Indoor Wireless Local Area Networks (WLAN) Location System Using a Kalman Filter

by

Lin Zheng

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

Master of Science

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

Copyright © 2009 by Lin Zheng

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES
*****
COPYRIGHT PERMISSION

Indoor Wireless Local Area Networks (WLAN) Location System Using a Kalman Filter

By

Lin Zheng

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

Lin Zheng©2009

Permission has been granted to the University of Manitoba Libraries to lend a copy of this thesis/practicum, to Library and Archives Canada (LAC) to lend a copy of this thesis/practicum, and to LAC's agent (UMI/ProQuest) to microfilm, sell copies and to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

Abstract

There are many location determination techniques that are used to determine a mobile user's position. These include the globe positioning system (GPS), infrared (IR) based systems, radio frequency (RF) based systems and ultra wideband (UWB) based systems. Considering the indoor environments, RF based systems, especially those based on IEEE 802.11, have garnered considerable attention because these systems have an increasingly ubiquitous coverage area and they are able to work with the existing infrastructure without requiring additional location determination hardware. The most challenging issues concerning wireless local area network (WLAN) based location systems are multipath fading and interference caused by variations in the indoor propagation environment as well as interference from other devices working on the same unlicensed frequency band at 2.45GHz. This thesis addresses the issue of indoor location estimation using an existing WLAN infrastructure. The location detection technique investigated here builds a signal radio map in the test area using the received signal strength index (RSSI). To alleviate the effect of multipath fading and interference, the Kalman filter is adapted in this location determination system, which effectively filters out various noise sources in the test environment. The Kalman filter is also investigated in combination with an accelerometer for improving the accuracy when tracking mobile users. Experimental results using real data collected from an office environment indicate that the proposed Kalman filter based method improves the accuracy both in locating and tracking mobile users than the measurement without Kalman filters.

Acknowledgements

Firstly, I would really like to thank my advisor Dr. R.D.McLeod for his patient supervising, support and encouragement over the course of the project. I would also like to thank the director of Internet Innovation Center (IIC) Blake Podaima for his advice and support. I must also express my appreciation to other staff and fellow students in IIC. Finally, I would like to thank my parents and my girlfriend for their support and encouragement during my Masters program.

Table of Contents

1. Introduction **1**

1.1 Background of Location System 1

1.2 Contribution of this thesis 7

1.3 Organization of this thesis 8

2. Related Works in Location System **9**

2.1 Location Methods 9

2.1.1 Proximity 9

2.1.2 Triangulation 10

2.1.3 Fingerprinting 13

2.2 Measurement of Metrics 14

2.3 Location Methods and Metrics used in this Thesis 17

3. WLAN-based Location System **19**

3.1 Overview of 802.11 19

3.2 The Properties of Indoor Signals 21

3.2.1 Temporal Variations 21

3.2.2 Spatial Variations 23

3.2.3 Statistical Properties of the RSSI 27

3.3 Impact of RSSI-based Systems Configuration upon Performance 30

3.3.1 Impact of the User's Orientation 30

3.3.2 Impact of the Number of Training Points 31

3.3.3 Impact of the Number of Samples 32

3.3.4 Impact of Access Points 32

4. Indoor Location Determination Methodologies..... 34

4.1 Free-Space Path Loss Model..... 34

4.2 Neural Network Model..... 35

4.3 Empirical Fingerprinting Model..... 38

4.3.1 K-Nearest-Neighbor Method..... 40

4.3.2 Probabilistic Method..... 42

5. Improved Location Estimation Accuracy Using the Kalman Filter..... 49

5.1 Introduction of the Kalman Filter..... 49

5.2 Extended Kalman Filter..... 53

5.3 Analysis of the Kalman Filter and EKF..... 56

5.4 Kalman Filter for Static Locating..... 58

5.5 Kalman Filter for Dynamic Tracking..... 61

6. Experiments and Discussion..... 66

6.1 Experimental Setups..... 66

6.2 Offline Phase..... 68

6.3 Online Phase..... 69

6.4 Results and Discussion..... 72

6.5 Tracking Results..... 75

6.6 Error Analysis..... 80

7. Conclusions and Future Work..... 82

Reference..... 85

Appendix A. Kalman filter for static locating in Matlab..... 90

Appendix B. A Kalman filter for dynamic tracking in Matlab..... 92

List of Figures

Figure 2.1 Proximity method..... 10

Figure 2.2 Determining a 2D position using lateration..... 12

Figure 2.3 Determining a 2D position using angulation..... 13

Figure 2.4 A functional block diagram of wireless indoor location systems..... 14

Figure 2.5 Four measurement based location techniques..... 16

Figure 3.1 Received signal strength from an access point at a fixed location..... 23

Figure 3.2 The relationship between distance and average RSSI..... 25

Figure 3.3 Signal strength contours..... 26

Figure 3.4 RSSI from 3 APs over 50 seconds (100 samples)..... 28

Figure 4.1 An example of multiple-layer perceptrons..... 37

Figure 4.2 Typical WLAN fingerprinting scenario..... 40

Figure 4.3 The RSSI-based fingerprinting algorithm..... 42

Figure 5.1 A complete work flow of Kalman filter..... 53

Figure 5.2 A complete work flow of Extended Kalman Filter..... 56

Figure 5.3 Kalman filtering for RSSI..... 60

Figure 5.4 Tracking moving object..... 62

Figure 6.1 Layout of experimental environment..... 67

Figure 6.2 Example of missing fingerprinting in radio map..... 71

Figure 6.3 Outline of location algorithm..... 72

Figure 6.4 Precision of the error distance for the two setups..... 74

Figure 6.5 Outline of tracking algorithm..... 78

Figure 6.6 The tracking results for a moving object..... 79

List of Tables

Table 3.1 RSSI from 3 different APs..... 28

Table 6.1 Accuracy of two setups..... 73

Table 6.2 Accuracy of two setups..... 80

List of Acronyms and Abbreviations

AOA	Angle of Arrival
AP	Access Point
CaDeT	Clustering and Decision Tree-based method
CCK	Complementary Code Keying
CDF	Cumulative Density Function
CSMA/CA	Carrier Sense Multiple Access with Collision Detection
DSSS	Direct Sequence Spread Spectrum
EKF	Extended Kalman Filter
FCC	Federal Communications Commission
GHz	Gigahertz
GPS	Globe Position System
IR	Infrared
ISM	Industrial, Scientific and Medical
KF	Kalman Filter
LOS	Line of Sight
MAC	Media Access Control
MHz	Megahertz
MLE	Maximum Likelihood Estimation
MLP	Multi-Layer Perceptrons
OSS	One-Step-Secant
PAF	Floor Attenuation Factor
PDA	Personal Digital Assistant
PDF	Probability Density Function
PF	Particle Filter
PHY	Physical

RF	Radio Frequency
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Index
RTLS	Real Time Location System
TDOA	Time Difference of Arrival
TOA	Time of Arrival
UWB	Ultra Wideband
VOR	VHF Omnidirectional Range
WAF	Wall Attenuation Factor
WLAN	Wireless Local Area Network

1. Introduction

1.1 Overview of Location Estimation Systems

Recently, many research efforts have begun to address on the issue of location-aware computing [5, 6, 9, 11, 13, 14, 17]. Location systems locate or track mobile objects or they facilitate locating objects for themselves for various purposes. Typical indoor location-aware computing applications range from monitoring a patient's movement or whereabouts, to tracking moving assets in a factory. Typical outdoor location-aware applications include supplying navigation information to vehicles for route planning and advising to providing location specific options for dining and entertainment purposes.

Location systems can calculate the positions of mobile objects using various types of information they collect. Usually, the positions of mobile objects are defined in two forms: absolute position and relative position. Absolute position refers the location of a mobile user in a 2D or 3D Cartesian coordinate system; while relative position is defined in relation to other objects such as a particular room or door. Based on the application environments, the location systems are typically classified into two basic categories: outdoor location systems and indoor location systems.

The signal often propagates in a wide open, unobstructed area in an outdoor environment.

Therefore, the transmitted signal suffers little degradation between a mobile user and the location systems. The Globe Positioning System (GPS) [35] is the most widely used outdoor locating technology which has been evolving since the 1980s that can determine a mobile terminal's location with accuracies of up to a few meters. The GPS system works with 24 orbiting satellites which can provide almost worldwide coverage. Each in range satellite continually transmits information messages to the receiver. The receiver estimates the distances from each satellite to the receiver from the time of flight. Based on these distances inferred from different satellites, a GPS receiver can calculate its position using geometric trilateration. GPS is mainly used for navigation purposes such as to locate aircrafts, automobiles, trucks, trains, ships as well as people. Enhanced GPS systems often utilize known references as well as correction data allowing for even greater accuracy. These are not discussed further as the primary interest here is indoor location and tracking systems. It should be pointed out however that the techniques developed here are not restricted to indoor location systems.

Another outdoor location estimation technology is based on cellular networks [36]. By operational necessity the location of a mobile terminal (cell phone) can be determined by the base stations which are serving it at the time. The estimation accuracy of this technology depends on the size of cell as well as the number of participating base stations. Additional information such as the radio signal's characteristics and relative position to other base stations can be used to determine a more accurate location of the

mobile terminal. For example, cellular phone subscribers are located when they associate with base stations using this technology; furthermore it is mandated by the FCC in response to requirements for E911 service. These location systems can often obtain satisfactory estimation results for outdoor applications; however, they are not applicable for many indoor location estimations due to the different signal propagation environments. It should be clear that the service provider knows a cell phone's approximate location when powered up as well as during a call.

The indoor environment has signal characteristics which are considerably different from outdoor signal characteristics. As such, different location technologies and sensors have to be developed for indoor locating and tracking. A number of indoor location systems are currently in use. These include technologies based on infrared (IR), traditional radio-frequency (RF), ultra-wideband (UWB), radio-frequency identification (RFID) tags and wireless local area network (WLAN) location systems.

IR is one of the earliest techniques used in the field of location-aware computing. The Active Badge system [37] is a typical IR based location system. In this system, the tracked object (e.g. person) wears a badge which can emit a unique IR code periodically (e.g. every 10 seconds). The IR receivers are distributed at known positions in a room area which can detect this IR information and determine the approximate position of the object. This system can locate a mobile object in a room size area with a considerable

accuracy. However, due to the characteristics of infrared signals, this system suffers from several drawbacks: (1) IR signals transmit over a limited range, so more receivers are needed if location estimation is required over a large area. (2) IR signals can not penetrate opaque obstruction such as walls, doors and partitions, hence, IR receivers need to be distributed in the line of sight with the tracked objects, (3) sunlight has negative impact on the IR signal causing interference, and (4) the high cost and complexity in installing the receivers.

UWB is another common location solution which is similar to an infrared location system. The UWB transceivers communicate with propriety scanners by sending radio signals having very wide bandwidth modulation. The location of transceiver can be determined with calculating the time of flight from the UWB transceivers to several scanners (at least three). A UWB location system can have a fairly high accuracy which may be up to 10 cm in laboratory environments. The main merits of UWB technology are its wide transmission bandwidth, typically in the 2-3 GHz unlicensed band, where it has been shown that the UWB signals suffer little from multipath propagation in the indoor environment. However, there are some obstacles in deploying UWB as a locating technology even though the use of UWB systems is licensed through the Federal Communications Commission (FCC). The UWB signals may impair other signals such as the signals from GPS receivers operating at approximately the 1.5 GHz frequency band. Moreover, UWB is still a new technology for most companies and enterprises.

Thus, the UWB devices are expensive to purchase and install.

RFID location systems are composed of RFID scanners and tags. There are two types of tags: active tags and passive tags. Active tags have batteries which can transmit and receive signals with scanners up to a 20 feet range, while passive tags have no batteries which only transmit their code to scanners when they are energized by the scanner, typically within a few feet. Due to their communication range, active RFID tags are more likely used in location purposes while passive tags used more for inventory. Each active tag can transmit information to a scanner, minimally including a unique ID code. When a person or other mobile client wearing an active tag is close to a scanner in a known position, the scanner can detect the signal from the tag and thus estimate the tag's position by its proximity. A centralized station is usually used to determine the position of the tag when in an environment of several scanners. This location system cannot locate the tag in its exact position but rather within an area corresponding to the range of scanners. RFID location systems need to install multiple scanners for finer resolution when locating and tracking. Considering the high cost of hardware and complexity of installation, RFID systems are also not easily deployed on a large scale.

RF location systems use radio frequency signal for communicating. Some of these systems include WLAN networks and Bluetooth networks adapted for use as RF location systems. The sensors in the networks detect radio signal information such as signal

strength, signal arrival time and signal arrival angle. Using these metrics, the position of the client can be extracted using triangulations on the time of arrival (TOA), time difference of arrival (TDOA) and angle of arrival (AOA) metrics. The details of these estimation methods are discussed in the next section. The advantages of RF devices over other location systems are that RF signals have a greater transmission range than other technologies, in addition, RF signals are easy to detect and RF devices are already widely deployed in many buildings.

WLAN location systems are interesting RF location systems. Unlike most other location systems which need propriety hardware such as scanners and tags, a WLAN location system can use off-the-shelf wireless devices and the existing WLAN infrastructure with no additional hardware. Since WLAN networks have been widely deployed in office buildings, airports, hospitals, shopping malls and many other indoor places, we can easily integrate existing Wi-Fi network infrastructure into the WLAN RF location systems. This reuse of infrastructure lowers the deployment cost compared to other location systems that require installing separate propriety hardware. Based upon the existing Wi-Fi network infrastructure all that is needed for a WLAN location system are Wi-Fi enabled tags and centralized software. The tags can be simple Wi-Fi devices working on IEEE 802.11 standard or adapters mounted in a laptop or PDA. They communicate with the Wi-Fi access points (AP) to detect the radio signal strengths. After collecting the signal strength, tags (client-based deployment) or access points

(infrastructure-based deployment) will report the measurements to a server, and then the centralized software will calculate the position of tags with corresponding algorithms. Obviously, there are many benefits of using WLAN location system such as lower costs, simple devices, faster deployment time and reasonably high location estimation accuracy. Hence, more and more research has focused on WLAN based location computing in recent years. Some IT enterprises such as AeroScout [39], Cisco [40], Ekahau [41] and PanGo Networks [42] are working in this market and rapidly developing this technique and associated technology.

1.2 Contribution of This Thesis

The contribution of this thesis is to develop a wireless local area network (WLAN) based location estimation prototype in a real world testbed (Engineering Building at the University of Manitoba) and to analyze the performance of the indoor location system based on the received signal strength index (RSSI) fingerprinting algorithm. To improve the estimation accuracy indoors, we embedded the Kalman filter into this model which has been widely used in outdoor tracking. Here, the Kalman filter is applied to this model not only for static locating but also for use in dynamic tracking, while most other papers only use the Kalman filter in the tracking process. Moreover, adding a measurement source (accelerometer) other than signal strength alone to decrease the estimation error is another novel contribution discussed in this thesis.

1.3 Organization of This Thesis

This thesis is organized as follows: Section 2 surveys common location estimation methods and measurement metrics. In section 3, we give an overview of WLAN based location systems and the signal properties in indoor environments. Section 4 introduces the indoor location determination methodology. In section 5, we explain the Kalman filter and how it is applied in the location estimation system. Section 6 shows experiment results in a real testbed which reveals an improvement in using the Kalman filter in the WLAN based location system. Finally, section 7 concludes this thesis and points out some problems for future work derived from this thesis.

2. Related Work in RF Location Systems

2.1 Location Methods

As mentioned, there are various methods that are used for location sensing applications. In this section, three basic methods, namely proximity, triangulation and fingerprinting are discussed. General location systems may use one or more of these methods. These methods were selected to be discussed in more detail as they represent common approaches and introduce terminology used throughout the thesis.

2.1.1 Proximity

The proximity method, namely the cell identification (Cell-ID) method as referred to in some papers, is the simplest and least accurate way to get an approximate location of a mobile device. In this method, the coverage area is divided into many small cells; each cell has a transceiver for communicating with the mobile clients. The proximity scheme estimates the position of mobile users by identifying which cell's transceiver receives the strongest signal from the mobile device (Figure 2.1). The estimated location of the object is within an area as large as the cell's transceiver or access point's signal coverage area. The accuracy depends on the cell coverage area and can be improved by deploying a dense architecture of low power access points. It is better to use this method to check the

existence rather than to track mobile objects [43]. This location method is commonly used in roughly determining a cell phone user's location by estimating which cell it is most closely associated with.

