國立交通大學

資訊學院　資訊學程

碩 士 論 文

一個智慧型 WiFi 無線區域網路定位系統之設計

Design of an Intelligent WiFi LAN Positioning System

研 究 生：黃建彰

指導教授：陳耀宗　教授

中 華 民 國 九 十 七 年 七 月

一個智慧型 WiFi 無線區域網路定位系統之設計

Design of an Intelligent WiFi LAN Positioning System

研 究 生:黃建彰　　Student:Chien-Chang Huang

指導教授：陳耀宗　　Advisor：Yaw-Chung Chen

國 立 交 通 大 學

資訊學院　資訊學程

碩 士 論 文

A Thesis

Submitted to College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Computer Science

July 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年七月

一個智慧型 WiFi 無線區域網路定位系統之設計

研究生：黃建彰　　　　　　　　　　　　　　　　　　　　指導教授：陳耀宗

國立交通大學　資訊學院　資訊學程碩士班

中文摘要

隨著網際網路和無線區域網路的蓬勃發展，與位置感知相關的服務和系統變成近年來相當熱門的議題。雖然GPS系統已經廣泛地使用在旅遊和駕駛導航上，但它並不適合應用於室內的環境或高樓的周邊，而且其精確度無法滿足某些需要面對面的服務或應用。此外目前的GPS系統主要是設計用來做純粹的定位或導航，如果要擁有網路傳輸的能力，必須另外結合具有網路傳輸功能的系統。基於以上GPS的不足，有很多定位的方法或技術在過去幾年相繼被提出來，但是其中大部分不是缺乏足夠的精確度，就是架設或維修的過程複雜，導致成本過高而不切實用。

在這篇論文中，我們針對室內的環境，在網路通訊的應用層上架構一個智慧型的無線區域網路定位系統，由於其架構在應用層上，這個系統並不需要修改或更新現有的無線網路設備。 而且，這個系統具有機器學習的能力，所以在延伸定位範圍和系統維護時將會非常實用。其本概念就是，當某個移動裝置被定位出

來之後，它的相關資訊可以加入系統的位置表(Location Tables)中，作為往後定位的參考樣本。但是這樣的方式在應用上會面臨一些問題，我們將會在論文中討論並且提出我們的解決方法。

我們也提出了幾種方法來改進我們系統的精確度、效率和彈性，當系統經過長時間的使用和訓練後，其準確度會相應地提升，但是位置表內相對的會有龐大的樣本資料。如果每次定位時都去搜尋並找出有用的樣本，對系統而言將會是相當大的負擔。在我們設計的系統裡，只有在一開始樣本數量較少時才需要一一比對位置表裡的樣本，一旦樣本成長到相當數量時，並不需要針對每次的定位去做樣本搜尋比對，只有在某些情況下才必須去執行完全搜尋比對動作，這將有助於系統效能的提升。

關鍵字：定位，WiFi 網路，無線區域網路，基地台，訊號強度，收訊強度指示，位置表

Design of an Intelligent Wireless LAN Position System

Student：Chien-Chang Huang Advisor：Yaw-Chung Chen

Degree Program of Computer Science

National Chiao Tung University

Abstract

Due to the fast growing in the user population of Internet and Wireless LAN, location-aware services and systems become a popular topic. Although GPS systems are widely deployed for traveling and driving guidance, it is not suitable for indoor environment. Further, it does not provide sufficient accuracy for face to face applications. Besides, current GPS system was solely designed for positioning purpose, so it needs to be combined with wireless communication service to implement location-aware function for mobile computing purposes. Because of these GPS disadvantages, quite a few indoor positioning schemes were proposed in the past years, but most of them are either expensive or featuring low accuracy.

In this thesis, we proposed an Intelligent WiFi LAN Positioning System which was implemented on the application layer to position the mobile station in indoor environment. Since it was implemented on application layer, it does not need any

change of wireless equipments. Further, it is intelligent because it has the machine learning capability, which is quite useful for extending the positioned area and system maintenance. The basic idea of this learning capability is that the positioned locations can be reused as new samples for future positioning. But there are some potential problems for using all of these positioned locations as future samples. We will discuss the problem and provide our solution in this work.

We also introduce schemes to improve the accuracy, performance and flexibility for our system. After the system has been trained for a certain period, the accuracy will be improved, but there will be a large number of samples in the Location Tables. Searching through the table one by one will be a time consuming task on Location Server, in our system, searching all samples is required only in the beginning or in few special cases.

Keywords: Positioning, WiFi LAN, Wireless LAN, Access Point, Signal Strength, Received Signal Strength Indication, and Location Table

誌謝

以現階段來說，對於WiFi LAN無線網路定位系統的研究，總算是有了初步的結果，並且促成這篇論文的完成。但這並非一己之力所能完成的，而是集眾人努力的成果，從早期的方向摸索、題目確定，到後期的實驗設計及論文完稿，總是有師長及諸位夥伴為我指引方向及提供寶貴的意見，才能將論文如期完成。

首先要感謝的是指導教授陳耀宗老師，由於因為在職的關係，這篇研究所花費的時間多於其他人許多，而老師在這麼長的時間裡仍能本著一貫的認真和熱誠，不時提供建議和意見，讓我獲益良多。另外老師也給予我最大的發揮空間，從想法、設計、實驗到論文完稿，一方面仔細了解我的想法和認知，一方面參與設計和提出建言，在老師的指導之下著實讓這篇研究趨於完整。

其次也要感謝口試委員蔡文能老師、留忠賢老師以及詹家泰老師，雖然只有短短幾天論文的接觸以及一個多小時的口試交流，仍能夠提供相當寶貴以及有力的意見，讓我對這方面有更深刻的體認和認識到自己的想法仍然有不足之處。

另外要感謝的是在事業上一起奮鬥的夥伴同仁們，因為你們的諒解、分擔和代勞，讓我在緊要關頭無後顧之憂。最後謝謝一直陪伴在身邊的家人與朋友們，因為你們的支持，讓我在遇到險阻時能夠持續走下去。

黃建彰

2008 年 7 月

Contents

Abstract in Chinese i

Abstract in English iii

Acknowledgements v

Contents ……... vi

List of Figures.. viii

List of Tables x

Chapter 1 Introduction 1

Chapter 2 Background and Related Works 3

2.1 Background 3

2.1.1 Measurement..3

2.1.2 Similarity..7

2.2 Related Works 8

2.2.1 GPS-Based 8

2.2.2 RFID 9

2.2.3 Cellular-Based..9

2.2.4 Bluetooth (IEEE 802.15) 9

2.2.5 WLAN (IEEE 802.11) 10

Chapter 3 Intelligent WiFi LAN Positioning System...................... 12

3.1 Components in the System 12

3.2 Architecture & Flow Diagram 15

3.3 Tasks in MN Site..15

3.3.1 Predicting by Potential Locations 16

3.3.2 Prediction based on Uniform Linear Moving 16

3.3.3 Predict by Signal Variation Ratio..18

3.3.4 Location Query without MN Predicting 18

3.3.5 Algorithm on MN Site 18

3.4 Location Server Site Tasks 20

3.4.1 Similarity Threshold 21

3.4.2 Confirming MN's Predicted Location 22

3.4.3 Evaluation by AP Signal Set of MN 22

3.4.4 Issues on Positioning and Location Tables Update....................................25
3.4.5 Algorithms on Server Site...27

