Park, and M. B. Srivastava. "The n-hop multilateration primitive for node localization problems". *Journal of Mobile Networks and Applications (MONET)*, Vol. 8, pp:443-485, 2003

7. Xiang Jim HOngyuan Zha, "Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling", *IEEE Infocom 2004*

8. Lingxuan Hu, David Evans. "Localization for mobile sensor networks", *MobiCom'04 ACM*, Sept 2004

9. Dragos Niculescu, Badri Nath, "Ad Hoc Positioning System (APS) using AoA", *IEEE Infocom San Francisco*, 2003

10. Alex Hills, Jon Schleel, Ben Jenkins. "Estimating Signal Strength in the design of an indoor wireless network", *IEEE Trans. On Wireless Communications*. Vol. 3, No.1, Jan. 2004

11. P. Krishnan, A.S. Krishnakumar, Wenhua JU, C. Mallows, S. Ganu. "a System for LEASE: System for location Estimation Assisted by Stationary Emitters for Indoor Wireless Networks", *in Proceedings of IEEE Infocom 2004,* Hong Kong, March 2004

12. P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, Ganga-Jamuna Press, MA, 2001

13. T. S. Rappaport, J. H. Reed, and B. D. Woerner, "Position location using wireless communications on highways of the future", *IEEE Commun. Mag.*, Vol. 34, No. 10, pp: 33-41, Oct. 1996

14. G. M. Djuknic and R. E. Richton, "Geolocation and assisted GPS", *IEEE Computer*, Vol. 34, No. 2, pp. 123-125, Feb. 2001

15. K. Pahlavan, X. Li, and J. P. Makela, "Indoor geolocation science and technology", *IEEE Communication Magazine*, Vol. 40, No. 2, pp. 112-118, Feb. 2002

16. T. Roos, P. Myllymaki, H. Tirri, P, "a Statistical modeling approach to location estimation", *IEEE Trans. Mobile Computing*, Vol. 1, No.1 pp:59-69 2002

17. Skellam, J.G., "Studies in statistical ecology. I. Spatial pattern", *Biometrica* 39, pp: 346-352

18. S. Saha, K. Chaudhuri, D. Sanghi, and P. Bhagwat, "Location determination of a mobile device using IEEE 802.11b access point signals", *In Proc. IEEE Wireless Communications and Networking Conference (WCNC'03)*, New Orleans, LA, pp: 1987-1992, Mar. 2003.

19. P. Prasithsangaree, P. Krishnamurthy, and P. K. Chrysanthis, "On indoor position location with wireless LANs", *In Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'02)*, Lisbon, Portugal, Sept. 2002

20. A. M. Ladd, K. E. Bekris, G. Marceau, A. Rudys, L. E. Kavraki, and D. S. Wallach, "Robotics-based location sensing using wireless Ethernet", *In Proc. ACM International Conference on Mobile Computing and Networking (MOBICOM'02)*, pp: 227-238, 2002

21. M. A. Youssef, A. Agrawala, and A. U. Shankar, "WLAN location determination via clustering and probability distributions", *In Proc. IEEE International Conference on Pervasive Computing and Communications (PerCom'03)*, Dallas-Fort Worth, TX, , pp: 23-26, Mar. 2003

22. Paramir Bahl, Venkata N. Padmanabhan, "RADAR: AN in-building RF-based user location and tracking system", *In proceedings of IEEE INFOCOM2000*, pp. 776-784, Mar, 2000

23. T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen, "A probabilistic approach to WLAN user location estimation", *International Journal of Wireless Information Networks*, Vol. 9, No. 3, pp: 155-164, July 2002

24. R. Battiti, Location-aware computing: "a Neural network model for determining location in wireless LANs", University degli Studi di Trento, Tech. Rep. DIT-0083, Feb. 2002

25. Roy Want, Andy Hopper, Veronica Falcao, Jonathan Gibbsons, "The active badge location system", *ACM Transaction on information systems*. Vol.10, No.1, pp: 91-102, Jan. 1992

26. Andy Ward, Alan Jones, Andy Hopper, "A new location technique for the active office", *IEEE Personal Communication*, Vol.4, No.5, pp:42-47, Oct. 1997

27. Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, Paul Webster, "The anatomy of a context-aware application", *in MOBICOM 1999*, pp: 59-68, Aug. 1999

28. Nissanka B. Priyantha, Anit Chakraborty, Hari Balakrishnan, "The cricket location-support system", *In MOBICOM 2000*, pp: 32-43, Aug 2000

29. Jay Werb, Colin Lanzl, "Designing a positioning system for finding things and people indoors", *IEEE Spectrum*, Vol. 35, No. 9, pp: 71-78, Sept 1998

30. P. Castro, P. Chiu, T. Kremenek, R. Muntz. "A probabilistic room location service for wireless networked environments", *In Ubicomp 2001: Ubiquitous Computing*, pp: 18-34. Springer, 2001

31. Jeffrey Hightower, Gaetano Borriello, Roy Wang, SpotON: "An indoor 3D location sensing technology based on RF signal strength", The Univ. of Washington, Technical Report: UW-CSE 2000-02-02, Feb 2000

32. R. Battiti, A. Villani, and T. Le Nhat, "Neural Network models for intelligent networks: deriving the location from signal patterns", *in Proceedings of AINS2002*, UCLA, May 2002