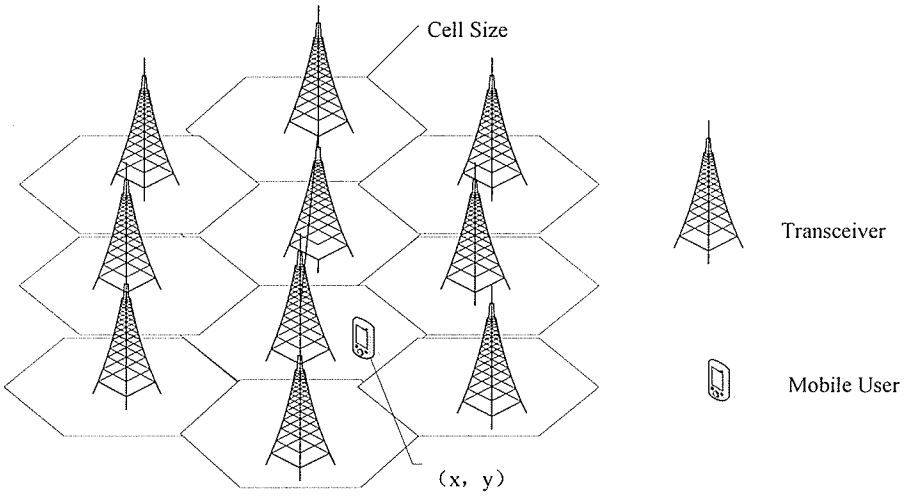


Figure 2.1 Proximity method

2.1.2 Triangulation

Triangulation is a method to determine the location of a mobile user by using geometric properties. Generally, triangulation can be classified into two types: lateration and angulation [44]. Lateration computes the position of mobile user from measured distances between the mobile user and at least three non-collinear reference points. Each reference point determines a circle in which the center is the location of the known reference point and the radius is the distance to the mobile object. Combining these

circles created by the reference points, the location of the mobile user is determined at the intersection of these circles. Figure 2.2 illustrates lateration in 2D space. Estimations of distances can be obtained from three general approaches: direct, time of flight and signal attenuation.

- Direct: measure the distance by using physical action or movements. For example, a robot can measure the distance by extending its probe to touch the solid object;
- Time of flight: estimate the distance by the product of the signal transmission velocity and the signal travel time from the mobile user to the reference point;
- Attenuation: estimate the distance by measuring the received signal strength. The emitted signal will attenuate as the transmission distance increases. As is known, the relationship between the signal travel distance and the signal power can be roughly represented by the following signal propagation model [7]:

$$P(r)[dBm] = P(r_0)[dBm] - (10n\log(\frac{r}{r_0})) + X \qquad (2.1)$$

where, $P(r)$ is the power received at distance r (meters), $P(r_0)$ is the power received from reference point which has a distance r_0 (meters) to the receiver, n is the propagation exponential value varying with different propagation environments, and X is the term accounting for shadow fading (a random variable with a normal distribution).

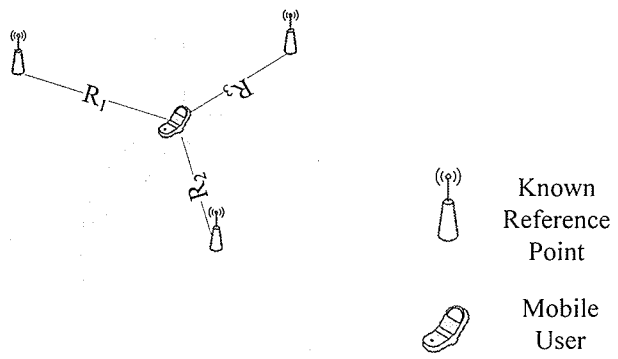


Figure 2.2 Determining position in 2d space using lateration

The GPS is the typical example that determines locations based on triangulation using lateration albeit in 3D space.

Angulation uses angles rather than the distances used in lateration to determine the position of mobile users. Generally, to locate a user, two angles to two reference points and the distance between two known reference points are needed to locate a mobile user in a 2D space. In this method, the measured results varied dramatically if any obstacle is present. The VHF Omnidirectional Ranging (VOR) navigation system is an example using the angulation technique [47].

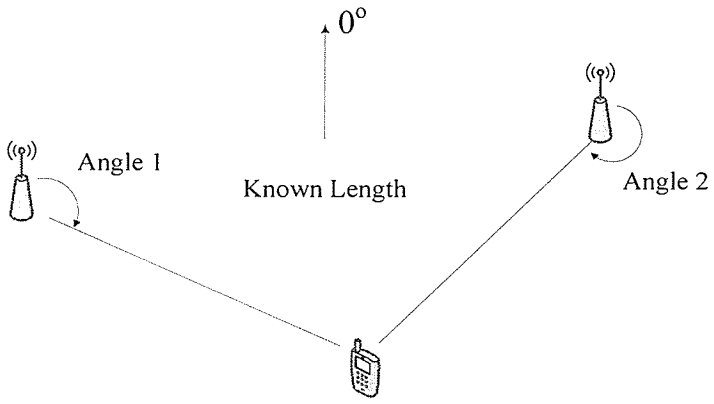


Figure 2.3 Determining a position in 2D space using angulation

2.1.3 Fingerprinting

The fingerprinting method is an empirical method [11] used to locate mobile users. In some papers, it is also called scene analysis which is one of the most commonly used methods in current location and tracking systems. The fingerprinting method first builds a grid map of a floor plan depending on the layout of the real environment including the number of doors and windows, positions of obstacles with the status of doors and windows being unknown. This is followed by a calibration process. This process builds a database of signal strength value at many reference sites in the grid map. To determine the object location, this method compares the observed signal strength value with the pre-stored values in the database, and then finds the closest match as its location. Since this method considers the effect of the real environment, it provides a better location estimation result than other methods. Also, the accuracy can be improved by adding more

reference points. However, this method is computationally costly due to the large number of reference points required if fine resolution is required. Moreover, the radio map is only applicable for a specific floor plan, and must be updated after any change in the radio environment. In the section 4.3, we will discuss this method in more detail.

2.2. Measurement of Metrics

In an indoor location system, the signal information is measured and then it is input into a location algorithm to compute the location of a mobile user. The typical functional block diagram of this system is shown in Figure 2.4. Wireless sensors measure a signal's metrics that can be used to locate a mobile user. Commonly used metrics are RSSI, AOA, TOA and TDOA. Each metric has its own advantages and is suitable for various positioning techniques in different measurement environments.

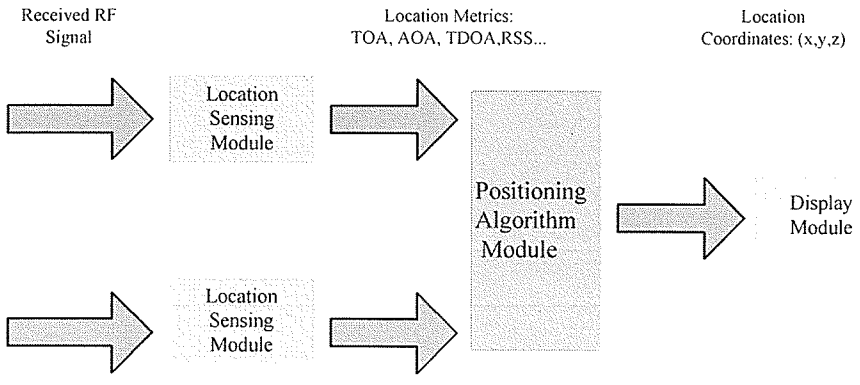
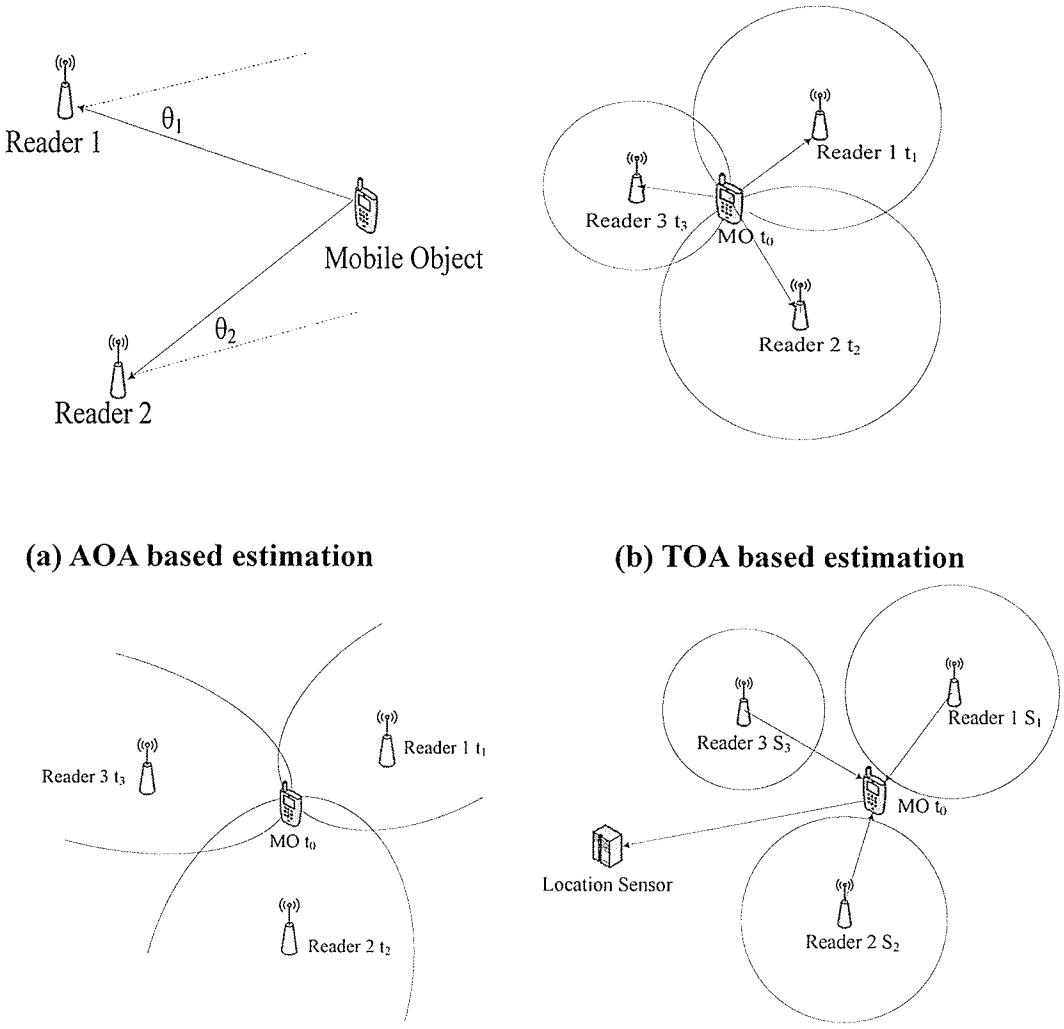


Figure 2.4 A functional block diagram of a wireless indoor location systems

More details about the four metrics are as follows:

- The AOA based system only needs two readers (transceivers) at known locations to determine the position of the mobile object [12]. The angle of arrival of the signal received from the same object is calculated and then a simple triangulation algorithm is used to estimate the position of the mobile object for each reader. The method is illustrated in Figure 2.4(a).
- The TOA based system is another common positioning system which uses the measurement of the propagation delay of radio signal between a mobile object and a number of readers (access points). Then the distance between the object and readers can be converted by the propagation time and the known velocity of the signal. A simple triangulation algorithm is used to calculate the position of the mobile object. Figure 2.4(b) gives the details of this method. Since TOA based system need a synchronization clock between readers and mobile objects, it leads to considerable costs and some technological problems in measurement.
- The TDOA based system is an advanced technique for solving problems in the TOA system. This system can locate a mobile object by computing the time difference of arrival of a signal transmitted from the object to three or more readers. Each metric received in the reader at a known location gives a hyperbola; the intersection of these hyperbolas can determine the location of the mobile object. Figure 2.4 (c) shows the details of the TDOA based system. The TDOA based system is also affected by the synchronization clock and multipath fading problems.



(a) AOA based estimation **(b) TOA based estimation**

(c) TDOA based estimation **(d) RSSI based estimation**

Figure 2.5 Four measurement based location techniques

- In a RSSI based system, the power received from several access points are used simultaneously to locate and track a mobile user. The RSSI information can be used to determine the distance between the mobile user and access points working either by triangulation algorithm or by a pattern matching algorithm. RSSI is an index of how strong the received signal from the transmitter is, which for example is

measured by the WLAN 802.11 adapter in dBm as follows:

$$RSSI[dBm] = 10\log\frac{\text{Re}\,ceived\ Power\ in\ watt}{0.001\ watt} \tag{2.2}$$

The AOA, TOA, and TDOA based systems all estimate the location based on the triangulation technique which requires a Line of Sight (LOS) between the transmitter and the receiver. However, in a complex indoor environment, there exist many Non-LOS situations between transmitters and receivers. Therefore, these systems would be affected by the attenuation and multipath effects, resulting in reduced positioning accuracy. In [45], they elaborate upon the difference of performance for each location system. The RSSI based system is best suited for tight indoor environments while the TDOA and TOA based system are better suited for outdoor or large open indoor environments. The big difference between the RSSI based system and other systems is that the RSSI system can be performed using the off-the-shelf hardware over existing 802.11 networks with no additional devices, whereas the AOA, TOA or TDOA system require special hardware to obtain timing information, thus the cost of these systems would be higher. Since the RSSI outperforms in other metrics for complex indoor environment positioning, the RSSI as the metric was selected to locate or track the mobile object.

2.3 Location Methods and Metrics used in This Thesis

From the above discussion, the fingerprinting location method arguably displays higher

accuracy for indoor location estimation due to its mapping of the real environment and that the RSSI based system provides reasonable location estimations with no additional hardware requirements. Therefore, we adopted the RSSI based location system with a fingerprinting technique in this thesis. In the commercial market, a popular terminology for location based on these systems is called Real Time Location Systems (RTLS) with the RSSI based location system being a typical example of the underlying technology.

3. WLAN-based Location System

3.1 Overview of 802.11

The RTLS standard has defined 2.4000-2.4835 GHz spread-spectrum transmissions as its air interface protocol specification and direct sequence spread spectrum (DSSS) as its modulation technique [1]. In this thesis, the IEEE802.11b is used as the air interface protocol which has the same frequency and same modulation technique as RTLS. As such, this section overviews the more common IEEE 802.11 standard in preparation to the use of IEEE 802.11 as an RTLS technology.

IEEE 802.11 is an IEEE standard for WLAN medium access control (MAC) and physical layer (PHY) specification [2]. The most popularly used today is 802.11b and 802.11g [2]. Both standards operate at the 2.4GHz ISM (Industrial, Scientific and Medical) band which is a crowded RF band. Microwave ovens, 2.4GHz cordless telephones and some Bluetooth devices also work within this band.

The 802.11b standard supports a maximum data rate to 11Mbit/s and the modulation technique is a complementary code keying (CCK) technique roughly based on previous DSSS modulation specification of 802.11 physical layers. 802.11b and others divide the 2.4GHz frequency range into 14 channels, each channel is 22MHz wide, 5MHz apart

from each other, therefore some adjacent channels overlap and interference will occur. To minimize the interference between access points, three non-overlapping channels: 1, 6 and 11 are usually used in the Americas. The transmitted beacon frame in IEEE 802.11 includes network information such as network identification, broadcasting network capabilities, synchronization and other control and management information [3]. Beacons will be transmitted within a fixed interval according to interval parameters. The IEEE 802.11 standard also defines a synchronization function which can keep all stations synchronized to the access point clock by using timestamp information of beacon frames.

The 802.11 standard has two basic topologies: independent and infrastructure topologies. In the independent networks which are also called ad-hoc networks, mobile stations communicate with each other directly, whereas in the infrastructure networks, all stations have to communicate with the local AP, which can communicate with other APs through a distribution system. The distribution system is the backbone of the wireless LAN and can be constructed either of wired or wireless networks. 802.11b devices usually communicate with omni-directional antennas. Typical transmission range is about 100m indoors and 300m outdoors. The coverage area can be enlarged through interconnected access points. This characteristic makes WLANs useful for indoor and campus area outdoor location estimation.

The IEEE 802.11 MAC applies a frame exchange protocol called carrier sense multiple

access with collision avoidance (CSMA/CA) to avoid error in the transmission of data. The carrier sensing is based on the energy detection and signal quality. The RF energy received by radio card can be measured, and is denoted the RSSI, which was introduced in last section. Usually, the unit of RSSI is dBm. The definition of dBm is described in Eq. (2.2).

3.2 The Properties of Indoor Signals

The signal characteristics of the wireless channel should be described before further work in location estimation. Unlike outdoor environments, indoor signal propagation is more difficult to estimate because of several phenomena that occur indoors such as reflection, diffraction and scattering [4]. These effects cause the indoor transmitted signals to suffer more noise and attenuation, which lead to more obvious signal variations. Generally, the variations of the wireless channel are divided into temporal variation and spatial variation [5].

3.2.1 Temporal Variations

The temporal variation is the result of received signal fluctuating over time at a fixed location. The variation is mainly caused by multipath fading, interference, shadowing and movements of humans indoors. Multipath fading is a simple term to describe the

radio waves transmitted between transmitters and receivers which have more than one path due to the reflection off obstacles, for example, windows, doors, tables and some other objects indoors. The multiple propagation paths lead radio waves having different arrival times to the destination, resulting in the different phases and amplitudes of signals at the receiver. These signals interfere with each other and they also interfere with the signals that were transmitted at a different time. Another cause of the temporal variation is the interference from some devices which also use 2.4GHz frequency band as their working frequency, for instance, microwave ovens, cordless telephones and Bluetooth devices. The propagation channels characteristics of 802.11 will be changed when such devices are put in a close proximity. Human movements in the wireless region, the change of temperature and humidity, and the time-varying transmitter power also contribute to signal temporal variations. The multipath fading also has a negative impact in outdoor estimation, but less than 5% outdoor location estimation error (like the GPS) is caused by the multipath fading effects due to its relatively more open propagation environments [7]. To see the temporal variations, an experiment was conducted to measure the received signal strength from a single access point over time at a fixed location indoors. In this experiment, the signal strength was sampled every 0.5 second and a total 400 samples are collected over 200 seconds. Figure 3.1 shows the received signal strength distribution over 200 seconds at a fixed location. We can see that the signals are centralizing around a mean value of -61dBm in this case from the figure. At the fixed location, a difference of 10dBm or more in the received signal strength is

considered as being time varied. Also from this figure, it shows that the probability of the RSSI from a single access point at a fixed location is similar to Gaussian distribution. In J.Small et al. [6], the conclusion stated that the RSSI is Gaussian distributed in dBm based on long duration sampling.