Chapter 4 Experiment and Numerical Results................................ 31

4.1 Experiment Setup – Simulation Model...31
4.1.1 20x20 m^2 with three Access Points..33
4.1.2 50x50 m^2 with 13 Access Points..34
4.1.3 100x20 m^2 with 12 Access Points..35
4.1.4 20x20 m^2 Extended Case: 36
4.2 Experiment Result 37
4.2.1 Basic Positioning Scenario 37
4.2.2 Extended Area Scenario...52

Chapter 5 Conclusions 53

Bibliography 55

List of Figures

Figure 1: AP packets captured by wireless Sniffer. 4
Figure 2: SS of packets measured at 1m from AP (Avg = -36.88 dbm) 5
Figure 3: SS of packets measured at next room from AP (Avg = -56.12 dbm). 5
Figure 4: SS of packets measured at lower floor from AP (Avg = -81.26 dbm) 6
Figure 5: SS of packets measured at outside (Avg = -91.33 dbm) 6
Figure 6: Error rate of packets SS measured at 1m from AP (Avg = 1.2%) 7
Figure 7: Error rate of packets SS measured at outside (Avg = 1.7%) 7
Figure 8: Proposed system architecture and flow diagram. 15
Figure 9: Prediction based on Moving Vector 17
Figure 10: Algorithm on MN 19
Figure 11: Similarity Threshold Update 21
Figure 12: Evaluating MN's location (Na) by trigonometry. 25
Figure 13: Worst case in positioning 26
Figure 14: Algorithm on Location Server 29
Figure 15: Evaluation_by_APSS( ): Subroutine for Evaluating Location by APSS 30
Figure 16: Environment for simulation model in 20x20 m^2 area. 33
Figure 17: Environment for simulation model in 50x50 m^2 area. 34
Figure 18: Environment for simulation model in 100x20 m^2 area. 35
Figure 19: Environment for extended simulation model in20x20 m^2 area. 36
Figure 20: Average distance error on "Sample Every Location" Case in 20x20 m^2 area. 37
Figure 21: Average distance error on "Sample Every Location" Case in 50x50 m^2 area. 38
Figure 22: Average distance error on "Sample Every Location" Case in 100x20 m^2 area. 38
Figure 23: Average distance error on "Without Learning" Case in 20x20 m^2 area. 39
Figure 24: Average distance error on "Without Learning" Case in 50x50 m^2 area. 39
Figure 25: Average distance error on "Without Learning" Case in 100x20 m^2 area. 40
Figure 26: Average distance error on the case with few Initial Samples in 20x20 m^2 area (1). 41
Figure 27: Average Distance Error on the case with few Initial Samples in 50x50 m^2 area(4). 41
Figure 28: Average distance error on the case with few Initial Samples in 100x20

m^2 area(4). 42
Figure 29: Average distance error on 20x20 m^2 area (One Initial Samples per AP). 42
Figure 30: Average distance error on 50x 50 m^2 area (One Initial Samples per AP). 43
Figure 31: Average Distance Error on 100x20 m^2 area (One Initial Samples per AP). 43
Figure 32: Average distance error on 50x50 m^2 area (33 Initial Samples). 44
Figure 33: Average distance error on 50x50 m^2 area (50 Initial Samples). 44
Figure 34: Average distance error on 100x20 m^2 area (32 Initial Samples). 45
Figure 35: Average distance error on 100x20 m^2 area (50 Initial Samples). 45
Figure 36: Comparison for the case on 20x20 m^2 area (all positions). 46
Figure 37: Comparison on 20x20 m^2 area (after training). 47
Figure 38: Comparison on 50x50 m^2 area (all positions). 48
Figure 39: Comparison on 50x50 m^2 area (after training). 49
Figure 40: Comparison on 100x20 m^2 area (all positions). 50
Figure 41: Comparison on 100x20 m^2 area (after training). 51
Figure 42: Average distance error on Extended Area Scenario. 52

List of Tables

Table 1: Location Information Format in Location Table 14
Table 2: Precision (%) table in 20x20 m^2 area (all positions). 46
Table 3: Precision (%) table on 20x20 m^2 area (after training). 47
Table 4: Precision (%) table on 50x50 m^2 area (all positions). 48
Table 5: Precision (%) table on 50x50 m^2 area (after training). 49
Table 6: Precision (%) tables on 100x20 m^2 area (all positions). 50
Table 7: Precision (%) table on 100x20 m^2 area (after training). 51

Chapter 1 Introduction

In recent years, because of the explosive growth of Internet population, network communication has become the most active topic in computer science. And, based on popular wireless technology and affordable cost, WLAN has become the most demanded equipment of the network environment in enterprise networks or SOHO networks. Following the anywhere availability of Wireless LAN, more and more related services and applications have been proposed based on WLAN environment. One of them is Wireless LAN positioning application. If there is an efficient and high accuracy positioning system, lots of other services and applications can be developed based on this technology. These include guided tour, meeting system, fast roaming and the active security of wireless network.

For a positioning system, the most popular one is GPS (Global Positioning System). But there are some limitations on GPS. cannot be used in indoor environment. So, in indoor environment, WLAN is the best choice for positioning. There are several advertences as follows: 1. Popularity: WLAN has been mass produced into market for several years, it can be found anywhere. 2. Affordable Cost: Even the positioning area without WLAN network, the cost of building an available one is pretty low. The AP (Access Point) in the market may be lower than $100, it also features low cost in maintaining. 3. Combined with network communication: Most of the positioning services need to be combined with network access. WLAN was designed for this purpose originally.

In this thesis, a novel WiFi LAN Positioning System was proposed. It was built based on current WiFi LAN and Application Layer. No modification on H/W or F/W for Wireless NIC and AP is necessary. This means it can be implemented on any brand of NIC and AP. Furthermore, any WiFi standard (802.11, 802.11a, 802.11b, 802.11g, 802.11n…) can be used together in this system.

During our study and development, we used wireless Sniffer to check the signal strength recorded in AP sent packets first. In our measuring, the difference of Signal Strength between each captured packet at the same location was within 5%. It proves that we can use it as our comparing measurement. Then, we built our proposed Intelligent WiFi LAN Positioning System based on this phenomenon. The basic concept of our system is: 1. Measure the AP Signal Strength Set at few selected locations. Use as initial samples in location tables. 2. When MN needs to be positioned, it scans the signal strength of available AP and sends this AP Signal Strength Set to Location Server to query the location. 3. At Location Server site, it receives MN's query and finds out its nearest samples in Location Tables by using Euclidean Distance. 4. Based on the variation between MN AP Signal Strength Set and the nearest location (recorded in Location Tables), Location Server positions the MN current location. 5. The positioned location was replied to MN. And Location Table was updated for it. According to this scheme, more and more samples will be added onto Location Table by mass positions. And the accuracy of system will be improved if there are lots of samples in the Location Table. This positioning system has such Machine Learning ability. That's why we call it "Intelligent". This ability is very useful for system maintenance and extended positioning area.

The remaining part of this thesis is organized as follows. In Chapter 2, we survey the background and related work in positioning techniques. In Chapter 3, we introduce our proposed system and methodology. Chapter 4 presents our simulation models and the experiment result. Finally, we conclude this thesis in Chapter 5.