33. G. Cybenko, "Approximation by superpositions of a sigmoidal function", University of Illinois, Department of Electrical and Computer Engineering, Technical Report 856, 1988

34. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, Vol. 2, No. 5, pp:359-366, 1989

35. K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks", *Neural Networks*, Vol. 3, No. 5, pp: 551-560, 1990

36. Dharma Prakash Agrawal, Qing-An Zeng, *Introduction to Wireless and Mobile Systems*, Thomson Brooks/Cole, USA, 2003

37. ETSI , *HIPERLAN Functional Specification*, ETSI Draft Standard, July, 1995

38. B. P. Crow, I. Wajaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 wireless Local Area Networks", *IEEE Communications Magazine*, pp:116-126, London, UK

39. IEEE, *IEEE Standard for IT-Telecommunications and information exchange between systems LAN/MAN - Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band*, IEEE, 2003

40. Bob O'Hara, Al Petrick, *802.11 Handbook: a Designer's Companion*, IEEE Press, New York, NY, USA, 1999

41. D. B. Johnson, D.A. Maltz, "The Dynamic Source Routing Protocol in Ad Hoc Networks", *Mobile Computing*, pp: 152-181, 1996

42. IEEE, *Information Technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 1999

43. Jeffrey Hightower, Chris Vakili, Gaetano Borriello, Roy Want, "Design and Calibration of the SpotON Ad-Hoc Location Sensing System", UW CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA, Feb. 2000

44. Frank Ohrtman, Konrad Roeder, *Wi-Fi Handbook – Building 802.11b Wireless Networks*, McGraw-Hill, New York, 2003

45. Amitava Mukherjee, Somprakash Bandyopadhyay, Debashis Saha, *Location Management and Routing in Mobile Wireless Networks*, Artech House, Boston, 2003

46. Robert K. Crane, *Propagation Handbook for Wireless Communication System Design*, CRC Press, New York, 2003

47. Robert Leese, Stephen Hurley, *Methods and Algorithms for Radio Channel Assignment*, Oxford University Press, London, 2002

48. T. Kurt, Y. L. Helloco, B. Breton, "Performance comparison of small-scale fading elimination methods", *Proceedings of the 8th ACM international symposium on Modelling, analysis and simulation of wireless and mobile system*, pp:219-222, Montreal, Canada 2005

49. Yu Hen Hu, Jenq-Neng Hwang, *Handbook of neural network signal processing*, CRC Press, 2002

50. Ramakrishna Janaswamy, *Radiowave propagation and smart antennas for wireless communications*, Kluwer Academic Publishers, 2001

51. Jogen Bach Andersen, Theodore S. Rappaport, Susumu Yoshida, "Propagation Measurements and Models for wireless communications channels", *IEEE Communications Magazine*, pp42-45, Jan 1995

52. Dan Dobkin, "Indoor propagation issues for wireless LANs", *RF Design*, pp:40-45, Sept 2002

53. T. S. Rappaport, Wireless *Communications: Principles and Practice*, 2nd edition, Prentice Hall, 2002

54. S. Y. Seidel, T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communication in multifloored buildings", *IEEE Transactions on antennas and propagation*, Vol. 40, No. 2, pp:207-217, Feb 1992

55. R. K. Morrow, T. S. Rappaport, "Getting In", *Wireless Review*, pp:42-44, Mar 2000

56. R. Katternbach, "Statistical modeling of small-scale fading in directional radio channels", *IEEE Journal on Selected Areas in Communications*, Vol. 20, No. 3, Apr 2002

57. S. Haykin, M. Moher, *Modern wireless communication*, ISBN 0-12-124697-6, Prentice Hall, 2005

58. H. Bertoni, *Radio propagation form modern wireless systems,* Prentice Hall, 2000

59. W.C.Y. Lee, "Estimate of local average power of a mobile radio signal", *IEEE Trans. Vehicle Technology*, Vol.34, pp.22-27, Feb, 1985

60. D. Wang, D. C. Cox, "Estimating local mean signal power level in a Rayleigh fading environment", *IEEE Trans. Vehicular Technology*, Vol. 48, NO. 3, May 1999

61. http://people.deas.harvard.edu/~jones/es151/prop_models/propagation.html

62. G. I. Wassi, C. Despins, D. Grenier, C. Nerguizian, "Indoor location using received signal strength of IEEE 802.11b access point", *Electrical and Computer Engineering, 2005 Canadian Conference on*, pp: 1367-1370, 1-4 May 2005

63. Eiman Elnahrawy, Xiaoyan Li, Richard P. Martin, "The limits of Localization using signal strength: A comparative study", *In Proceedings of the first IEEE international conference on Sensor and Ad Hoc Communications and Networks (SECON 2004)*, Santa Clara CA, Oct 2004.

64. A. J. Goldsmith, L. J. Greenstein, G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing", *IEEE Trans. Vehicular Technology*, Vol. 43, No. 1, pp: 439-446, Aug, 1994

65. W. C. Y. Lee, "Estimate of local average power of a mobile radio signal", *IEEE Trans. Vehicular Technology*, Vol. 34, pp:22-27, Feb. 1985

66. J. Bardwell, "Converting Signal Strength Percentage to dBm Values", White Paper, WildPackets Company, Nov. 2002