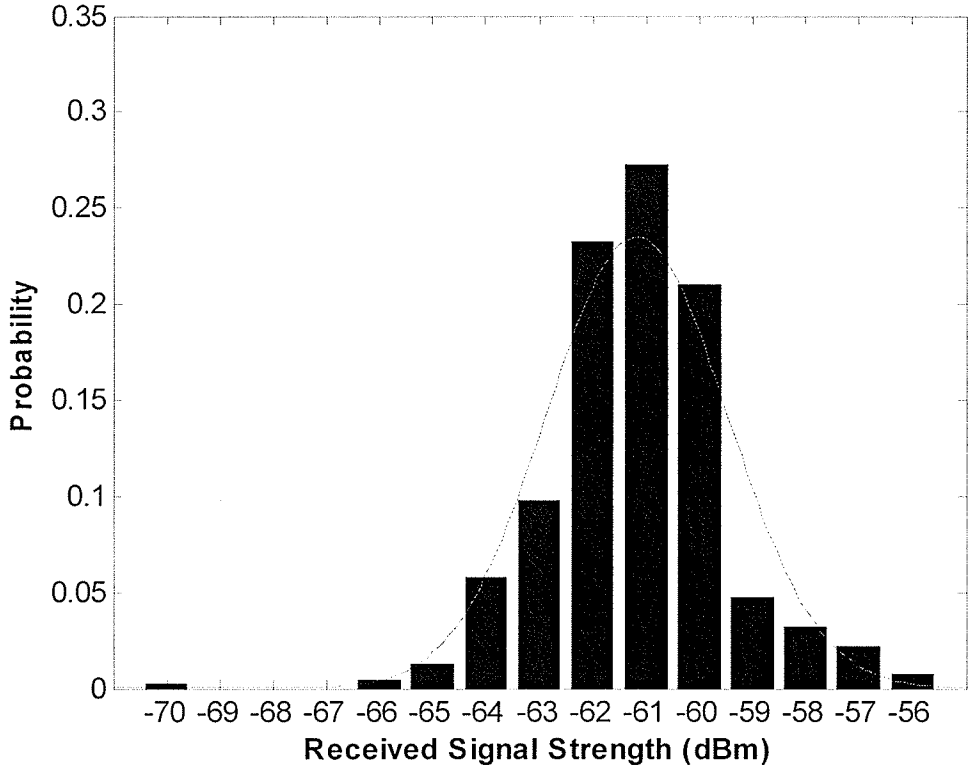


Figure 3.1 Received signal strength from an access point at a fixed location

3.2.2 Spatial Variations

The spatial variation is mainly caused by the combined effects of the large-scale fading and the small-scale fading.

The large-scale fading refers to the attenuation of the received signal strength as the propagation distance increases. The statistics associated with the large-scale fading provides a way to estimate the mean value of the RSSI as a logarithmical function of the distance. In a simple indoor environment, the RSSI at distance r can be expressed by the known log-distance path loss model represented in Eq. (2.1), which is reproduced in Eq. (3.1) [7]

$$P(r)[dBm] = P(r_0)[dBm] - (10n\log(\frac{r}{r_0})) + X_\sigma \qquad (3.1)$$

where, r is the distance between the transmitter and receiver,

r_0 is the nominal reference distance, typically 1 meter,

n is the propagation exponential value varying with different propagation environments,

X_σ refers the shadow fading which is a zero-mean Gaussian random variable with standard deviation σ.

X_σ is a site- and distance-dependent factor [8] which is mainly a consequence of the wall attenuation factor (WAF) and the floor attenuation factor (PAF) [7]. The WAF and PAF are often empirically obtained based on measured data.

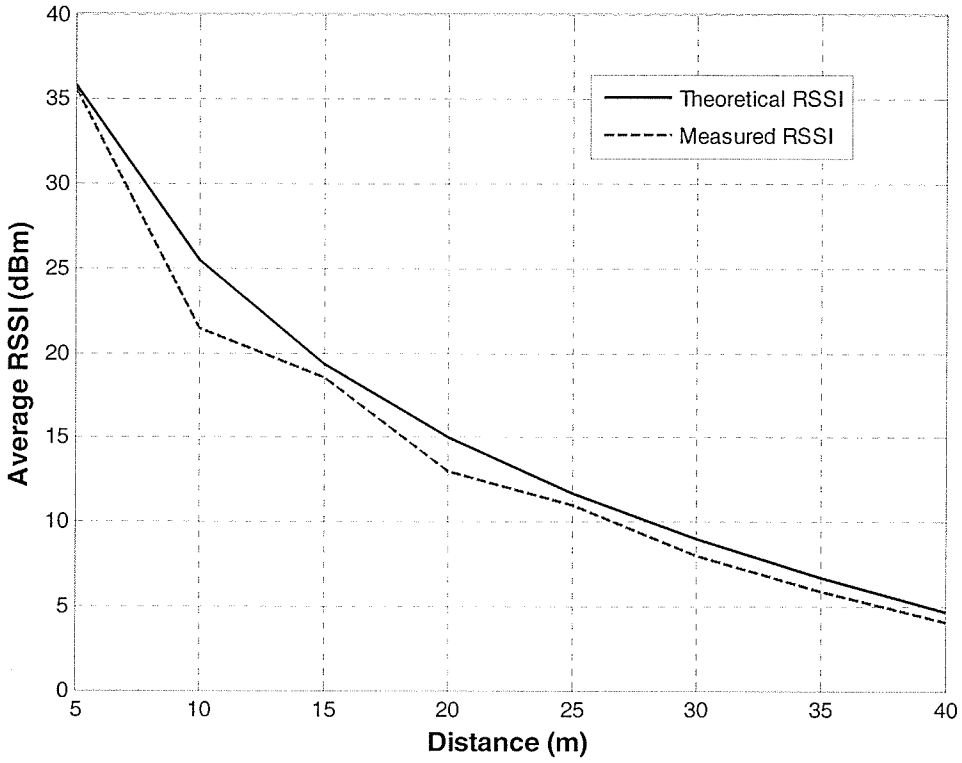


Figure 3.2 The relationship between distance and average RSSI

Since the random variable X_σ exists, the RSSI attenuates at different rates in different directions. Figure 3.2 shows the relationship between the distance and the RSSI in a building. In Figure 3.2, the red dashed shows that the RSSI decreases as the distance between transmitter and receiver increases. However, the decreasing variation did not follow a logarithmical model which is represented by the blue line in the figure because of higher environmental variability indoors such as asymmetry of building layouts, building materials and movements of doors. The large-scale fading leads to different locations having different RSSI values for each AP at comparable distances. This

characteristic helps to justify building the signal map for the RSSI based location system used in this thesis.

The small-scale fading refers to the dramatic changes of the signal amplitude and phase after short distance change or short time change. Figure 3.3 reproduced from [5] shows the RSSI variations can be nearly 10 dBm within a small distances of about 3 inches. Multipath propagation, the orientation of mobile user, the motion of mobile users and the transmission bandwidth may all cause small-scale fading. Multipath fading is a main cause of small-scale fading which not only causes temporal variations, but also spatial variations. Usually, if there is a strong line of sight among multiple propagation paths, the small-scale fading is denoted Rician fading; otherwise it is called Rayleigh fading.

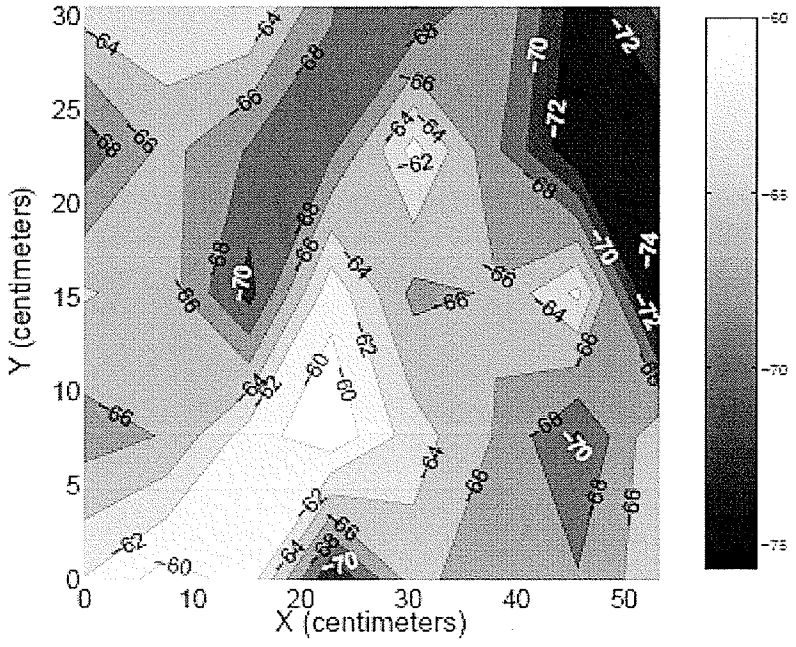


Figure 3.3 An example of signal strength contours [5]

3.2.3 Statistical Properties of the RSSI

Because of the complex nature of indoor propagation, it is hard to depict the RSSI with a standardized path loss model. Several indoor experiments were conducted to uncover statistical properties of RSSI within the specific area of interest.

The RSSI is traditionally considered being Gaussian distributed. In fact, due to the many uncertain factors like human movements in a real test environment the RSSIs are not always Gaussian distributed. The results in [9] noted that some of the RSSI distributions were skewed to the left, only a few distributions fitted a good Gaussian distribution approximation.

Based on data collected in experiments, another statistical property of the RSSI was identified. The standard deviation is in direct proportion to the average RSSI. Figure 3.4 revealed this trend. 100 RSSI samples were collected from 3 different APs at the rate of 2 samples per second, the AP with highest mean signal strength value had the largest standard deviation and the AP with lowest mean signal strength value had the smallest standard deviation. Table 1 shows the mean values and standard deviations of RSSI from three different APs at a fixed location. The mean values from three AP were -63.73, -78.75 and -86.86 dBm respectively. Their corresponding standard deviations were 3.19, 2.23 and 1.58 dBm. This is because the stronger signal strength, the more serious the

interference it suffered. Considering the RSSI values from one AP, the large signal variations may cause two nearby locations to have dissimilar RSSI values, which increase the difficulties in estimating position [9]. On the other hand, two distant (but different) locations may have similar RSSI values because their low level RSSI suffers little signal variations. Therefore, using multiple RSSI values and different power level APs may lead to a more accurate location result.

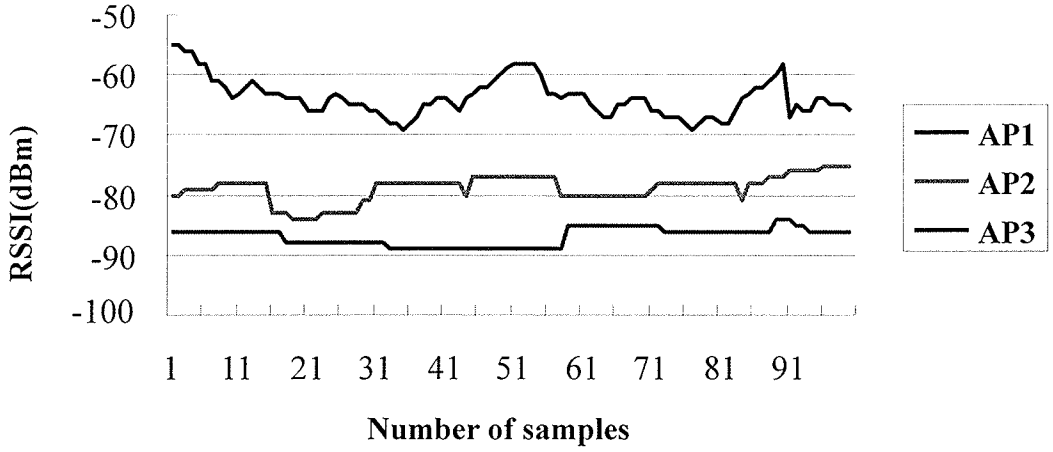


Figure 3.4 RSSI from 3 APs over 50 seconds (100 samples)

	AP 1	AP 2	AP3
Sample Mean (dBm)	-63.73	-78.75	-86.86
Standard Deviation	3.19	2.23	1.58

Table 3.1 RSSI from 3 different APs

Based on a large number of samples collected from one AP at the fixed location, it was also found that the consecutive RSSI samples were highly auto-correlated. A result in [5] observed that the autocorrelation of consecutive samples can be as large as 0.9. This particular characteristic of RSSI samples is taken into consideration when using multiple RSSIs to enhance location accuracy (i.e. the Kalman filter).

The RSSI samples are time dependent. Measurement taken at different time of day or different day showed variable results. Previous work [9, 10] showed different distributions of RSSIs from an AP at a fixed location over different time periods. Indoor temperature, humidity and the sunlight were associated with the reasons for this problem. The experiments conducted at the same time over a number of different days are used to reduce the impact of this problem.

In RSSI based location estimation systems, a large number of RSSI values were obtained from different APs at fixed locations. The experiments in [9] that were based on one hour tests revealed that the RSSI from different APs were independent. This uncorrelated property is very important in the probabilistic algorithm which will be discussed later.

Although various factors lead some RSSI distributions to be non-Gaussian, it is still assumed that the indoor location system is roughly governed by a Gaussian distributed model. In this thesis, the following assumptions are made:

- All RSSI distributions are Gaussian
- The standard deviations of all RSSI are constant
- All RSSIs are mutually independent

3.3. Impact of RSSI-based Systems Configuration upon Performance

As is well known, the most important criterion to judge a location system is accuracy, which is the degree of closeness of the user's estimated location to the user's real location [46]. In a complicated location environment and system, many factors will contribute to the final location estimate.

3.3.1 Impact of the User's Orientation

In several experiments, the user's orientation was found to have a significant impact on the final location estimation result. In the experiments of [9, 11], the results showed a significant difference for four orientations at a fixed location. In some cases, the difference can be as large as 9dBm because of the impact of different orientation. This is because mobile devices are usually held or worn by a person. The human body can absorb RF energy when the human body obstructs the signal, causing signal attenuation and a reduction in the RSSI. Therefore, a user's orientation should be considered when computing a user's location. Dealing with a user's orientation is challenging and no

effectual location determination system has perfectly solved this issue so far. To limit the impact of user's orientation in this experiment, we arbitrarily restrict the users to one orientation at all times. In the future, an electronic compass may be used to estimate the user's orientation, or previous estimates of the user's velocity may also be used.

3.3.2 Impact of the Number of Training Points

In any fingerprinting location assisted estimation, the number of training points and the granularity between these points also has an impact on the estimation accuracy. Generally, the greater the number of training points used, the more accurate the location estimation. The experimental results in [11, 12] showed that as the number of training points increased, the estimation accuracy improved significantly. In their experiments, the mean error decreased from 6.73 m to 1.23m when the training points increased from 4 to 27. On the other hand, the large number of training points does not always benefit the location estimation results. With a grid of training points over a threshold, the accuracy improves little, and even degrades in some cases. Dense training points also leads to a higher computational cost. Therefore, an appropriate number of training points should be considered in any fingerprinting location assisted estimation system. Moreover, the distribution of training points would impact the system estimation due to the complicated indoor environment. Indoor signals attenuate at different rates, and rapid attenuation occurs in the presence of obstacles such as walls, tables and doors. To obtain a more

accurate estimation, we set dense granularity in hallways or steps, while we set sparse granularity in open floor areas. Usually, a 1 to 3 meter spacing of training points is adopted depending on the environment [13].

3.3.3 Impact of the Number of Samples

As discussed previously, various noise sources exist in the indoor environment. A transmitted signal may be corrupted by noise combined with the true value, which leads to poor location estimation. Therefore, one way to deal with this phenomenon is to use multiple signal samples from each AP at all locations. This can help to reduce the effect of noise. Similar to training points, too many samples cause a high computational cost, as well as long sampling time. For real-time location purposes, we only have limited sampling time to locate a mobile user. As with most location systems in use, in each single test 20 samples from each AP were taken at each fixed location.

3.3.4 Impact of Access Points

The selections of AP include the number of APs and the location of the APs. In a triangulation method, at least three APs are needed to locate the mobile user. In fact, in a real WLAN environment the number of APs is often much more than three. Obviously, more APs used in location estimation leads to a more accurate result and a corresponding

higher computational cost. Therefore, this motivates us to select a subset of available APs in the location system and find a tradeoff between the accuracy of estimation and complexity of computation. Previous work [14] introduced two methods to reduce the computational complexity via clustering spatial locations. The work of [15] represented an intelligent selection of a set of clusters using K-means known as *CaDeT* (Clustering and Decision Tree-based method) to improve the power efficiency and reduce computations. Commonly, the subsets of APs with the strongest signal strengths are chosen for locating due to their highest probability of coverage. In this thesis, the APs are selected to cover the whole test area and show distinct RSSI characteristics at fixed locations.