Chapter 2 Background and Related Works

2.1 Background

2.1.1 Measurement

For wireless positioning, the basic assumption is that the measured values should be similar at the same location in the positioned environment. There are several different types of sensor measurements relevant to wireless network. As Stuart A. Golden introduced in [1], some of them are TOA (Time of Arrival), TDOA (Time Difference of Arrival), AOA (Angle of Arrival) and SS (Signal Strength).

TOA: It directly provides a distance estimate from an MN to a BS. The desired MS location is determined by the intersection of at least three circles formed by multiple measurements between the MN and several BSs. TOA measurements need to timestamp the signals and require synchronization among the BSs

TDOA: Positioning systems can be grouped into two different categories. The First one uses many synchronized transmitters and one receiver. GPS uses this approach. However, this approach does not work well for WLAN. It is because all of the access points do not transmit on the same channel simultaneously. The second one is the inverse from first one. In this case, there are many synchronized receivers but only a single transmitter. Here, the client transmits a signal and the access points all listen and then collaborate to determine the location of the client's transmitter. Although this approach can be applied to WLAN, the access points must all be located on the same channel.

AOA: Because of MIMO technology, WLAN is becoming more attractive. With angle information and distance measurement, client can be located with only a single access point. Although MIMO is currently motivated to increase throughput and range, it can easily be modified to accentuate the direct-path and attenuate other paths.

SS: Distance can be measured by the power loss between the transmitter and receiver. That is, only with two nodes, SS can be used to estimate the distance.

Since our purpose is implementing our system on the application layer of WiFi LAN, we adopt SS in our scheme. The SS information can be obtained from analyzing the packets. Figure 1 shows the packets sent from AP and captured by wireless Sniffer. According to the packet information, we can use the SS information of received packets in MN for our measurements.

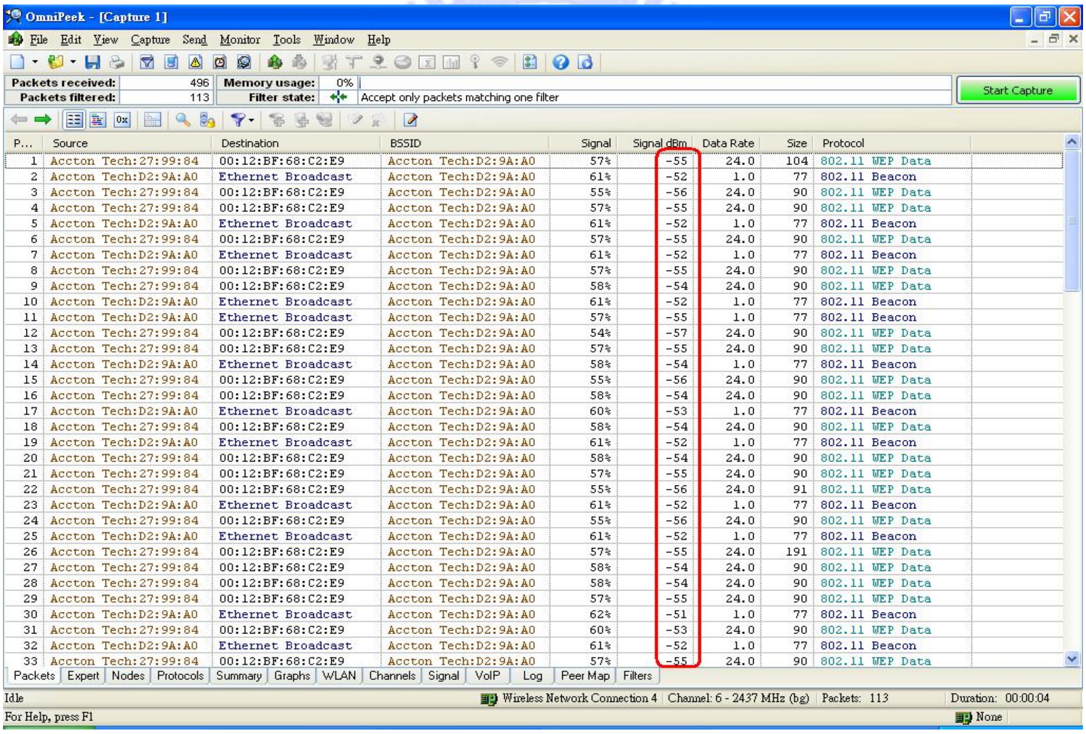

Figure 1: AP packets captured by wireless Sniffer.

We captured AP packets at different locations and make statistics as in Figure: 2, 3, 4 and 5. Each of the statistic data shows in the same location received packets with similar SS.

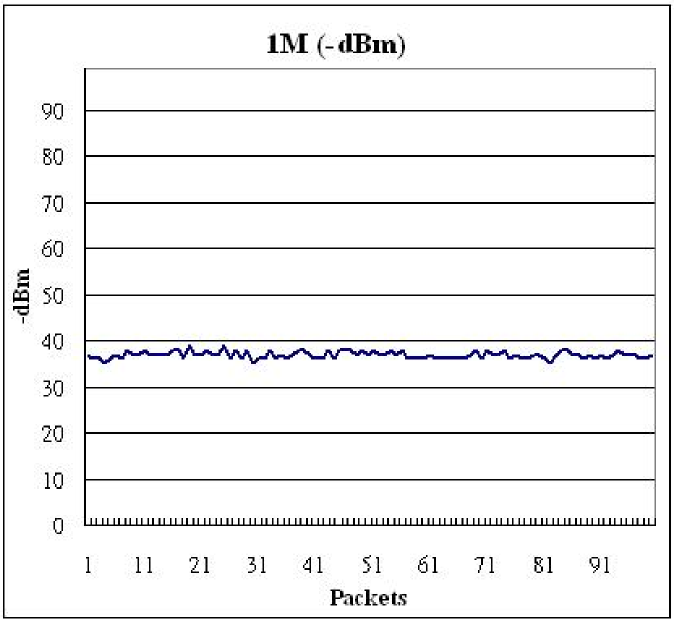

Figure 2: SS of packets measured at 1m from AP (Avg = -36.88 dbm)

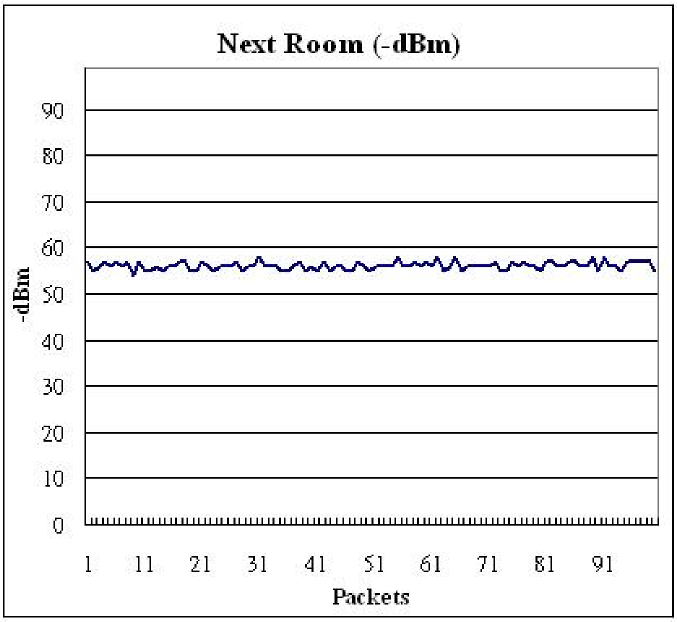

Figure 3: SS of packets measured at next room from AP (Avg = -56.12 dbm).

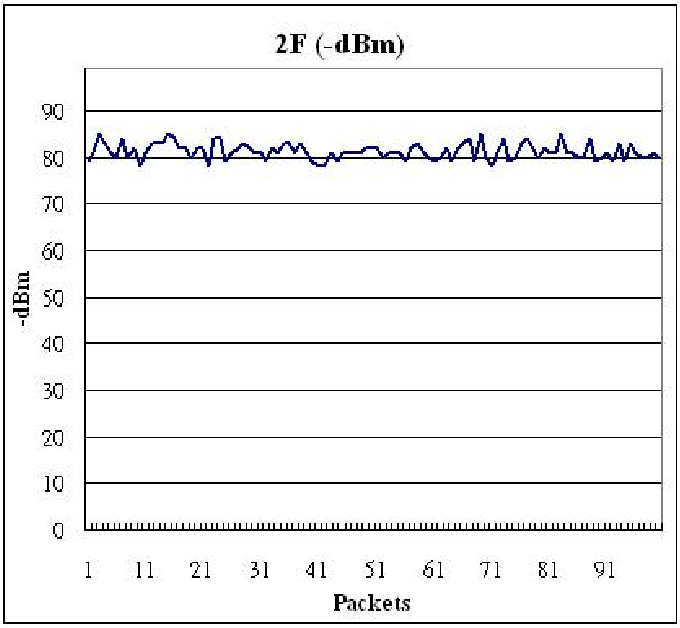

Figure 4: SS of packets measured at lower floor from AP (Avg = -81.26 dbm)

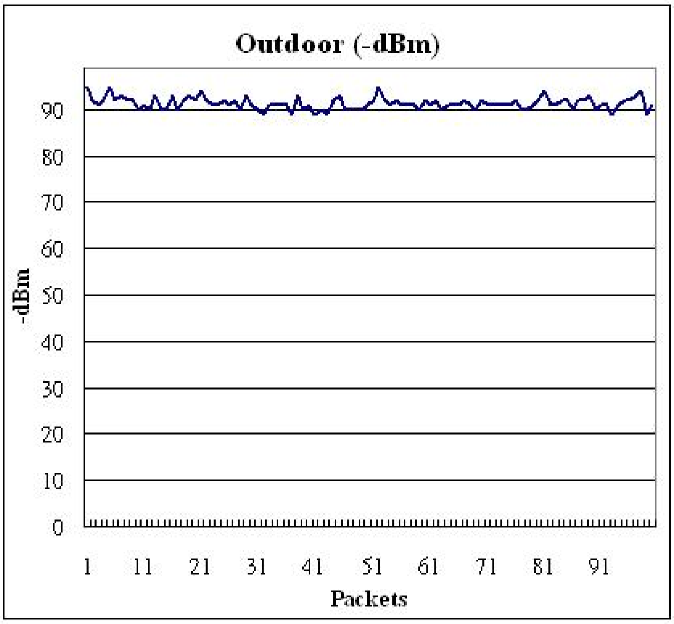

Figure 5: SS of packets measured at outside (Avg = -91.33 dbm)

According to Figure 2 and 5, we calculate the difference (Error Rate) and show the data on Figure 6 and 7. On both figures, the error rate for most of packets is within 5% for the same location. In our simulation, we will use it as the error rate (±5%) for simulated signal strength.

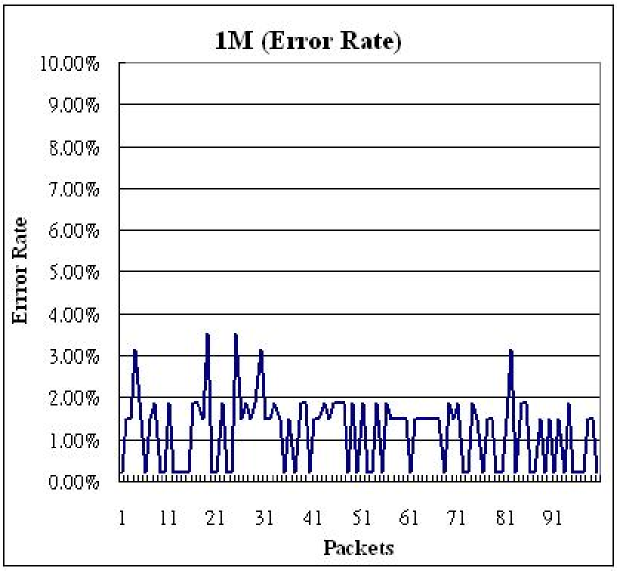

Figure 6: Error rate of packets SS measured at 1m from AP (Avg = 1.2%)

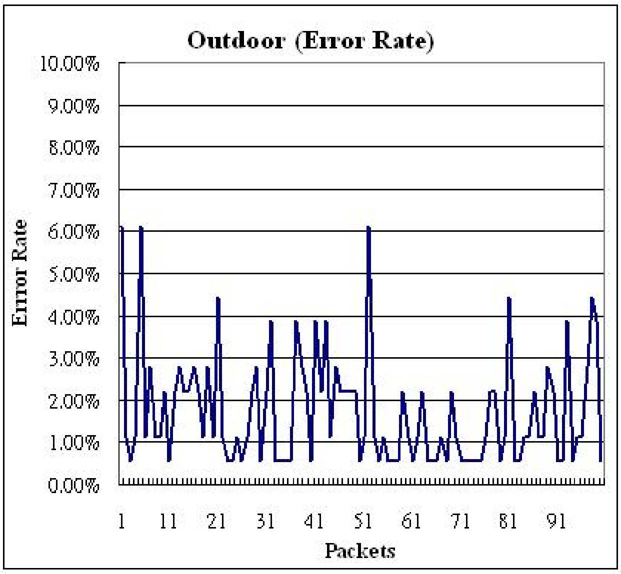

Figure 7: Error rate of packets SS measured at outside (Avg = 1.7%)

2.1.2 Similarity

As Bahl pointed shown in [2], pattern match approaches have higher accuracy than propagation models. We use it as the basic concept in our system. To perform pattern match scheme, we adopt similarity as the indicator in our system. It is calculated by **Euclidean Distance** as below.

- If 2 nodes, a = [a1, a2, …, ad], b = [b1, b2, …, bd], Euclidean Distance is:

$$dist(\mathrm{a},\mathrm{b}) = \sqrt{\sum_{1}^{d} (a_i - b_i)^2} \tag{1}$$

- Using [a1, a2, …, ad] & [b1, b2, …, bd] to represent AP signal strength set of node a and node b,
- The larger the dist (a, b), the lower the Similarity (a, b).

2.2 Related Works

Several positioning systems have been proposed in the literature. There are two basic approaches to designing a wireless positioning system. The first one is to separate the signaling system and network infrastructure which was focusing on wireless location application. The second approach is to use an existing wireless network infrastructure to locate a target. The advantage of the first approach is that the designers are able to control physical specification and the quality of the location sensing results. The advantage of the second approach is that it single and low cost.