The location of APs is a critical issue in some location estimation methods such as TOA and TDOA triangulation estimation. Some of the location details of APs were discussed in Section 2. In this thesis, the RSSI-based fingerprinting method is justified and used. In this method, the location of APs is unnecessary, only the RSSI values at each training point are the metrics that need consideration. Of course, there needs to be APs but provisioning their location utilizes the original efforts when they were deployed for the WLAN itself.

4. Indoor Location Determination Methodology

As discussed in previous sections, the RSSI method appears to provide reasonable estimation for indoor location determination. There are at least three main models for RSSI based indoor location systems: the free-space path loss model, the neural networks model and the empirical fingerprinting model.

4.1 Free-Space Path Loss Model

In this model, the RSSI at a location can be represented by a function related to the distance between APs and mobile users. This function is listed in Eq. (3.1), re-written here as follow:

$$P(r)[dBm] = P(r_0)[dBm] - (10n\log(\frac{r}{r_0})) + X_\sigma \qquad (4.1)$$

Since the signal strength has the relationship with the distance between the AP and the mobile user, we can estimate all RSSI values received and estimate a user's location using this model. Noise and obstacles are taken into account by a zero-mean Gaussian variable X_σ. Once the general environment characteristics such as the wall attenuation factor and the floor attenuation factor are known, only a small number of measurements are required in this model. Compared to the empirical fingerprinting model discussed later, the advantage of this approach is fewer measurements. After obtaining a small

number of general environment measurements, the free-space path loss model uses simple computations to estimate the RSSI at each location. However, as RSSI in an indoor environment has many environment dependent characteristics, the function model does not express the RSSI accurately. Moreover, the work of [16] pointed out that the sensitivity of path loss is severely influenced as the distance increases. That is, at two large distance distinct locations, the model does not express a corresponding RSSI difference clearly. Therefore, this model has not been adopted in any commercial RSSI RTLS systems.

4.2. Neural Network Models

Although we estimate a random variable in the propagation function, the free space path loss model does not describe the RSSI accurately due to the complex and heterogeneous indoor environment. Another location determination method called a neural network model builds a more flexible model to deal with this issue. Neural networks can build a non-linear mapping between the input and output. In location estimation systems, the RSSI at each test location from different APs are taken as the inputs of the network and the real test locations are taken as the outputs. Therefore, a non-linear mapping from the RSSI to the location is built. In the training phase, the network is trained using pairs of RSSI values and their actual locations. The internal structure of the neural network is adjusted after each input/output pair in order to obtain a more accurate result when the

network is input with unseen RSSI values. The most common neural network used in location estimation systems is the multi-layer perceptron (MLP) which uses supervised training algorithms [17]. The general MLP includes the input layer, one or more hidden layers which are constituted by a large number of processing units and the output layer. An example of a MLP is shown in Figure 4.1. The units that are not treated as output units are called hidden units. In this one hidden layer example, there are m RSSIs as the inputs, k locations as outputs and n hidden units between inputs and outputs. The outputs of the jth hidden unit are obtained by the following functions:

$$a_j = \sum_{i=1}^{m} w_{ji}(1)x_i + x_0 w_{j0}(1) \tag{4.2}$$

$$z_j = g(a_j) \tag{4.3}$$

Here, $w_{ji}(1)$ denotes a weight in the first layer from the input i to the hidden unit j,

$w_{j0}(1)$ denotes the weight from the bias x_0 to the hidden unit j,

a_j is the output value of the jth hidden unit

z_j is the activation value of the jth hidden unit

The activation function $g(\cdot)$ is a smooth non-linear function. In neural networks, the continuous sigmoid function (4.4) and the hyperbolic tangent function (4.5) are two principal forms:

$$g(x) = \frac{1}{1+e^{-x}} \tag{4.4}$$

$$g(x) = \tanh(x) = \frac{1-e^{-x}}{1+e^{-x}} \tag{4.5}$$

Similarly, the output of the whole network can be obtained by the summation of the hidden units, then transforming the linear combination using a non-linear activation function (4.4) or (4.5). After the MLP model is built, the internal structure of the network needs to be trained to obtain correct estimation. During the training phase, the weights and biases of the units are modified with each input/output pair to minimize the error.

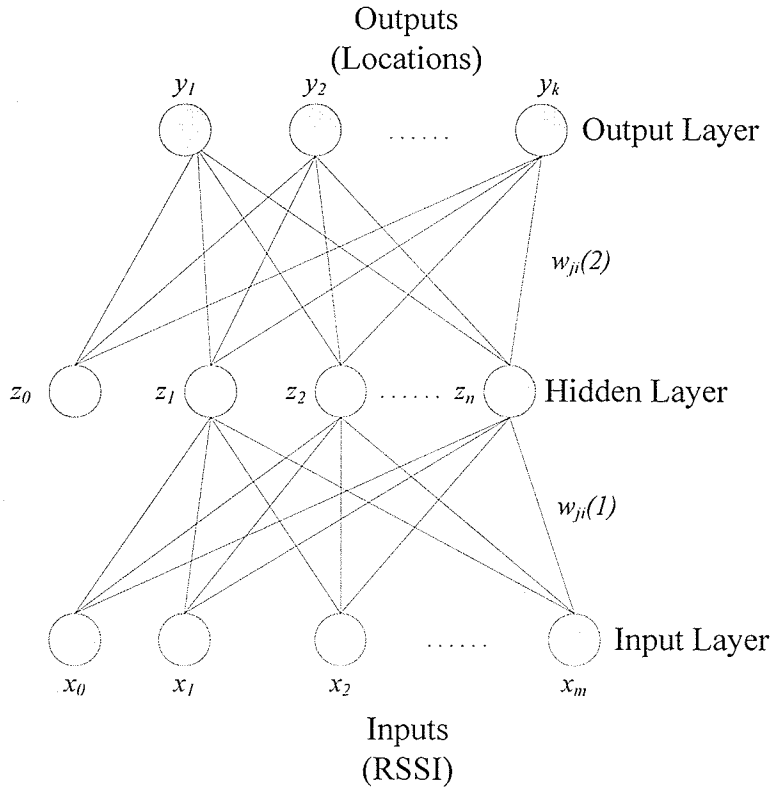


Figure 4.1 An example of a multiple-layer perceptron (MLP)

The MLP model can be viewed as an m-dimensional to k-dimensional Euclidean space non-linear mapping. It has been proven that a single layer MLP network can approximate any continuous function to any desired accuracy with a large number of hidden units and a long enough training phase. In the work of [18], the best generalization value (about 1.52 meters) could be reached after 4000 iterations using a one-step-secant (OSS) learning algorithm. This result is acceptable for many indoor location estimation purposes. However, the defect of this method is that adjusting the model is rather complicated; there is no regular rule for the adjustment, which mostly depends on experience.

4.3 Empirical Fingerprinting Model

The empirical fingerprinting model is based on identifying the location of user's location by comparing the measured RSSIs with a pre-built radio map (Figure 4.2). Generally, there are two basic phases in the empirical fingerprinting model: the offline phase and the online phase. In the offline phase, a floorplan of the test area with pre-determined training points is needed. The test area is configured with APs, which provide overlapping coverage in portions of the floor, and together cover the entire floor. Then a number of training points in this area are selected depending on the physical environment. The training point selection requirements were discussed in the Section 3.3.2. Summarized here, a dense distribution is selected in crowded areas and a sparse

distribution is selected in wide open areas. At each training point, the RSSI features from different APs are collected to build a signal location database, which is also called a radio map. Due to the presence of noise in the indoor signal channel, one sample does not contain sufficient signal information for a location. Usually, a mean value of samples over a period or multiple samples is used as the signal fingerprint at the location to reduce the effect of noise and interference. In the online phase, the mobile user compares the RSSI feature at its location with the radio map pre-created in the offline phase. Then through a location determination algorithm, the probable location of mobile user is estimated by minimizing either the Manhattan distance or Euclidean distance between the measured RSSI features in the online phase and pre-stored RSSI features in the offline phase. For example, if the RSSI vectors are $(s_1, s_2, ..., s_n)$ and $(m_1, m_2, ...m_n)$ respectively, for the offline phase and online phase, the Manhattan distance and Euclidean distance can be expressed as:

Manhattan distance: $d_m = \sum |s_i - m_i|$

Euclidean distance: $d_e = \sqrt{\sum (s_i - m_i)^2}$

Depending on the different forms of RSSI fingerprinting stored in the offline phase and measured in the online phase, we can further divide the empirical model into mean K-Nearest-Neighbor method and a probabilistic method.

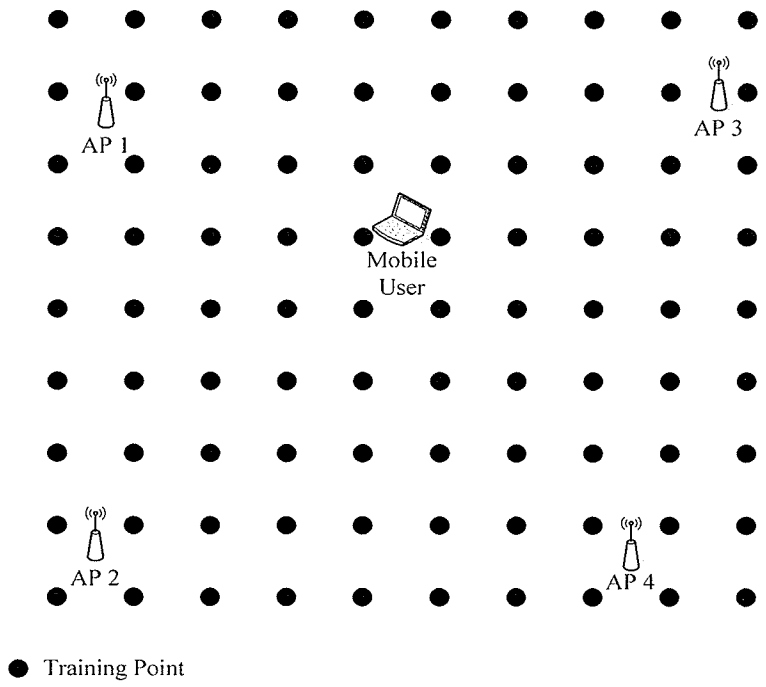


Figure 4.2 Typical WLAN fingerprinting scenario

4.3.1 K-Nearest-Neighbor Method

From the previous discussion, taking only single sample of the RSSI at a location as the fingerprint leads to the loss of information and affects the accuracy of location estimation significantly. Hence, using the mean value of RSSI samples at a fixed location over a period time instead of a single RSSI sample from AP can make the estimation more accurate. During the offline phase, a signal strength database is built by collecting a large number of RSSI features at n distinct training points: $l_i = (x_i, y_i),\ i = 1,...,n$. The RSSI is stored as $s_{ij} = (s_{i1}, s_{i2},..., s_{im})$, where m is the number of access points and s_{ij} is the received signal strength information (which is the mean value of N signal strength

sampled here) from the ith training point at the jth access point AP_j. In the online phase, the RSSI features observed at the desired location along with the signal database are fed into a matching algorithm to find an approximate location estimation of the mobile user. The location is estimated by minimizing the Euclidian distance d_k between the measured set of received signals and signal samples in the database.

$$d_k = \sqrt{\sum_{i=1}^{m} (s_{ki} - s_{ki}')^2} \tag{4.6}$$

where s_{ki} is the signal information stored in the offline phase and s_{ki}' is the average signal strength measured in the online phase. After calculating the Euclidean distance with respect to all RSSI data stored in the database, K samples in the database with smallest distance are found. The location of the mobile user can be calculated by using a weighting factor summed over the K best matching neighbors.

$$< \hat{x}, \hat{y} >= \sum_{k=1}^{K} w_k \cdot < x_k, y_k > \tag{4.7}$$

where w_k is the kth weighting factor defined as follow:

$$w_k = \frac{\frac{1}{d_k^2}}{\sum_{k=1}^{K} \frac{1}{d_k^2}} \tag{4.8}$$

The architecture of RSSI-based fingerprinting algorithm is showed in Figure 4.3. This

method is easy to implement due to its simple computation requirements. However, the accuracy of this estimation method is not very high because the mean value of RSSI is not sufficient to express the RSSI distribution from APs at a fixed location. Considering these reasons, an advanced method called a probabilistic method is presented [5].

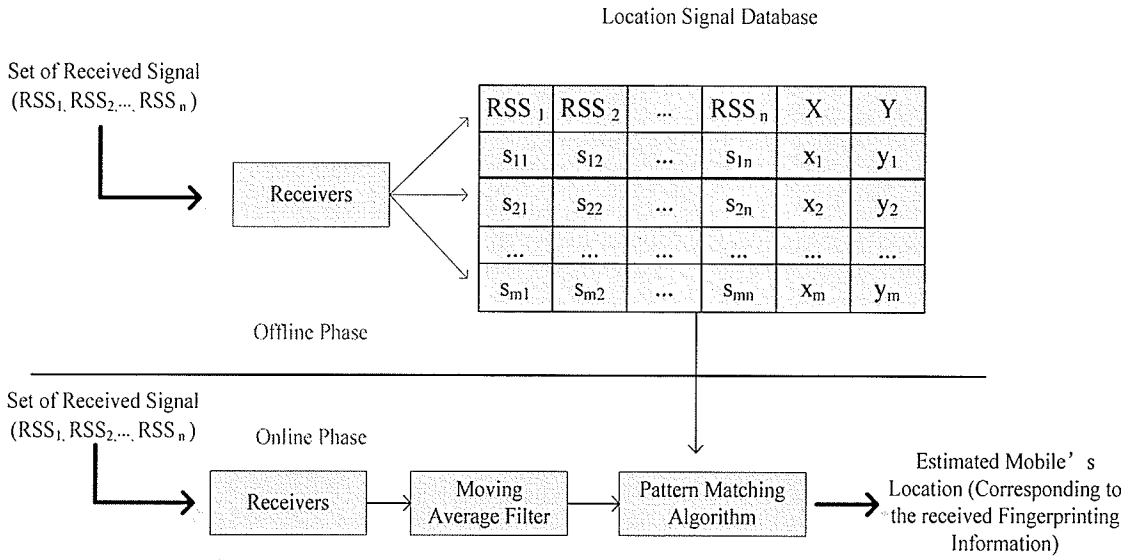


Figure 4.3 The RSSI-based fingerprinting algorithm

4.3.2 Probabilistic Method

Instead of using a mean value to express the RSSI features from the APs at a fixed location, the probabilistic method uses the signal strength probability distribution at a fixed location as the fingerprinting function [14]. At each location l, we obtained the observation of the signal information vector from k APs as: $s = (s_1, s_2, ..., s_k)$. The

location determination algorithm attempts to maximize the conditional probability value: $P(l|s=(s_1,s_2,...,s_k))$. The probabilistic method provides more information about the RSSIs at a fixed location; hence it results in a better estimation. Based on the RSSI distribution, we consider two approaches in this method: non-parametric distribution and parametric distribution.

Non-parametric Distribution

When the RSSI at a fixed location is treated as a non-parametric distribution, the details of operation can be depicted in two phases with this approach.

In the training phase, APs are chosen based on the balance between the coverage issue and computation issue. Selected APs should have coverage of the whole test area and the number of APs should be limited to reduce the computation cost, for example, if there are k APs available in the test area. Then, the RSSI probability distributions from those APs at each training point are collected, the fingerprinting feature of the RSSIs is expressed as a joint probability form: $P(s_1,s_2,...,s_k)$, where s_i denotes the RSSI from the ith AP at the training point. The joint probability distribution can be calculated by the maximum likelihood estimation (MLE) as shown in Eq. (4.9)

$$P(s_1, s_2, ..., s_k) = \frac{Count(s_1, s_2, ..., s_k)}{Size\ of\ Training\ Data} \tag{4.9}$$

Using maximum likelihood estimation, the joint probability distribution needs a large amount of training data. Also as the number of APs increase, the size of training data increases exponentially with respect to k. To reduce the training size, another probability distribution is considered instead of the joint distribution. Assume the test APs are working independent of each other; as such the RSSI from different APs at the fixed location should also be independent. The joint distribution can be transformed into a product of marginal probability distribution as shown in Eq. (4.10)

$$P(s_1, s_2, ..., s_k) = P(s_1)P(s_2)...P(s_k) \tag{4.10}$$

$P(s_i)$ is the probability of RSSI $= s_i$ appearing in the entire training set at a given location which can be obtained by using the normalized histogram from AP i at a fixed location. In fact, as we have discussed in Section3.1, there are only 3 non-overlapping channels in an 802.11 WLAN. The independence of APs is not completely accurate if more than 3 APs are implemented because some APs may work at the same channels.