2.2.1 GPS-Based

GPS (Global positioning system) is one of the most successful positioning systems in outdoor environments. However, poor coverage of satellite signal for indoor environments decreases its accuracy and makes it unsuitable for indoor location estimation. The other problem is the precision issue. Although there are some schemes of GPS proposed for indoor positioning [3], SnapTrack[1], Atmel[2] and U-blox[3], but most of them either have low accuracy or need additional system.

[1]SnapTrack: http://www.snaptrack.com/

[2]Atmel Corporation: http://www.atmel.com/

[3]U-blox AG. http://www.u-blox. com

2.2.2 RFID

RFID system has several basic components, including RFID readers, RFID tags, and communication between them. The RFID reader is able to read the data emitted from RFID tags. Both readers and tags use defined RF and protocol to transmit and receive data. RFID tags are categorized as either passive or active.

Passive RFID tags operate without a battery. However, its ranges are limited in 1 to 2 meters, and the cost of the readers is relatively high. Active RFID tags are small transceivers, which can actively transmit their ID (or other additional data) in reply to an interrogation. The advantages of active RFID are the smaller antenna and its longer range (>10m). But comparing to GPS or WiFi LAN, this range is still not large enough for positioning. For positioning in a wide area, lots of RFID tags need to be allocated and the cost of maintaining the positioning environment is high. SpotON[4] and LANDMARC[5] are the indoor location sensing systems which were constructed based on this technology.

2.2.3 Cellular-Based

A number of systems have used Global System of Mobile/Code Division Multiple Access (GSM/CDMA) mobile cellular network to estimate the location of outdoor mobile clients. However, the accuracy of the method using cell-ID or enhanced observed time difference (E-OTD) is generally low (in the range of 50–200 m), depending on the cell size. Generally speaking, the accuracy is higher in densely covered areas (e.g. urban places) and much lower in rural environments [6]. Otsasen et al. presented a GSM-based indoor localization system in [7].

2.2.4 Bluetooth (IEEE 802.15)

Bluetooth operates in the 2.4-GHz band. Compared to WLAN, the gross bit rate is lower (1 Mbps), and the range is shorter (typically 10–15 m). On the other hand, Bluetooth is a “lighter” standard, highly ubiquitous (embedded in most phones, personal digital assistants (PDAs), etc.) and supports several other networking services in addition to IP. Bluetooth tags are small size transceivers, like any other Bluetooth device, each tag has a unique ID. This ID can be used for locating the Bluetooth tag. [8].

Antti et al. present the design and implementation of a Bluetooth Local Positioning Application (BLPA) [9]. The accuracy of BLPA is reported to be 3.76 m. A similar work has been done by Hallberg et al. [10].

2.2.5 WLAN (IEEE 802.11)

Wireless local area network (WLAN) operating in 2.4-GHz has become very popular in public hotspots and enterprise environments in the last few years. It is, therefore, appealing to use an existing WLAN infrastructure for indoor location as well, by adding a location server. The accuracy of typical WLAN positioning systems using RSS is 3-30 m, with an update rate in the range of few second.

RADAR: Bahl et al.[2] proposed a location and tracking system—RADAR, which adopts the nearest neighbor(s) in signal-space technique. The authors proposed empirical measurement and signal propagation modeling. Its accuracy is about 2~3 m. In the enhanced RADAR the result is around 2.37~2.65 m and its 90 percentile is around 5.93~5.97 meters.

Horus system [11], [12] proposed a joint clustering technique for location estimation. Each candidate location coordinate is regarded as a class or category. The location with highest likelihood is chosen for result. The experiment results show the accuracy is more than 90% within 2.1 meters. Roos et al. [13] developed a grid-based

Bayesian location-sensing system over a small region of their office building, the result is within 1.5 m over 50% of the time. Nibble [14], used a probabilistic approach (based on Bayesian network) to estimate a device's location.

Chapter 3　Intelligent WiFi LAN Positioning System

In this chapter, we introduce the architecture of our proposed system, describe the methodology and explain its working process.

3.1 Components in the System

The following is the description of each component in our system:

- **LS (**Location Server) - The most important component in our system. It is responsible for the following:

  - Maintain Location Tables: The Location Server manages the Location Table based on the AP Signal Strength Set measured by mobile nodes (MNs). It includes AP Signal Strength Set provided by the initially sampled and positioned MN.

  - Confirm MN's predicted location: If MN's positioning query is performed with MN's predicted location, check if it is a correct location based on similarity threshold. If not, performs location server prediction as next one.

  - Evaluate current MN location: If the positioning query is performed without MN's predicted location or failure in confirming MN's predicted location, it evaluates the new MN's location based on our scheme.

- **Posited AP:** Wireless base stations with location information recorded on Location Server, and broadcast beacons continuously. If necessary, beacon interval can be configured shorter.

- **Initial Samples:** This is to measure AP Signal Strength Set at known location address. It provides the initial samples with AP Signal Strength Set.

- **MN:** Mobile Nodes to be positioned.
  - Maintain traveling history (4 location addresses and 1 AP Signal Strength Set): Four location addresses are used for **Prediction by Uniform Linear Motion**. One AP Signal Strength Set is used for **prediction using signal variation ratio**. Each of them will be discussed in Section 3.3.2 and 3.3.3.
  - Scan wireless channels and get the AP Signal Set: At each position, MN needs to perform all channels scanning to measure each available AP Signal Strength. It may spend 2~3 seconds to scan all channels. This time duration can be shortened by modifying wireless driver.
  - Predict possible current location of MN: Based on our scheme, MN can predict its location first and send it to Location Server for confirmation. Through this feature, it is not necessary to perform fully search on Location Table to find out the most suitable sample.
- **Location Table:** It is centrally maintained by Location Server. Location server uses it to manage location information of positioned samples and evaluating MN's current location.
  - The location table for each AP features
    - ✓ A 60 x 60 matrix for each AP with each entry in the matrix representing the location related to this AP.
    - ✓ It does not need to rebuild location table when position area is extended. It just adds AP and the associated location table.
    - ✓ The absolute location is represented by location table and AP

location.

➢ Data and information on each entry of Location Table:

Evaluation Base	AP Signal Strength	Similarity Threshold	Next 5 Locations	AP List

Table 1: Location Information Format in Location Table

- ✓ **Evaluation Base**: Besides initial samples, most of the samples in Location Table were based on other samples. Recording the previous sample for "Exception Location Detection" when Location Server is used for positioning MN's location. We will explain this in Section 3.4.4.
- ✓ **AP Signal Strength:** In our design, each AP has its own Location Table. This field records the Signal Strength of the selected AP at this location.
- ✓ **Similarity Threshold:** The average similarity of this entry (Location) compared with its 2meter-distance neighbors. When samples in Location Table were increased up to a high threshold, it will be very useful to speedup the position process. How to generate and how it works will be described in each prediction and confirmation process.
- ✓ **Next 5 Location:** It is used for "Potential Locations Prediction" on MN prediction.
- ✓ **AP List:** Each Location Table only records the location information for its AP. To perform positioning, AP Signal Strength Set is required. Referring to this field can get other AP's Signal Strength of its absolute location.

3.2 Architecture & Flow Diagram

Figure 8 shows the architecture and flow diagram of our proposed system. Generally, it can be divided into two phases. The first phase is performed in MN site and it includes wireless channels scanning, predicting location before sending position query and sending MN position query. The second phase is executed on Location Server site and its task includes confirming or evaluating MN's current location, responding MN position query and maintaining Location Table. We will use Figure 8 in the discussion in next two sections.

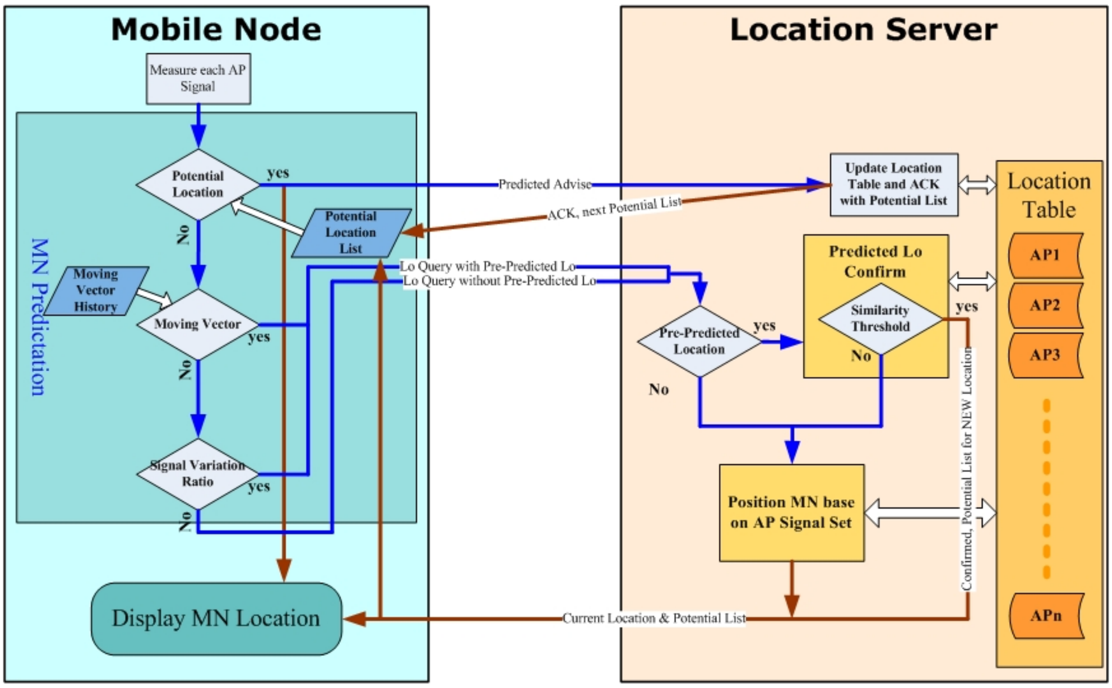

Figure 8: Proposed system architecture and flow diagram.

3.3 Tasks in MN Site

As shown in Figure 8, the first task on each positioning is that MN scans wireless channels and measures the Signal Strength of each detectable AP. After collecting the AP Signal Strength Set, positioning process can be started.

3.3.1 Predicting by Potential Locations

This is the first stage to predict MN's location. Considering the moving behavior, most of them may follow the same tracking. That is, if someone traveled to location A and went to location B, the next one who comes to A may follow this track and go to B. As discussed in Location Table data, each location records its 5 next locations. These data will be updated in recent positioning (Machine Learning). These 5 locations and their related information (AP Signal Strength Set and Similarity Thresholds) will be sent to MN at each succeed position. So, when an MN starts to perform positioning, it checks whether any potential location has been sent by last positioning first. If so, it compares MN's current measured AP Signal Strength Set with these locations' AP Signal Strength Set and selects the one with maximum similarity. If the maximum similarity is larger than the selected location's Similarity Threshold, we can consider this selected location as its current location. MN sends an informing message to Location Server to update the Location Tables by its measured AP Signal Strength Set. Location Server updates this location information on Location Tables and re-calculates the Similarity Threshold of this location and its neighbors with 2 meter distance. Then AP replies an ACK with next potential location list to MN for later usage. The Location Server needn't to position or confirm MN's location in this stage.

If this is not the previous position or there is no any potential location sent by Location Server, or predicting by potential locations was failed, it will go to next stage: Predicting by Uniform Linear Moving.

3.3.2 Prediction based on Uniform Linear Moving

Considering the moving objects, most of them are moving in uniform linear behavior. If this behavior is detected, it can be used for location prediction. As we discussed in Section 3.1, MN keeps four history locations, these locations can make 3 vectors. In our design, if these 3 moving vectors are similar, we can use it to predict MN's current location. Please refer to Figure 9.

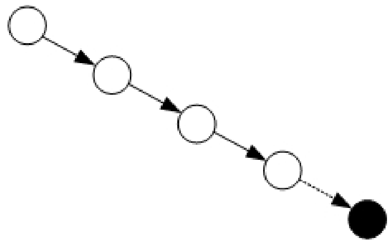

Figure 9: Prediction based on Moving Vector

When Uniform Linear Moving happened and was detected, MN sends a Location Query with measured AP Signal Strength Set and the predicted location to Location Server. Location Server confirms this predicted location by the location information (AP Signal Strength Set and Similarity Threshold) recorded in Location Table. If it succeeded, that is, the similarity of AP Signal Strength Set between "MN measured" and "location table recorded" is larger than the recorded Similarity Threshold, the Location Server updates this predicted location's information in Location Tables and re-calculated the Similarity Threshold of this location and its 2-meter neighbors. Then the Location Server replies an ACK with next potential location list to MN for next positioning.

If the confirmation was failed, Location Server performs **Evaluating by AP Signal Set of MN** to get MN's current location, and updates location table then replies an ACK with next potential location list to MN for next positioning, too.

If the condition in this stage does not exist, it will go to next stage: **Predicting by Signal Variation Ratio.**

3.3.3 Predict by Signal Variation Ratio

Before an MN sends a query to Location Server to perform positioning on Location Server, it can predict its own location by its signal variation from previous location. That's why we keep previous location information (Location and AP Signal Strength Set) in MN. That is, if an MN is not doing the first time positioning in this area, it can predict its location based on signal variation ratio first. This predicted location also can be added into position query to be sent to Location Server for confirmation. If confirmation is successful, complement positioning process is not needed on Location Server. This can reduce large amount of loading for Location Server especially when number of entries in the Location Table is large.

3.3.4 Location Query without MN Predicting

If all of above stages cannot be satisfied, the MN will send location query to Location Server with AP Signal Strength Set only. Location Server will evaluate the MN current location based on AP Signal Strength Set, and response to MN.