In the online phase, the RSSI vector is measured at the mobile user's location within the test area: $s' = (s_1', s_2', ..., s_k')$. The purpose of online phase is to determine a location in the test area which maximizes the conditional probability: $P(l|s')$, it can be expressed as:

$$\arg\max_l [P(l|s')] \tag{4.11}$$

From the Bayes' theorem, the posterior distribution of the location can be obtained; hence, we can rewrite the Eq. (4.11) as Eq. (4.12):

$$\arg\max_l [P(l|s')] = \arg\max_l [\frac{P(s'|l)P(l)}{P(s')}] \tag{4.12}$$

Here, as $P(s')$ is constant for all locations, Eq. (4.12) can be simplified as Eq. (4.13):

$$\arg\max_l [P(l|s')] = \arg\max_l [P(s'|l)P(l)] \tag{4.13}$$

P(l) is the prior probability of location *l*, which depends on the personal mobile user's profile. If there is no mobile user profile available, we simply assume the prior probability is a uniform distribution here. That is, the mobile user has the same probability for all locations. Based on this reasoning, we can rewrite the probability of Eq. (4.13) to that of Eq. (4.14):

$$\arg\max_l [P(l|s')] = \arg\max_l [P(s'|l)] \tag{4.14}$$

Substituting Eq. (4.10) into Eq. (4.14), the location determination system can be calculated as a product of the marginal probability distribution:

$$\arg\max_l [P(l|s')] = \arg\max_l [\prod_{i=1}^{k} P(s_i'|l)] \tag{4.15}$$

Therefore, the location *l* which satisfies Eq. (4.15) can be treated as the mobile user's

estimated location.

Parametric Distribution

When the RSSI from APs at a fixed location can be modeled by a parametric distribution, we can also describe the estimation approach in two phases, similar to the non-parametric distribution case.

The RSSI from each AP at a fixed location can be modeled as a parametric distribution in the training phase, for example, a Gaussian distribution model is used in [5]. Other parametric models such as the Kernel method can also be implemented in location estimation [13]. The probability density function (pdf) of a Gaussian signal distribution can be expressed as:

$$f(s) = \frac{1}{\sigma\sqrt{2\pi}} \exp[(s-\mu)^2/(2\sigma^2)] \tag{4.16}$$

where μ is the mean value of signal samples from a AP and σ is the standard deviation of this distribution.

In the online phase, the objective is to obtain the maximum conditional probability $P(l|s')$. Similar to the non-parametric method, from Eq. (4.14) and Eq. (4.15), we know that:

$$\arg\max_l [P(l|s')] = \arg\max_l [P(s'|l)] = \arg\max_l [\prod\nolimits_{i=1}^{k} P(s_i'|l)]$$

in the Gaussian parametric distribution, the conditional probability can be obtained by using Gaussian pdf [5],

$$P(s_i'|l) = \int_{s_i-0.5}^{s_i+0.5} f(s)ds \tag{4.17}$$

The upper and lower limit of integral used here is because the received signal strength s_i is quantized to integer values.

Both non-parametric and parametric methods enhance the accuracy of location estimation compared to previous methods. However, each has its own advantages over the other. The non-parametric method collects and stores a large number of signal probability distributions in the training phase, then finds the most appropriate location given an observed signal in the online phase using a simple array lookup. In the online phase, the calculation load for the nonparametric method is lower than the parametric method. Although the parametric method reduces the size of radio data storage by only storing the mean values and standard deviation in the radio map, the computation of radio map is increased because mean values and standard deviation add some complex computations such as exponentiation and square root calculations. Compared to the non-parametric method, the parametric method improves the estimation results slightly and can avoid the effect of missing signals in some locations [5]. In this thesis, we adopt

the non-parametric method and we improve the non-parametric method using a Kalman filter. The improved non-parametric method will be discussed in later sections.

5. Improved Location Estimation Accuracy Using a Kalman Filter

As discussed in previous sections, there exist various noise sources in the RSSI measurement and location estimate, which leads to inaccurate estimate results. Some noise influences can be avoided, or at least can be reduced to an acceptable level. To minimize the measurement error and improve estimate accuracy, many post processing methods have been proposed for use in improving the locating and tracking estimation. The Kalman filter (KF), extended Kalman filter (EKF) and particle Filter (PF) are often used to filter and smooth the motion of mobile users [19], [20], [21]; Kalman filters are used in [22], [27], [28] to reduce the effect of measurement noise in tracking linear moving users; extended Kalman filters are used in [23], [29], [30] for linearizing the nonlinear system model; the particle filter is another approach for tracking an object using a nonlinear system model [31]; another possible method similar to H-infinity filtering is used in [24]. In this thesis, we adopt Kalman filter to improve the performance of location system from two aspects: locating the static mobile user and tracking the dynamic mobile user.

5.1. Introduction of the Kalman Filter

In the process of measurement and estimation, the current state of the system may be related to previous states. Therefore, the information of previous states can contribute to

the estimation of the current state. The Kalman filter is the most commonly used in building the relationship between the previous information and current estimation. The Kalman filter is a recursive algorithm [25] used to obtain the most optimal estimation of current state, combining measurements and predictions in a linear dynamic system interfered by noise. A brief review of the Kalman filter is described as following [26].

The Kalman filter is used to estimate the variables of a wide range of processes by minimizing the variance of the estimation error. Generally, the processes we are measuring must be described by a linear system. The received signal strength and moving mobile user both can be approximated as linear systems. A linear system can be described by two equations:

Process Equation:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \quad , \qquad p(w) \sim N(0, Q) \tag{5.1}$$

Measurement Equation:

$$z_k = Hx_k + v_k \qquad p(v) \sim N(0, R) \tag{5.2}$$

In the above equations,

x_k and z_k is the estimate value and measurement value vectors respectively at time step k.

u is the known control input to the system

A, B, H are parameter matrices in the system

w and v are noise vectors, w is called process noise, v is called measurement noise.

Both noise sources are independent random variables which are assumed white with zero autocorrelation. Process noise w is assumed to be Gaussian distributed with zero mean and variance of Q. Measurement noise v is also assumed Gaussian distributed with zero mean and variance of R.

The Kalman filter predicts a process state for the next time step, and then incorporates the noisy measurement to correct its prediction. Hence, the Kalman filter has two distinct stages: time update stage and measurement update stage. In the time update stage, the filter predicts the next time step state and error covariance using the current state. An initial value of state variable and error matrix should be decided before the Kalman filter is run. The equations are presented below:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k \quad , \tag{5.3}$$

$$P_k^- = AP_{k-1}A^T + Q \quad , \tag{5.4}$$

where, $\hat{x}_k^-$ is the a-priori state estimate at time step k

$\hat{x}_k$ is the a-posteriori state estimate at time step k

P is the estimate error covariance given by

$$P_k^- = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_x^-)^T] \tag{5.5}$$

Q is the covariance of process noise w_k

In measurement update stage, an improved estimate is obtained by taking into account a new measurement during the time interval to the a-priori estimate predicted in the time update stage. The equations are presented blow:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad , \tag{5.6}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad , \tag{5.7}$$

$$P_k = (I - K_k H)P_k^- \quad , \tag{5.8}$$

where, R is the covariance of measurement noise

K_k is the Kalman gain which plays a role in setting the weighting factor between the measured value and the updated value when the state vector x_k is updated.

The final estimation algorithm is an iteration of the Predictor-Corrector algorithm, for obtaining improved estimation. The complete work flow diagram of the Kalman filter is showed in Figure 5.1 [26].

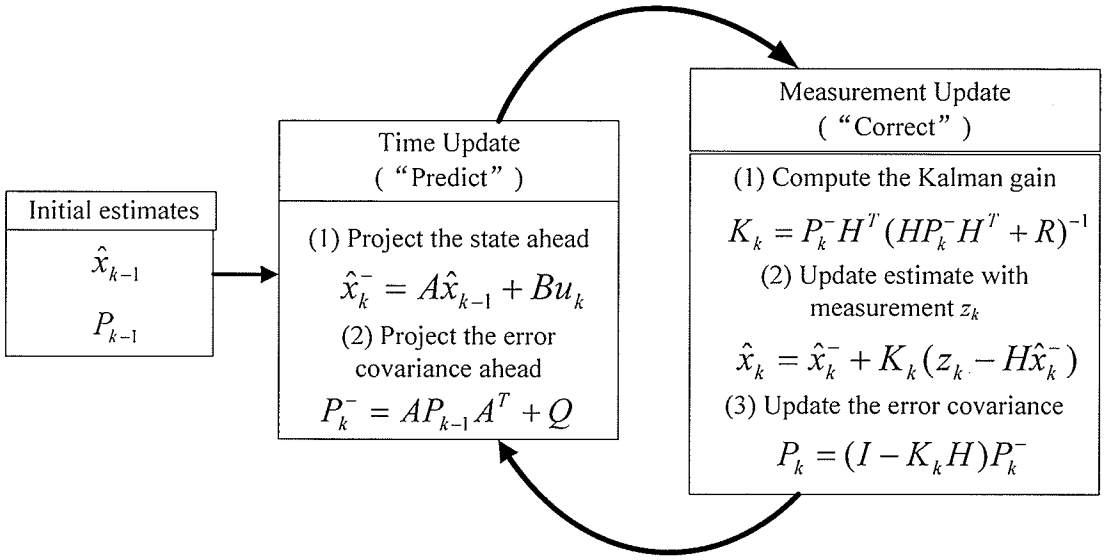


Figure 5.1 A complete work flow of the Kalman filter

5.2 Extended Kalman Filter

The Kalman filtering problem discussed so far addresses the process when the relationship between inputs, state variables and measurements are linear. However, in some cases where this relationship does not always exist, one may then extend the use of Kalman filter by linearizing about the current mean and covariance [26]. In this type of situation, a different Kalman filter known as extended Kalman filter (EKF) has to be used in dealing with the non-linear case.

Similar to the linear system, the general non-linear system model can be represented by two equations:

Process Equation:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \tag{5.9}$$

Measurement Equation:

$$z_k = h(x_k, v_k) \tag{5.10}$$

In the above equations, w and v still represent process and measurement noise as in the linear system, while they may have different values at each time step. The process equation $f(\cdot)$ and measurement equation $h(\cdot)$ are non-linear functions such as power functions, trigonometric functions and reciprocals or some combination of non-linear functions and linear functions. First-order Taylor series [32] are usually used to expand both process equation and measurement equation around a nominal state, which is very close to the actual non-linear state. The Taylor series linearizes the non-linear system to a reasonably accurate linear system. Hence, the linearized Kalman filter algorithm can be summarized as two update stages as in the previous discussion.

Time update stage:

$$\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}, 0) \tag{5.11}$$

$$P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \tag{5.12}$$

A, W, H, V are Jacobian partial derivative matrices which are different at each time step. They can be computed as following:

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0) \tag{5.13}$$

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_j}(\hat{x}_{k-1}, u_{k-1}, 0) \tag{5.14}$$

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}} h(\tilde{x}_k, 0) \tag{5.15}$$

$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{x}_k, 0) \tag{5.16}$$

where, $\tilde{x}$ is the approximate state. Other values have similar definitions as in the Kalman filter algorithm.

In the measurement update stage, taking account of new measurements corrects the state and covariance estimates. This process can be shown in the following equations:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \tag{5.17}$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0)) \tag{5.18}$$

$$P_k = (I - K_k H_k) P_k^- \tag{5.19}$$

So, the basic operation of EKF can be illustrated in a similar manner to the Kalman filter. Figure 5.2 shows the complete work flow diagram of EKF [26].

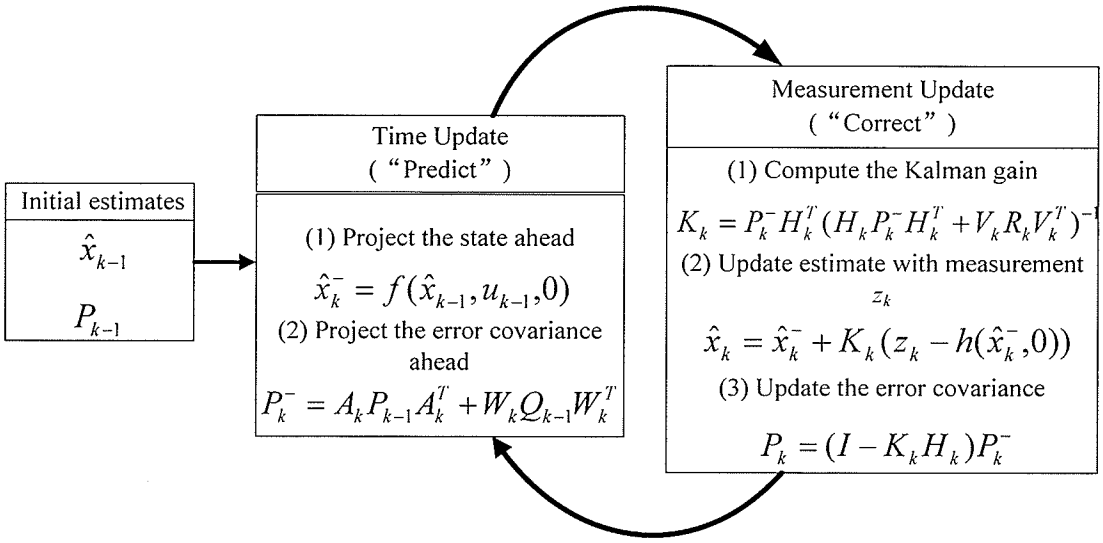


Figure 5.2 A complete work flow of the extended Kalman filter

However, the limitation of the EKF is that the random variables in this filter are no longer Gaussian distributed after the nonlinear transformations. This leads EKF to provide only an approximation to optimal estimation.

5.3 Analysis of the Kalman Filter and the EKF

The most important aspect in implementing a Kalman filter is that the noise in the measurement is Gaussian distributed. From the previous discussion, it can be noted that the signal strength is nearly Gaussian distributed. Hence, we opt to use the Kalman filter in our location estimation. However, in reality the noise in measurement is not ideally Gaussian distributed, which may lead to inaccurate results in estimation.

There are two types of noise sources in the Kalman filter and EKF: process noise and measurement noise. Process noise is caused by the model uncertainty and input disturbances. An accurate model can help reduce process noise. The other noise, measurement noise is produced by the measurement method, equipment and measurement operation.

To run the filters, a number of parameters are required which have different effects on the estimation. These parameters include:

- Parameter matrices: ***A, B, H, W, V***. These matrices decide the system model.
- The control input $\boldsymbol{u}$ at each time step may lead the system model to an appropriate estimate.
- Sensor measurement z at each time step which incorporates and corrects the predicted state value.
- Measurement noise covariance $\boldsymbol{R}$. This is usually determined offline by computing the difference between measurements and known true values. These values give the confidence to the weight of measurements of the filter. The smaller the $\boldsymbol{R}$, the more we trust the measurements in estimation.
- Process noise covariance $\boldsymbol{Q}$. This noise reflects how the filter model performs in estimating. It is usually decided empirically by testing different values of $\boldsymbol{Q}$ until finding a value which can optimize the filter. Generally, the value of $\boldsymbol{Q}$ is very small.
- Error covariance matrix $\boldsymbol{P}$. $\boldsymbol{P}$ indicates the difference between real values and

estimated values. $\boldsymbol{P}$ is generally reduced as more measurements are input. An accurate model will lead the value of $\boldsymbol{P}$ to approach zero after many iterations.

- Kalman Gain $\boldsymbol{K}$. This value depends on other parameters such as $\boldsymbol{R}$, $\boldsymbol{P}$ and $\boldsymbol{H}$. The gain determines the effect of innovation on the estimate. If the measurement noise $\boldsymbol{R}$ is small, the gain will be large which means that filter output provides more confidence on the measurement. If error covariance matrix $\boldsymbol{P}$ is small, the filter will trust the prediction more than the measurement.
- Initial value of state $\boldsymbol{x}_0$ and $\boldsymbol{P}_0$. Before running the filter, we have to know the initial state values of system. Usually we use the first measurement as the initial state $\boldsymbol{x}_0$. As for $\boldsymbol{P}_0$, we can assume a large initial value; it will be reduced as the iterations increases.

The Kalman filter provides the best estimate when some assumptions are satisfied. However, these assumptions are not always valid in reality, for example, noise sources are non-Gaussian distributed, and the system is not linear. In such situations, the Kalman filter still works for estimation, but the results may not be optimal.

5.4 The Kalman Filter for Static Locating

Due to the existence of various noise and interference during the measurement of received signal strength, the RSSIs collected by the mobile user suffer serious error from

their true values. Another noise is caused by the transmitted system, the process noise, also adds a noise component to the received signal strength. As discussed in the previous section, the RSSI distribution is nearly Gaussian. Hence, this noise induced error can be alleviated using the Kalman filter [22]. The Kalman filter can reduce measurement noise and process noise to obtain a reliable value of RSSI from an AP at a fixed location for locating a static user. First, we need to decide two noise parameters Q and R. Process noise Q is generated by the unstable transmitter power, which is negligible when compared to measurement noise R. The measurement noise is the result of multipath fading, interference by other devices and movement of the user in the signal channel. In this estimation, the state values are all scalars and there is no input control modifying the state in the system. With these assumptions, the simplified Kalman filter equations are obtained as follows.