3.3.5 Algorithm on MN Site

Definition:

Pot_Lo: potential locations;

Pot_Lo_SminTh: Similarity Threshold of each Pot_Lo;

Pot_Lo_APSS: AP Signal Set of each Pot_Lo;

TRK_Lo: 4 MN last locations,

PAPSS: MN previous AP Signal Set;

Prd_Lo: Predicted Location;

Prd_Inf: MN predicted informing message. It asks Location Server to update Location Tables;

Lo_Query: MN Location Query message. Query for MN current location

Simi(APSS1,APSS2): Similarity between APSS1 (AP Signal Strength Set) and APSS2

```
Input: CAPSS: MN's current AP Signal Set;
Output: MN's location;

Begin:
    Scan all Wireless Channels to get AP Signal Set(CAPSS);
    if (Pot_Lo exists && Max[Simi(CAPSS, Pot_Lo_APSS)] ≧ Pot_Lo_SminTh)
    {
        send Prd_Inf(Pot_Lo(with Max Similarity),CAPSS);
        return (Pot_Lo(with Max Similarity));
    }

    else if (there are 4 TRK_Lo && all moving vectors are similar)
    {
        Prd_Lo (4th moving vector);
        send Lo_Query (Prd_Lo & CAPSS);
    }

    else if (PAPSS exists)
    {
        Prd_Lo(variation between PAPSS and CAPSS);
        send Lo_Query (Prd_Lo & CAPSS);
    }

    Else
    {
        send Lo_Query (CAPSS);
    }

    receive ACK or response from Location Server (positioned location, Pot_Lo)
    {
        renew Pot_Lo for next positioning;
        push positioned location into TRK_Lo queue;
        save CAPSS to PAPSS;
    }
    return (positioned location);
End
```

Figure 10: Algorithm on MN

3.4 Location Server Site Tasks

As shown in Figure 8, several different messages may be sent by MN. They are **Predicted Advise** (Sent by “MN’s Predicting by Potential Locations” stage), **Location Query with Predicted Location** (Sent by “MN Predicting by Uniform Linear Moving” or “MN Predicting by Signal Variation Ratio” stages), and **Location Query without Predicted Location** (Sent if MN has not performed prediction). Location Server must perform different tasks according different messages. We give a brief description below, and more details will be addressed in the following sections.

- **Predicted Advise:** If it can satisfy the “Predicting by Potential Locations” stage on MN site, the only tasks of Location Server are updating Location Tables and ACK with next potential location list. Location confirmation or positioning is not necessary.

- **Location Query with Predicted Location:** As we discussed in MN tasks, MN may predict its location through several methods. Most of the position queries may include its predicted location. Once the Location Server received a query with predicted location, it will decide whether to do confirmation by Similarity Threshold check first. If the similarity check is passed, it responses with confirmation and the next potential location list. If failed, it performs “Positioning by AP Signal Set of MN” task and feedbacks with positioned location and next potential location list.

- **Location Query without Predicted Location:** In the Position Query without predicted location, Location Server performs “Positioning by AP Signal Set of MN” task directly. It feedbacks with positioned location and next potential location list.

3.4.1 Similarity Threshold

Similarity Threshold has been addressed in our previous discussion. Before describing Location Server's tasks, we define and explain what it is and how it works first.

- **Purpose:** To check whether a predicted location is close to a sample location in Location Tables.

- **Definition:** If the similarity between the predicted location and the selected sample > the Similarity Threshold of the selected sample, the predicted location can be considered as equal to this selected sample.

- **Calculation:** Average Similarity between selected location and its 2-meter neighbors. For example, L_a is one of the locations in Location Tables. N_i is L_a's 2-meter neighbor. The Similarity Threshold of L_a is:

$$\frac{\text{Similarity}(La, N_1) + \text{Similarity}(La, N_2) + \ldots + \text{Similarity}(La, N_k)}{k} \quad \textbf{(2)}$$

- **Update Operation:** If any Location Tables updates for a successful positioning,

  1) If there are more than 2 neighbors at 2-meter distance, it calculates similarity for this updated location.

  2) Recalculate the Similarity Threshold for each 2-meter neighbor.

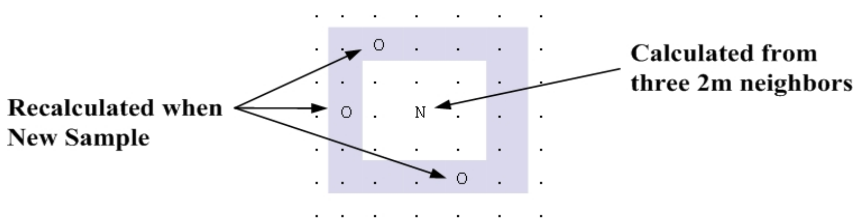

Figure 11: Similarity Threshold Update

3.4.2 Confirming MN's Predicted Location

As described above, a Location Server may receive MN's Position Query with predicted location. Location Server must confirm whether it is correct or not. Once Location Server receives such query, it will do the following:

1. Find out all neighbors within 2 meters of predicted location and make calculation to check whether the similarity with MN's AP Signal Strength is larger than their Similarity Threshold or not. If there are more than 50% successful, consider this confirmation is successful, then it performs the following process:
   - Location Server updates the location information onto the Location Tables with measured AP Signal Strength Set.
   - Recalculate Similarity Threshold of this location and its 2-meter neighbors.
   - Feedback confirmation to MN with potential list if any.
2. If the above process was failed, it is a failed confirmation and Location Server will perform **Evaluation by AP Signal Set of MN** to be described in the next section.

3.4.3 Evaluation by AP Signal Set of MN

This task was performed when predicted location confirmation failed or received Location Query without Predicted Location from MN. It can be divided into 2 steps: **Nearest Sample Discovery** and **Location Evaluation from Nearest Sample**.

Step 1. Nearest Sample Discovery

On Location Server, it tries to find out the nearest known node (Sample) in the Location Tables based on MN's measured AP Signal Strength Set. That is, to calculate each AP Signal Strength Set for the similarity between MN's measurement and each sample's record in Location Tables. The sample with maximum similarity will be the nearest sample.

$$\textbf{Nearest Sample} = \textbf{Max}\ [\textbf{Similarity}(\textbf{MN}, \textbf{Sample}_i)] \quad \textbf{(3)}$$

Step 2. Location Evaluation from Nearest Sample

Once the nearest sample was discovered on Location Table, the MN's location can be evaluated by the difference of each AP Signal Strength. Based on radio power Equation (4), we can obtain that in the same condition (same transmitter, receiver and other environment condition) the radio power is directly proportion to (1 / distance^2) as follows.

$$\frac{P_R}{P_T} = \left(\frac{\lambda}{4\pi d}\right)^2 * G_R G_T \quad \Rightarrow \quad P_R \propto \frac{1}{d^2} \quad \textbf{(4)}$$

And based on dBm (decibel relative to one milliwatt) Equation (5), we can predict the distance between each AP and MN by nearest sample distance to AP, nearest sample's AP signal strength and MN's AP signal strength as follows.

$$\textbf{dbm} = 10 * \log_{10}\left(\frac{P}{1mW}\right) \quad \textbf{(5)}$$

Replacing P by 1/d^2,

$$\Rightarrow \textbf{dbm} = 10 * \log_{10}\left(\frac{1/d^2}{1mW}\right) \quad \textbf{(6)}$$

We get the distance d:

$$\Rightarrow d = \sqrt{(1 / (10^{(dbm/10)} * 1mw))} \quad \textbf{(7)}$$

If we have the distance between the nearest sample (d_b), AP Signal Strength of nearest samples (dbm_b) and AP Signal Strength of MN (dbm_a), based on Equation (7), we can predict the distance between MN (a) and AP as:

$$=> d_a = d_b * \sqrt{\frac{10^{(dbm_b/10)}}{10^{(dbm_a/10)}}} \tag{8}$$

For an MN, it selects 3 located AP with strongest signal strength and calculates the distance to each of them by Equation (8). We can get MN's location based on trigonometry as shown in Figure 12. Where N_a denotes MN and N_b denotes Nearest Sample on Location Tables.