In the time update stage:

$$\hat{x}_k^- = \hat{x}_{k-1}^- \tag{5.9}$$

$$P_k^- = P_{k-1} + Q \tag{5.10}$$

In the measurement update stage:

$$K_k = P_k^-(P_k^- + R) \tag{5.11}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{x}_k^-) \tag{5.12}$$

$$P_k = (I - K_k H)P_k^- \tag{5.13}$$

Experientially, we assume that $Q=1e^{-4}$ and set $R=4$ as estimated by an offline test. The value of noise covariance does not need to be exact because they will converge when many samples are iterated. We used the first noisy measurement of RSSI values as the initial state estimate. The final estimate mainly depends on the number of samples iterated. In our case, we use 20 samples from each AP at a fixed location. The simulation (the code is attached in Appendix 1) of Kalman filter is shown in the Figure 5.3:

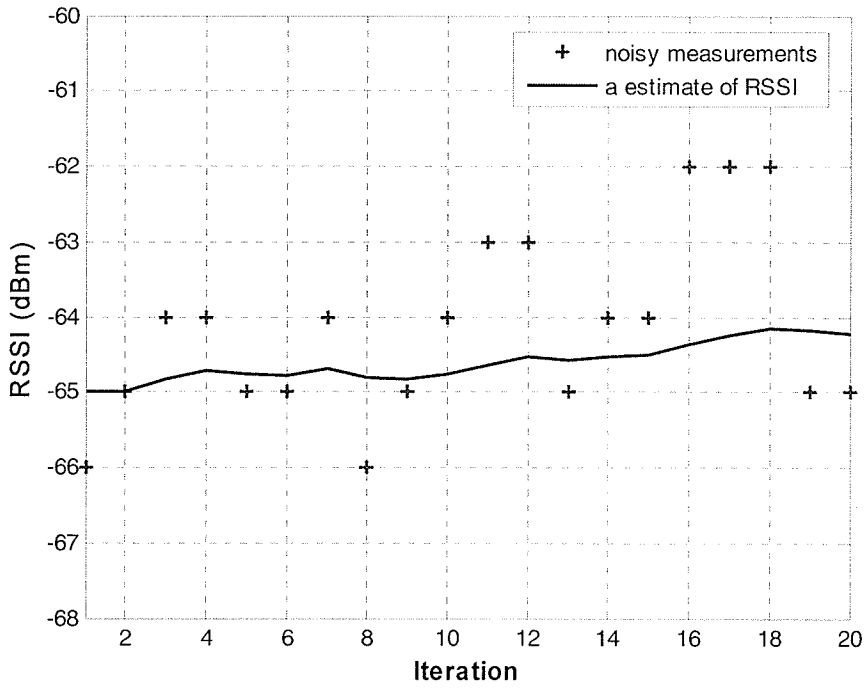


Figure 5.3 Kalman filtering for RSSI

The Figure 5.3 shows that the final RSSI value converges to -64dBm after 20 samples are iterated. We took this value as the final filtered result. Although this estimate value still has some error with respect to the true value, the Kalman filter reduces the total error. That is, the difference between the filtered estimate and true value is less than either the

difference between the mean value of 20 samples or one single signal sample and true value.

5.5 The Kalman Filter for Dynamic Tracking

Similarly, various noise sources in the indoor environment may lead to an error when the system is tracking a mobile object. The noisy measurement makes the mobile object appear as if it is "jumping" rather than "walking" around within the area. Figure 5.4 shows a tracking trajectory of a moving object by the location system without the Kalman filter. From this figure, we can see that the object jumps a lot between consecutive time steps. Obviously, it does not look like a normal moving trajectory. However, considering the characteristics of a moving object, its current location must be related to its previous location, which means that the location of next time step should be in a close vicinity of the current location. The Kalman filter can build a relationship between the current locations with the previous locations. In this thesis, we use Kalman filter again to track the mobile user more accurately.

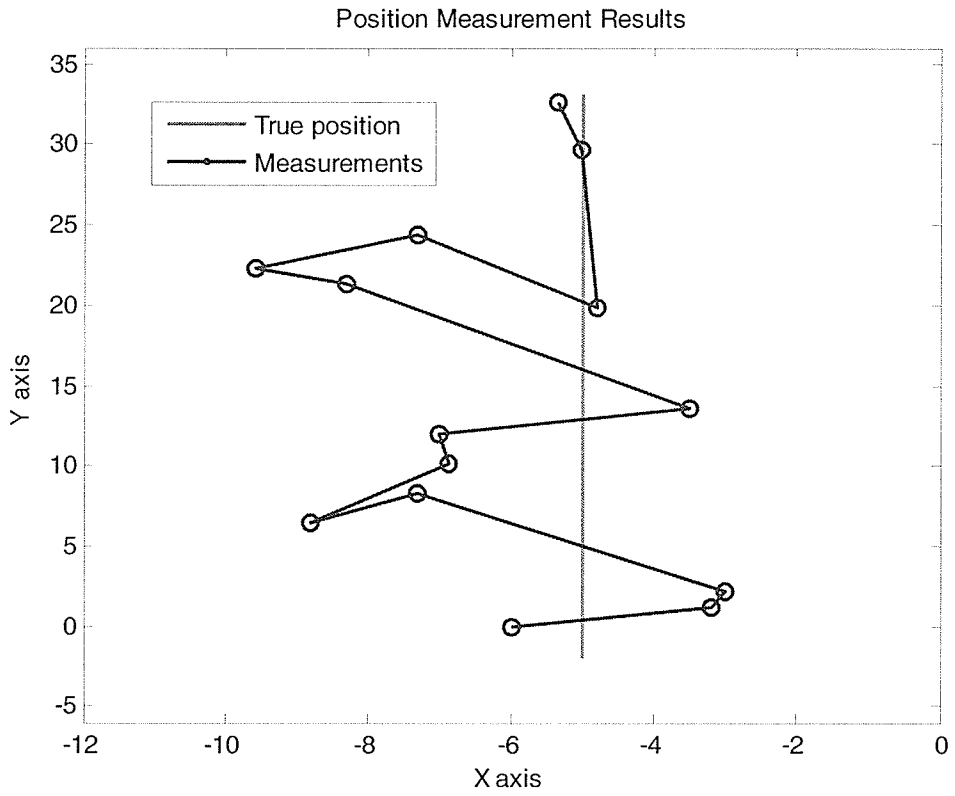


Figure 5.4 Tracking moving object

We consider the object which moves in a 2D space. The measurements are the object's positions in Cartesian coordinates system x and y which can be obtained from the system's locating function. If the time duration of observation is short enough, the movement of the object can be considered a linear motion governed by Newton's laws of motion. Based on the Newton's laws of motion, we can build a motion model for dynamic tracking. In other words, some parameters such as position, velocity and acceleration are needed to model a moving object. We denoted $s_k = (x_k, y_k, \dot{x}_k, \dot{y}_k, \ddot{x}_k, \ddot{y}_k)^T$ as the state of the moving object. Where, x_k and y_k are

the object's position in Cartesian coordinates system at time step k; $\dot{x}_k$ and $\dot{y}_k$ are the velocities of the object at both coordinate axes and $\ddot{x}_k$, $\ddot{y}_k$ are the acceleration, respectively. $z_k = (xm_k, ym_k)^T$ is the measurement of object's location. A common method [33] to define the state equations is:

$$s_k = A \cdot s_{k-1} + w_{k-1} \tag{5.14}$$

$$\begin{bmatrix} x_k \\ y_k \\ \dot{x}_k \\ \dot{y}_k \\ \ddot{x}_k \\ \ddot{y}_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & t & 0 & \frac{t^2}{2} & 0 \\ 0 & 1 & 0 & t & 0 & \frac{t^2}{2} \\ 0 & 0 & 1 & 0 & t & 0 \\ 0 & 0 & 0 & 1 & 0 & t \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \dot{x}_{k-1} \\ \dot{y}_{k-1} \\ \ddot{x}_{k-1} \\ \ddot{y}_{k-1} \end{bmatrix} + w_{k-1}$$

$$z_k = H \cdot s_k + v_k \tag{5.15}$$

$$\begin{bmatrix} xm_k \\ ym_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_k \\ y_k \\ \dot{x}_k \\ \dot{y}_k \\ \ddot{x}_k \\ \ddot{y}_k \end{bmatrix} + v_{k-1}$$

In above equations, only the positions can be obtained from the measurements. Velocities and accelerations have to be calculated by the known positions' first derivatives and second derivatives, respectively. The disadvantage of this method is that precise

information like position, velocity and acceleration are needed. However, the noise in the measured positions is not ideally Gaussian distributed, which leads the velocities and accelerations derived from non-linear transformation of positions to be non-Gaussian distributions. In this situation, the accuracy of estimation by the Kalman filter will seriously collapse. Therefore, it is desirable to combine information from other measurement sources other than position to improve the accuracy. Here, we consider adding control inputs such as accelerations and velocities to the Kalman filter to correct the location prediction. The accelerations and velocities can be obtained from a different measurement source like an accelerometer. Hence, we can rewrite the Kalman Filter equations as following:

$$s_k = A \cdot s_{k-1} + B \cdot \dot{s}_{k-1} + C \cdot \ddot{s}_{k-1} + w_{k-1} \tag{5.15}$$

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} \cdot \begin{bmatrix} \dot{x}_{k-1} \\ \dot{y}_{k-1} \end{bmatrix} + \begin{bmatrix} \frac{t^2}{2} & 0 \\ 0 & \frac{t^2}{2} \end{bmatrix} \cdot \begin{bmatrix} \ddot{x}_{k-1} \\ \ddot{y}_{k-1} \end{bmatrix} + w_{k-1}$$

$$z_k = H \cdot s_k + v_k \tag{5.16}$$

$$\begin{bmatrix} xm_k \\ ym_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_k \\ y_k \end{bmatrix} + v_{k-1}$$

$$\dot{s}_k = \dot{s}_{k-1} + \ddot{s}_{k-1} \cdot t \tag{5.17}$$

where, z_k is the positions measurement at time step k; $\ddot{s}$ is the accelerations

measurement from the accelerometer and $\dot{s}$ is the velocity calculated by the product of acceleration $\ddot{s}$ and time slot duration t. A is the transmit matrix of the system; B and C are the control matrices of the system.

6. Experiments and Discussion

In this section, some experiments were conducted in the real world indoor environment to test the performance of a location estimation system. The experiments included two parts: static locating and dynamic tracking.

6.1 Setup

The testbed is a typical WLAN-based area which is located in the main floor of the Engineering Building E-3 at the University of Manitoba. The layout of the main floor is shown in Figure 6.1. The floor has dimension of 126 ft by 89 ft. In this experimental area, there is average of 43 APs detectable covering the test area depending on the training locations. All these fixed APs are deployed in this building and operate on channels 1, 6 and 11 of IEEE 802.11 standard. The locations of the APs are not required to be known for RSSI-based location estimation. In the experimental area, a total of 103 training points were taken. The granularity of each training point is 6 feet selected, due to the existing floor tiles in the building. The detail layout of training points can be seen in Figure 6.1. The training points are represented by black dots in the floorplan and we used a 2-D reference coordinate to locate them with an origin point (0, 0) at the left lower side of the figure. The RSSI data are collected by an ASUS laptop with a Pentium M CPU, a TP-LINK 510G wireless network card, and the Windows XP operating system. The RSSI

features were obtained by using publicly available network software, Netstumbler. Netstumbler is a free Windows tool to detect WLANs using IEEE 802.11. This software can be downloaded from [34]. Netstumbler's maximum sampling rate of 2 samples/second was adopted. In the process of the experiments, the laptop is held by a person with the height to the ground of approximately 3 ft. The data collections were taken after 10:00 p.m. to reduce the effects of human movement in the test area.

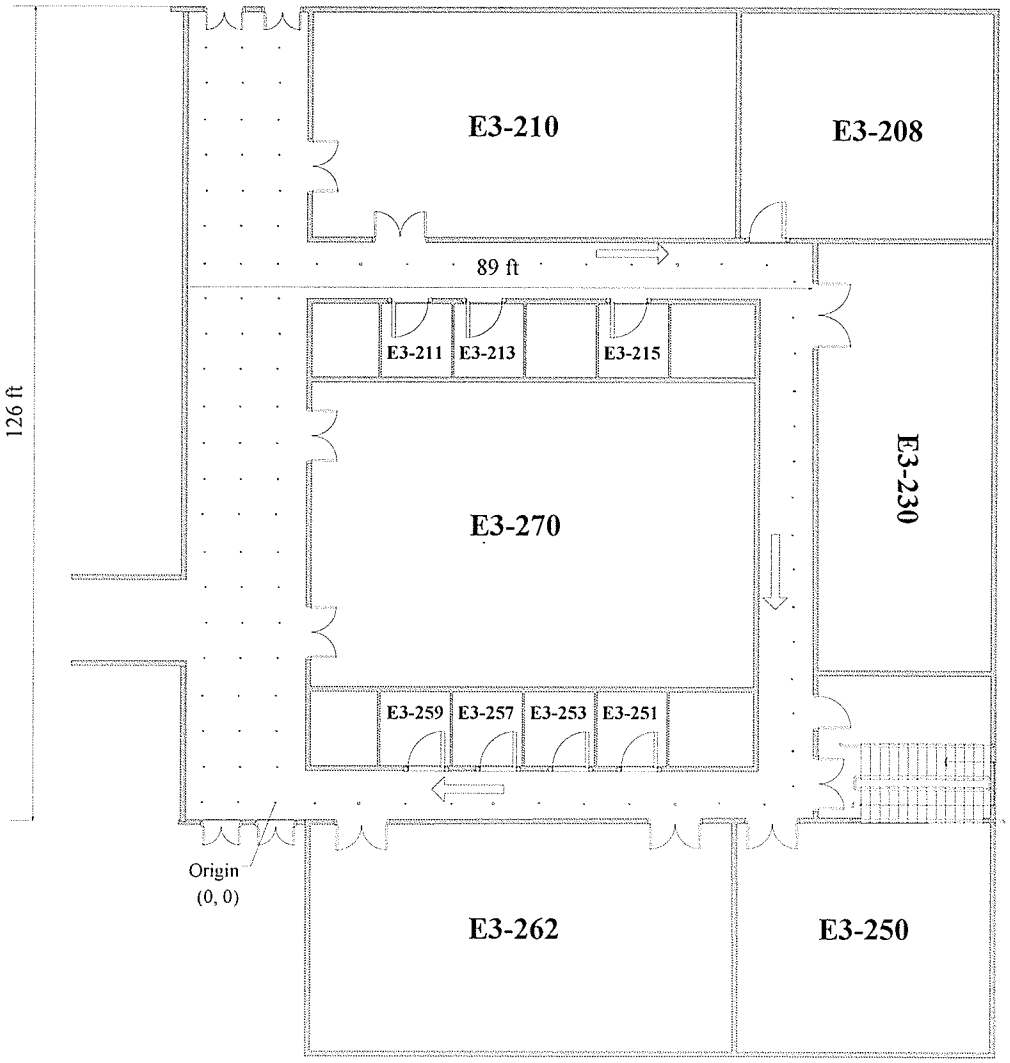


Figure 6.1 Layout of the experimental environment

6.2 Offline Phase

The objective of the offline phase is to construct a radio map. As discussed in Section 4, the non-parametric distribution method was adopted in our experiments. The orientation of training was kept the same and is indicated by the arrow in the Figure 6.1 during the test to eliminate the effects of orientation. In the test area, a total of 43 APs are available for positioning which is much more than the minimum requirement of 3 APs. Intuitively, the more APs used in the location system, the higher estimate accuracy tends to be. In fact, increasing the number of APs contributes little to the accuracy of estimation while leading to high cost of computation. Some research [14, 15] focuses on the selection of APs in location estimation. This thesis does not address this issue in detail. Therefore, 6 APs were used to provide a complete coverage for the experimental area. In the offline phase, the RSSI samples from each selected AP were measured 20 times at one training point. Such that 20 samples were collected every time and they were classified as a group. Then, the Kalman filter was applied to each group of samples to filter the noise in the RSSI samples. As such, the process was repeated for 20 groups, so 20 such filtered results were obtained. This process was repeated for each AP at each training point. Hence, a total of $120 = 20(\textit{results}\text{ for each AP}) \times 6(\textit{number}\text{ of APs})$ fingerprints are stored for each training point used to construct the radio map. Ultimately, the same measurements were taken for every training point to complete the building of the complete radio map.

6.3 Online Phase

In the online phase, a group of 20 RSSI samples from each AP are measured at an unknown location within the experimental area. Similar to the offline phase, we also used the Kalman filter to remove noise from these samples. After repeating this procedure for each AP, we can obtain 6 such filtered estimates. Then, these filtered results were used to query the radio map built in the offline phase to calculate the location. The location determination compares the product of marginal probability distribution with different training points. The training points which have larger product values may be considered as estimates of the location of mobile user. Here, we compound the KNN method with probabilistic method to improve the accuracy of estimation. Instead of the only nearest training point, several nearest training points were chosen depending on the product to estimate the location. In this thesis, we set the $K=4$ for the KNN method. Here, weighting factors can be defined as follow which are similar to those of Eq. (4.8):

$$w_l = \frac{P_l}{\sum_{l=1}^{K} P_l} \tag{6.1}$$

P_l is the product of the marginal probability of training point l and from Eq. (4.15) the product can be defined as:

$$P_l = \prod_{i=1}^{k} P(s_i' | l) \tag{6.2}$$

Hence, the final estimated location is obtained from Eq. (4.7). This equation is rewritten here as Eq. (6.3):

$$< \hat{x}, \hat{y} >= \sum_{l=1}^{K} w_l \cdot < x_l, y_l > \tag{6.3}$$

One problem should be noted in the process of queries. The values of RSSIs in the radio map are not consecutive. This is because there are only 20 filtered RSSI fingerprints in the radio map for each AP, which is not sufficient to express the distribution of fingerprints at a location. For example, in Figure 6.2, a group of signal strengths are approximately Gaussian distributed around the center of -74dBm with the range of fingerprints from -80dBm to -72dBm, while the fingerprints are not consecutive between this ranges, a fingerprint at -79dBm is missing in this case. When the RSSI measured in the online phase cannot find the corresponding fingerprint in the radio map, the value of the product in the Eq. (6.2) will be zero because one of the marginal probabilities is zero, so that we cannot compare the products of different training points properly. To avoid this situation, we empirically set the number of signal samples at the missing strength as a nonzero value, for example 0.5, if there is no sample found at this strength in the process of query. That is, in the example when the online measured signal strength query the radio map at -79dBm, then 0.5 is taken as the number of signal strength in calculating Eq. (6.2).

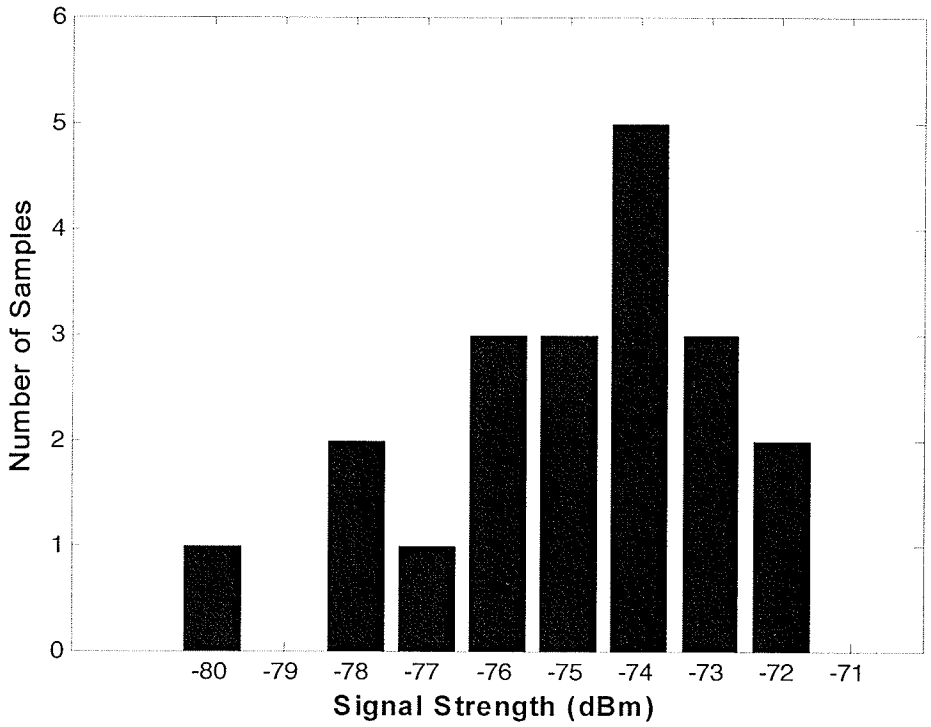


Figure 6.2 Example of missing fingerprinting in radio map

Therefore, the details of locating algorithm can be described in Figure 6.3.

Inputs:

s: Measured RSSI vector from k AP: $s = (s_1, s_2, ..., s_k)$;

N: training points l;

M: Radio Map: construct by fingerprinting at N training points.

Outputs:

$< \hat{x}, \hat{y} >$: Estimated location.

Algorithm:

1. Collect 20 RSSI samples from a AP at a training point;
2. Filter these 20 samples by Kalman filter, then obtain a filtered result;
3. Repeat step 1, 2 for 20 times, obtain 20 filtered results;
4. Repeat these procedure for 6 APs, store them in radio map;
5. Repeat the above steps for each training point to complete the radio map;
6. Measure a group of RSSI samples at a unknown location;
7. Filter them by using Kalman filter;
8. Find the 4 maximum values of P_l by using $P_l = \prod_{i=1}^{k} P(s_i' | l)$;
9. Calculate the weighting factors: $w_l = \frac{P_l}{\sum_{l=1}^{K} P_l}$;
10. Estimate the final location: $< \hat{x}, \hat{y} >= \sum_{l=1}^{K} w_l \cdot < x_l, y_l >$.

Figure 6.3 Outline of location algorithm

6.4 Results and Discussion

In our experiments, we first estimated more than 20 locations of the static mobile user at several different locations covered by the location system. We executed the experiments

in two setups. The first setup used the non-parametric probabilistic method proposed in [5] to estimate locations. The second one is our location system depicted in Figure 6.3 which added the Kalman filter process. Usually, we compared these two setups from two criteria: accuracy and precision. Accuracy describes the closeness of the measurement to the true value. Here, it is considered as the mean error distance which is the mean distance between the estimate location of mobile user and the true location. Precision considers the reproducibility of many measurements. In this experiment, the cumulative distribution function (CDF) of error distances is considered to depict measurement precision. Table 6.1 shows the average error distance of two setups and Figure 6.4 shows the CDF of error distance for the two setups. The two performance criteria showed that using Kalman filter results in a better estimate accuracy and precision than the non-Kalman filter method.

	Average Error	Standard Variance	Min Error	Max Error
Non-Kalman	4.65 ft	2.24 ft	0.99 ft	9.53 ft
Kalman	3.14 ft	1.45 ft	1.05 ft	5.94 ft

Table 6.1 Accuracy of the Non-Kalman and Kalman setups

From the results of the above table, we can see that in our experimental environment, the average error distance without the Kalman filter is about 4.65 ft while with the Kalman

filter, the average error distance decreased to 3.14 ft. The Kalman filter obviously improved the estimate accuracy by 32% in this case. Compare with most other location systems with the average error distances of around 6 ft to 9 ft, the accuracy results of our system is acceptable for most indoor location applications. From the view of precision, we can see that Kalman filter process has smaller standard variance. Although the Kalman filter may get the minimum error of 1.05 ft which is a little bit larger than the non-Kalman filter, the maximum error is 5.94 ft which is much smaller than the error of non-Kalman filter of 9.53 ft. Therefore, the Kalman filter method provides a smaller error value, which means it is more stable than the non-Kalman filter. Figure 6.4 show the CDF of distance error of two setups.

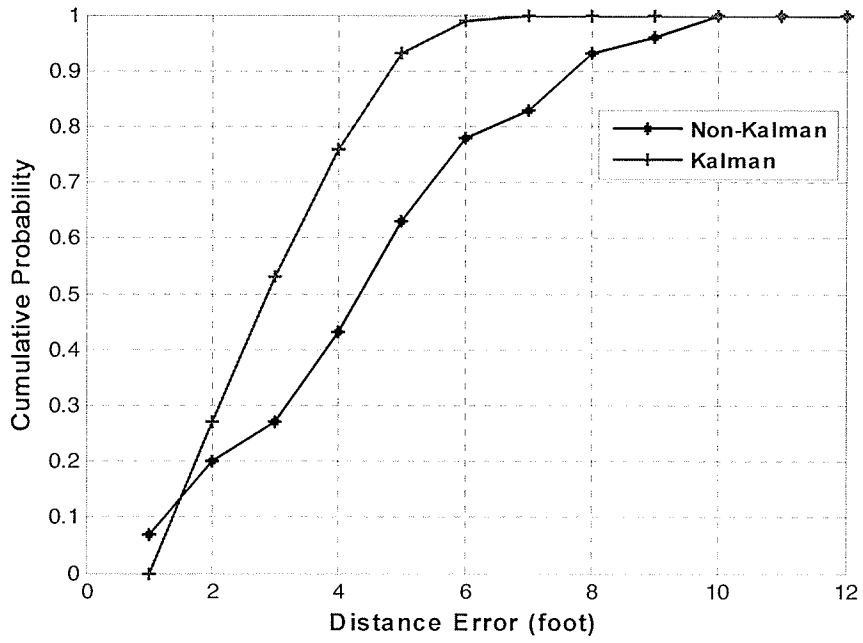


Figure 6.4 Precision of the error distance for the two setups

From the Figure 6.4, more than 90% of distance errors are less than 5 feet by the Kalman filter method, while the non-Kalman filter reports nearly 8 feet to reach the same cumulative probability.

6.5 Tracking Results

Besides locating the static user, the other important function of a location system is to track a moving object. This function is based on the locating function. After locating the mobile user at each time step, we want to see the tracking performance for the moving objects during a period of time.

In the moving situation, a group of RSSI values is measured twice a second during the trip. Since a mobile user moves at a limited speed (on average, people walk at *2-4* feet per second), the user moves little and the signal strength would not change too much over a very short period of time. Considering the computation cost, it is unnecessary to locate the object after every sample obtained. We divided the trip time into many small time slots, each time slot is a length of 2 seconds. So, the location of the tracked object is updated every 2 seconds. To alleviate the effect of noise in measurement, we first use the Kalman filter to remove noise in signal strength measurement. In every time slot (2 seconds), we can obtain 4 RSSIs from each AP. Due to the limitation of measurement hardware and software (the maximum sampling rate is 2 samples/second), only 4 RSSI

samples are used as a group in applying the Kalman filter. In some high level location system, more complex hardware and software are implemented, so more signal strength samples can be gathered in a short time slot. After inputting these samples to the Kalman filter, a group of filtered results is exported to calculate the location of the moving object. The trajectory of the object was described in Figure 5.4. In this section, the calculated locations are usually called measurements. In Figure 5.4, the red line represented the measurements trajectory which has big errors compared to its real trajectory expressed by the green line.

Here, the Kalman filter is used again to improve the tracking accuracy. Now, we need to determine the parameters in equations (5.15) and (5.16). The inputs of the Kalman filter are the location measurements and acceleration measurements. The sequence of locations came from the results of the first Kalman filter, and the accelerations would be obtained from an accelerometer. In this experiment, the group of acceleration measurement values is assumed to be Gaussian distributed. Generally, the measurement error of most available accelerometers is less than *10%*. Hence, it is postulated that the error of acceleration is *8%*. Considering people moved at a constant normal speed in this experiment, we set the acceleration to be about $1.35 feet/s^2$ during the initial time slot and be zero during the following time slots with the measurement error of $0.1 feet/s^2$. With the measurement of accelerations, the velocities of moving object can be calculated. An important parameter for the Kalman filter is the initial value. The initial location of

object $(x_1, y_1)^T$ usually take the first measurement value, that is $z_1 = (xm_1, ym_1)^T$. The initial velocity is set to be zero in both orientations. The time slot t is 2 seconds. The variance of location measurement is similar to the error of static locating, here we set:

$$R = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

and the process noise Q is set to small elements matrix which is:

$$Q = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

The initial error matrix covariance P_0 does not have too much effect on the Kalman filter, as it will converge as more samples are iterated. Like most Kalman filter applications, we set P_0 as:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hence, the details of tracking algorithm are described in Figure 6.5.

Inputs:

s: Measured RSSI vector from k AP: $s = (s_1, s_2, ..., s_k)$;

ac_k: Measured acceleration vector from a portable accelerometer.

M: Radio Map: construct by fingerprinting at N training points.

Outputs:

T: Estimated trajectory.

Algorithm:

1. Measure a group of RSSI samples for a moving object;
2. Divide these samples to sub groups every 2 seconds;
3. Filter RSSI samples by the Kalman filter;
4. Using these filtered samples, calculate the location of moving object at every 2 seconds;
5. Combine calculated locations and measured accelerations to a Kalman filter, the estimated trajectory T is obtained.

Figure 6.5 Outline of tracking algorithm

Incorporating the parameters above and the algorithm described in Figure 6.5, we can plot the tracking results in Figure 6.6:

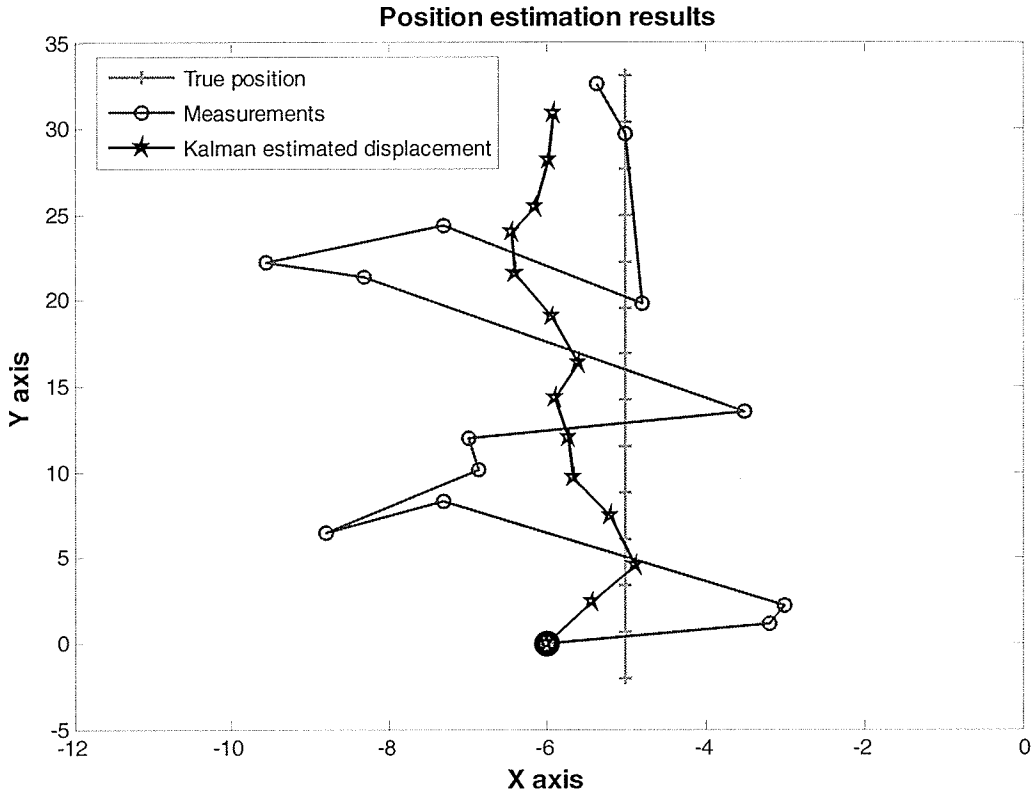


Figure 6.6 The tracking results for a moving object

In figure 6.6, the green straight line represents the actual path of moving object, which is pre-known by the user. The red circle curve is the trajectory of connecting the measured locations (red circles) at every 2 seconds. The blue star curve is the estimated trajectory combining measured locations and measured accelerations by the Kalman filter, the blue stars indicates filtered locations. The start point is expressed by a black circle, which is the initial measurement location. Figure 6.6 shows that the blue curve is smooth and much closer to the green line than the red curve. That means the Kalman estimated displacement can improve the tracking accuracy by a considerable degree. This is

because that the measurements only calculate the location at each time step, while the Kalman filter builds a relationship between the next location and last location. The predicted location is always in the vicinity of the last one. Also, the control inputs accelerations and velocities from other source measurements will correct the Kalman filter prediction at each time step.

6.6 Error Analysis

In the experiments, the laptop was held to move in the test area from point (-5, -2) to point (-5, 35) and 14 measurements were obtained during this trip. The Gaussian distributed accelerations were simulated from Matlab. We calculated the distance between measurements (or Kalman displacements) and true locations as the errors. Based on calculating these 14 locations, we also find that the Kalman estimated displacements have less error than the measurements. The average distance between the measurements and true trajectory is about 3.06 feet, while the average distance about Kalman estimated is 1.44 feet. The details of error are shown in Table 6.2:

	Average Error	Min Error	Max Error
Measurements	3.06 ft	0.58 ft	7.9 ft
Kalman Estimated	1.44 ft	0.71 ft	2.24 ft

Table 6.2 Accuracy of two setups

From Table 6.2, we can see that even though the minimum error of measurements is smaller than the Kalman estimated displacement; the average error of measurements is around 2 times that of the Kalman estimate. Hence, the Kalman estimate improves the accuracy in tracking a moving object.

With the application of the Kalman filter in static locating and dynamic tracking, the measurements show an improvement in accuracy. However, error still exists in the system. There are several errors that contribute to the location system which include:

- Measurement errors: These errors include the measured RSSI data and measured acceleration data. Various noises in the test environment and inexact test equipments caused these error;
- Algorithm errors: The error derived from the character of location algorithm and parameters setup in the location algorithms. For example, the number of APs, the granularity of training points and the sampling rate all can lead to the error;
- Model error: In this thesis, the moving object is modeled as a linear system. In reality, the object may move irregularly, a more complicated motion model may be needed to describe it.

7. Conclusion and Future Work

In this thesis, the issue of location aware computing is investigated by reviewing a number of location methods such as proximity, triangulation and fingerprinting. The RSSI is the best metric to measure signal information compared to metrics like TOA, AOA and TDOA for an indoor environment. Three main models are used for RSSI based location system. Among these three models, the empirical fingerprinting model is preferred because it provides better accuracy, although it costs more in computations. The Kalman filter is used to filter the RSSI measurement in locating mobile users and can also improve the tracking accuracy when combined with the information obtained from another measurement source (accelerometer). An experiment deployed at the University of Manitoba showed that an accuracy of 90% with precision of 5 feet is achieved when using Kalman filter comparing an accuracy of 90% with precision of 8 feet of non-Kalman filter. The average error also decreased 32% by using the Kalman filter. Moreover, the tracking accuracy can be improved by using the Kalman filter. These experimental performances were affected by various training parameters such as the number of samples, the number of APs, the spacing of the radio map and the direction of a mobile object.

Future work will consist of applying other techniques or algorithms to improve estimation accuracy. The Kalman filter is not the perfect filter in the location system

because the signal in the real indoor environment is not ideally Gaussian distributed and the motion of object may not follow a linear model over some time period. Other filters such as extended Kalman filter and particle filter may achieve more accurate results.

In this thesis, only one orientation of the reference point in the training phase was considered. In fact, the mobile object may be located at a point with different orientations, which has its corresponding signal information. Hence, in future experiments, the RSSI measured for different orientations at a reference point should be considered. Moreover, other training parameters should be configured to obtain the best accuracy according to the real test environment.

Some other techniques will also contribute to improve the location system. The floor layout may help to determine the location of a mobile user. Mobile objects cannot cross a wall or door in a short time, hence, the location of object is determined within a restricted area. The sampling rate is another issue in the location system. The rate can be set to speed-adaptive. The location system will detect the speed of mobile object and adjust the sampling rate according to its speed. The faster object moves, the higher rate the system samples. In aspects of computation cost, reductions can be achieved using clustering location algorithms.

The other measurement source (acceleration) was obtained through simulation with

Matlab in this thesis. In fact, this measurement can be obtained in an accelerometer. Accelerometers are mounted in many devices such as iPhones, PDAs and thinkpad laptops. Integrating the accelerometer and location system into a whole system is the future work of ongoing location aware computing. The measurements from another device will reduce the computation costs and improve the estimate accuracy.

References:

[1]. ANSI, INCITS 371.1 "for Information Technology-Real Time Locating System (RTLS), part 1: 2.4-GHz Air Interface Protocol", 2003.

[2]. IEEE 802.11 Working Group, http://www.ieee802.org/11/, (available as of April 2009).

[3]. A.Kotanen, M.Hännikäinen, H.Leppäkoski, T.D.Hämäläinen, "Position with IEEE 802.11b Wireless LAN", *Personal, Indoor and Mobile Radio Communications, 2003, vol.3, pp.2218-2222*, Sept, 2003.

[4]. K.Pahlavan and P.Krishnamurthy, *Principles of Wireless Networks: A Unified Approach*, Prentice Hall PTR, Upper Saddle River, New Jersey, 2002.

[5]. M.Yousief and A.Agrawala, "The Horuss WLAN Location Determination System", *Proceedings of the Third International Conference on Mobile Systems, Applications, and Services (MobiSys 2005), pp.205-218*, Seattle, WA, USA, June 2005.

[6]. J.Small, A.Smailagic, and D.P.Siewiored, "Determining User Location For Context Aware Computing Through the Use of a Wireless LAN Infrastructure". *Project Aura Report*, Carnegie Mellon University, Dec 2000.

[7]. T.S.Rappaport, *Wireless Communications Principles and Practice*, Second Edition, Prentice Hall Communications Engineering and Emerging Technologies Series, 2002.

[8]. B.Sklar, "Rayleigh fading channels in mobile digital communication systems: I. Characterization," *IEEE Communications Magazine, vol.35, pp.90-100*, July, 1997.

[9]. K.Kaemarungsi and P.Krishnamurthy, "Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting", *the 1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous 2004), pp.14-23*, Boston, Massachusetts, USA, August 22-25, 2004.

[10]. J.Yin, Q.Yang and L.Ni, "Adaptive Temporal Radio Maps for Indoor Location Estimation", *In Third IEEE International Conference on Pervasive Computing and Communications (PerCom 05'), pp.85-94*, March, 2005.

[11]. P.Bahl and V.N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," *in Proc. IEEE INFOCOM, 2000, vol.2, pp. 775-784,* 2000.

[12]. M.Heidari, "A Testbed for Real-Time Performance Evaluation of RSS-based Indoor Geolocation Systems in Laboratory Environment", *Master T*hesis, online source: http://www.wpi.edu/Pubs/ETD/Available/etd-050405-144528/unrestricted/Thesis_heidari.pdf, (available as of April 2009) Worcester Polytechnic Institute, April 2005.

[13]. A.Kushiki, K.N.Plataniotis and A.N.Venestanopouls, "Kernel-based Positioning in Wireless Local Area Networks", *IEEE Transactions on Mobile Computing, vol.6, no.6, pp.689-705*, 2007.

[14]. M. Youssef, A. Agrawala, and A.U. Shankar, "WLAN Location Determination via Clustering and Probability Distributions", *in Proceedings of the First IEEE Int. Conf. on Pervasive Computing and Communications, 2003, pp. 143-150*, March 2003.

[15]. Y. Chen, Q. Yang, J. Yin, and X. Chai, "Power-Efficient Access-Point Selection for Indoor Location Estimation", *IEEE Trans. on Knowledge and Data Engineering, vol. 18, no.7, pp.877-888*, July 2006.

[16]. B.Pearn, "Location Awareness in Wireless Networks", *Master Dissertation*, Online source: http://eprints.utas.edu.au/165/1/Thesis.PDF, (available as of April 2009) University of Tasmania, November 2004.

[17]. C.M.Bishop, "Neural Networks for Pattern Recognition", Oxford University Press, 1995.

[18]. R.Battiti, T.L.Nhat and A.Villani, "Location-Aware Computing: A Neural Network Model for Determining Location in Wireless LANs", *Technical Report DIT-02-0083*, University of Trento, Feb 2002.

[19]. F.Evennou, F.Marx and E.Novakov, "Map-aided indoor mobile positioning system using particle filter", *IEEE Wireless Communications and Networking Conference, vol.4, pp. 2490-2494*, March 2005.

[20]. C.Guvenc, R.Jordan Abdallah, and O.Dedeoglu, "Enhancements to RSS Based Indoor Tracking Systems Using Kalman Filters", *GSPx & International Signal Processing Conference*, Dallas, TX, 31 Mar – 3 April 2003.

[21]. P.O.Gutman and M.Velger, "Tracking Targets using Adaptive Kalman filtering", *IEEE Transactions on Aerospace and Electronic Systems, vol. 26, no. 5, pp. 691-699,*

1990.

[22]. Kumar K.R, Apte.V, Powar.Y.A, "Improving the Accuracy of Wireless LAN Based Location Determination Systems Using Kalman Filter and Multiple Observers", *IEEE Wireless Communications and Networking Conference, vol.1, pp.463-468,* 2006.

[23]. J. Latvala, J. Syrjärinne, S. Niemi and J. Niittylahti, "Patient Tracking in a Hospital Environment Using Extended Kalman-filtering", *Proceedings of IEEE Middle East Workshop on Networking*, Beirut, Lebanon, 2-4 November 1999.

[24]. D. Simon, "From Here to Infinity", *Embedded Systems Programming, vol.14, no.11, pp. 20-32,* October 2001.

[25]. P.S.Maybeck, "Stochastic models, estimation and control", *vol. 141 of Mathematics in Science and Engineering*, Academic Press, 1979.

[26]. G.Welch and G.Bishop, "An Introduction to the Kalman Filter", *Technical Report TR 95-041*. http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf

[27]. E.Cuevas, D.Zaldivar, and R.Rojas, "Kalman filter for vision tracking", *Technical Report B 05-12*, August 2005.

[28]. D.Simon, "Kalman Filtering", *Embedded Systems Programming, vol.14, no.6, pp.72-79*, June 2001

[29]. E.Kiriy and M.Buehler, "Three-state Extended Kalman Filter for Mobile Robot Localization", *CIM Technical Report TR-CIM 05.07,* MiGill University, April 2002.

[30]. E.Ivanjko and I.Petrovic, "Extended Kalman Filter based on Mobile Robot Pose Tracking using Occupancy Grid Maps", *IEEE MELECOM 2004, May 12-15 2004, vol.1, pp.311-314,* Dubrovnik, Croatia, 2004.

[31]. M.M.Olama, S.M.Djouadi, and C.D.Charalambous, "Position and Velocity Tracking in Cellular Networks Using Particle and Kalman Filtering with Comparison", *Proceedings of the 45th IEEE Conference on Decision & Control, pp.1315-1320,* Manchester Grand Hyatt Hotel, San Diego, CA, USA, December 13-15, 2006.

[32]. D.Simon, "Using nonlinear Kalman Filtering to Estimate Signals", *Embedded Systems Design, vol.19, no.7, pp.38-53*, July 2006.

[33]. J.Hartikainen and S.Sarkka, "Optimal filtering with Kalman filters and smoothers — a Manual for Matlab toolbox EKF/UKF", February 25, 2008.

[34]. www.netstumbler.org.

[35]. P.Enge and P.Misra, "Special issue on global positioning system", *Proceedings of the IEEE, vol. 87, no.1, pp.3-15*, Jan 1999.

[36]. M.McGuire and K.Plataniotis, "Dynamic model-based filtering for mobile terminal location estimation", *IEEE Trans. on Vehicular Technology, vol. 52, no.4, pp. 1012-1031*, July 2003.

[37]. R.Want, A.Hopper, V.Falco, and J.Gibbons, "The Active Badge Location System," *ACM Trans. on Information Systems, vol.10, no.1, pp.91-102*, Jan. 1992.

[38]. "First Report and Order 02-48", Federal Communication Commission, Washington, D.C., 2002.

[39]. AeroScout, www.aeroscout.com.

[40]. Cisco, www.cisco.com/en/US/products/ps6386/index.html, (available as of April 2009).

[41]. Ekahau, www.ekahau.com, (available as of April 2009)

[42]. PanGo, http://www.pangonetworks.com/release_34.htm, (available as of April 2009)

[43]. H.Kim, "A Speed-Adaptive Location Estimator for Wireless LAN-based RTLS Systems", *Communications 2006 APCC'06. Asia-Pacific Conference on, pp.1-5,* 2006.

[44]. J.Hightower and G.Borriello, "Location Sensing Techniques", http://seattle.intel-research.net/people/jhightower/pubs/hightower2001techniques/hightower2001techniques.pdf, (available as of April 2009) *University of Washington, WA*, July 30, 2001.

[45]. A.Hatami, B.Alavi, K.Pahlavan, and M.Kanaan, "A Comparative Performance Evaluation of Indoor Geolocation Technologies", *Interdisciplinary Information Sciences, vol.12, no.2, pp.133-146*, 2006.

[46]. Wikipedia, "Accuracy and precision", http://en.wikipedia.org/wiki/Accuracy_and_precision, (As available of May 2009).

[47]. J.Hightower and G.Borriello, "Location Systems for Ubiquitous Computing", *Computer, vol.34, pp.57-66*, Aug 2001.

Appendix A: Kalman filter for static locating in Matlab

```
% Kalman filter in Matlab;
% This code is modified from Andrew D. Straw's Python implementation of
% Kalman filter algorithm.
% The original code is here:
% http://www.scipy.org/Cookbook/KalmanFiltering

% by Lin Zheng
% This code here is used to estimate the "actual" RSSI at an unchanged
% location of our signal map

clear all;
close all;


% initial parameters
z = [85 85 85 85 85 85 85 85 85 85 89 89 89 89 89 89 89 89 95 95]; % noisy
    measurements
N = length(z); % number of measurements
sz = [N, 1]; % size of array
Q = 1e-4; % process variance, Kalman filter does not need this value to be exact

% allocate space for arrays
xhat=zeros(sz);          % a posteri estimate of x
P=zeros(sz);             % a posteri error estimate
xhatminus=zeros(sz); % a priori estimate of x
Pminus=zeros(sz);        % a priori error estimate
K=zeros(sz);             % gain or blending factor

R = 4; % estimate of measurement variance, this value can be obtained by the measured
    noisy RSSI

% initial guesses
xhat(1) = z(1); % usually use the first noisy measurement of RSSI value as the xhat(1)
P(1) = 1.0; %

for k = 2:N
    % time update
```

```
    xhatminus(k) = xhat(k-1);
    Pminus(k) = P(k-1)+Q;

    % measurement update
    K(k) = Pminus(k)/( Pminus(k)+R );
    xhat(k) = xhatminus(k)+K(k)*(z(k)-xhatminus(k));
    P(k) = (1-K(k))*Pminus(k);
end
xhat(N) % display the final estimate value of RSSI
figure();
plot(z,'k+');
hold on;
plot(xhat,'b-')
legend('noisy measurements', 'a estimate of RSSI');
xlabel('Iteration');
ylabel('RSSI');
hold off;
grid on;
```

Appendix B: Kalman filter for dynamic tracking in Matlab

```
close all
clear all

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Ground truth %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Set true trajectory
N=14;          % set the number of measurements
dt=2;          % set the time slot of two consecutive measurements
t=1:N;

% the true trajectory of mobile user is from the position (-5,-2) to the position (-5,35)
% with the real velocity is about (0, 1.35)feet/second;

Vtrue = [0;1.35]; % the real velocity of mobile user, it contains x axis velocity and y axis
      velocity
Strue = zeros(2,N); % set the size of true positions
Strue(:,1)=[-5;-2]; % set the initial position of mobile user
k=2:N;
    Strue(1,:) = -5*ones(1,N); % x axis of true position
    Strue(2,k) = -2+Vtrue(2,:)*dt*(k-1);    % y axis of true position

%    the noisy measurement

Z = zeros(2,N); % Set the size of noisy measurements

% the positions measurements came from the calculation results of fingerprinting

Z(1,:)=[-6,-3.19,-3,-7.3,-8.8,-6.86,-7,-3.51,-8.31,-9.57,-7.31,-4.8,-5.01,-5.36];
Z(2,:)=[0,1.13,2.21,8.35,6.43,10.19,12,13.59,21.38,22.29,24.38,19.8,29.73,32.64];

% position measurements

AC=zeros(2,N);
r1=0.05;    % set the noise of accelerometer, it is usually 5%-10% of the measurement.
AC(1,:)=normrnd(0,0.05,1,N);
```

```
AC(2,:)=[1.36,normrnd(0,0.05,1,N-1)];
V_average(:,1) = [0;0];
for i=2:N;
    V_average(:,i)=V_average(:,i-1)+AC(:,i-1);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Motion equations %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


% Motion equation: Sk = A*Sk_prev + B*Uk_prev+Noise, that is Sk(n) = Sk(n-1) +
    Vk(n-1) * dt
% Of course, V is not measured, but it is estimated
% A represents the dynamics of the system: it is the motion equation
A = [1   0;
     0   1];

% B represents the control matrix which controls the velocity to the system
B=[dt 0;
    0 dt];
% C represents the control matrix which controls the accelerations to the
% system
C=[.5*dt^2 0;
    0 .5*dt^2];

% The error matrix (or the confidence matrix): P states whether we should
% give more weight to the new measurement or to the model estimate
sigma_model = 1;
% P = sigma^2*G*G';
P = [sigma_model^2                    0;
                    0 sigma_model^2];

% Q is the process noise covariance. It represents the amount of
% uncertainty in the model. In our case, we arbitrarily assume that the
% model affects a little by the process noise, here, we set Q as follow:
Q = [0.1 0;
              0 0.1];

% H is the measurement matrix.
% We measure 2D position, the x axis position and y axis position,(Sx,Sy) so
```

```
H = [1 0;
     0 1];

% R is the measurement noise covariance. Generally R can be estimated by
% some known samples, here, we set R as follows:
R = [5   0;
     0   5];


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Kalman iteration %%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Previous state (initial guess): Our guess is that the train starts at the
% first measurement: (-5,-2)
Sk_prev = Z(:,1);

% Buffers for later display
Sk_hat = zeros(2,N);
Sk_hatminus = zeros(2,N);
Sk_hat(:,1) = Sk_prev;
Vk_prev = V_average(:,1);
% Sk_hatminus(:,1) = A*Sk_prev + B*Vk_prev;

for k=2:N

    % Kalman iteration
    Sk_hatminus(:,k) = A*Sk_hat(:,k-1)+B*V_average(:,k-1)+C*AC(:,k-1);
    Pk_minus = A*P*A' + Q;
    T = H*Pk_minus*H' + R;

    % K is Kalman gain. If K is large, more weight goes to the measurement.
    % If K is low, more weight goes to the model prediction.
    Kk = Pk_minus*H'*inv(T);
    P = Pk_minus - Kk*H*Pk_minus;
    Sk_hat(:,k)=Sk_hatminus(:,k) + Kk*(Z(:,k)-H*Sk_hatminus(:,k));
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Plot resulting graphs %%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
figure;
plot(Strue(1,:),Strue(2,:),'g');
hold on;
plot(Z(1,:),Z(2,:),'r-');
plot(Sk_hat(1,:),Sk_hat(2,:),'m');
plot(Sk_hat(1,1),Sk_hat(2,1),'bo','MarkerSize',10);
title('Position estimation results');
xlabel('X axis');
ylabel('Y axis');
legend('True position','Measurements','Kalman estimated displacement');
```