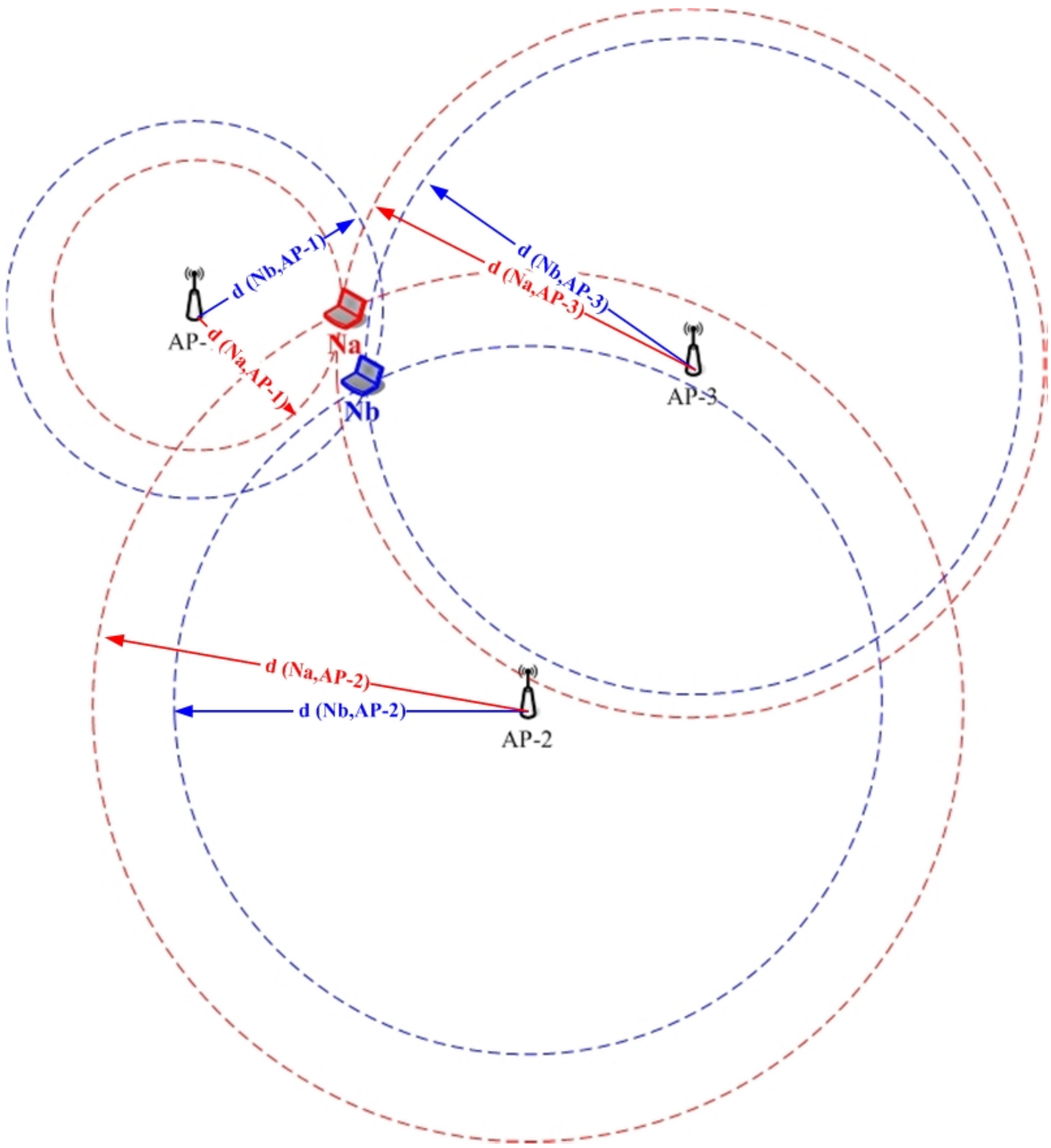

Figure 12: Evaluating MN's location (Na) by trigonometry.

3.4.4 Issues on Positioning and Location Tables Update

Since Signal Strength may have ±5% inaccuracy, the positioned location is not always precise. It is not proper to add any positioned into Location Tables as sample. Take an example as the worst case as shown in Figure 13, at 1st position reference from Initial Pattern, it may have 5% inaccuracy. At 2nd position reference from 1st positioned location, it may have even larger inaccuracy because of the 1st and 2nd 5% inaccuracy. And the inaccuracy may be become larger and larger because of the accumulation of inaccuracy in each previous positioning.

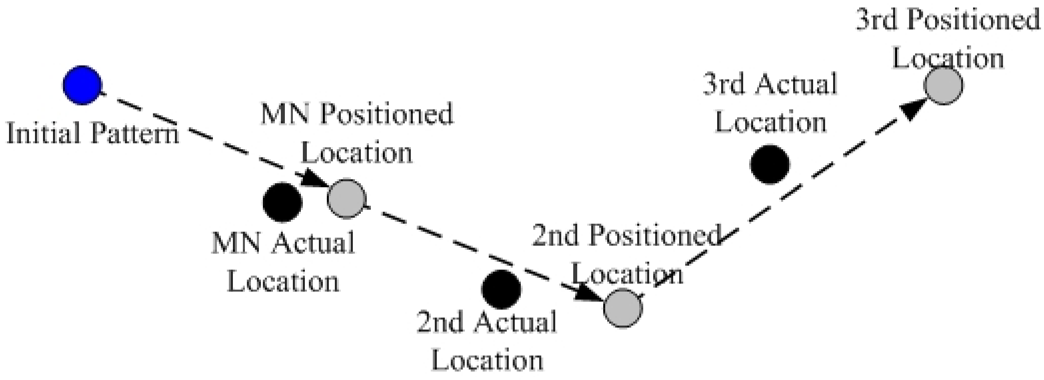

Figure 13: Worst case in positioning

To prevent such case from impacting the positioning accuracy, we setup 2 conditions and propose their solution in our system. "**Exceptional Location Detection"** and "**Trusted Location Checking"**.

Exceptional Location Detection

In each positioning, the system will check the positioned location to verify whether it is reasonable. Exceptional location means that it is unlikely an MN at this location to sense some AP, but it did and recorded in its measured AP Signal Strength Set. That is, after positioning, the system calculates the distance between positioned location and each of the sensed AP. If there is any one larger than a specific distance (we give 40m in indoor environment), it must be too far from actual location. When this situation happened, it means the nearest sample cannot be trusted. Therefore the system will remove it from location table and re-calculate it again. Furthermore, any positioned location on Location Tables which reference to this error sample also needs to be removed. This is why we give an "Evaluation Base" field in location entry of Location Tables. According to this method, number of error samples can be reduced in Location Table.

Trusted Location Checking

Due to unstable signal in indoor environment, sometimes positioned location may not be trusted for being used as sample recorded in Location Table. This is because the nearest sample may be far away from MN location, and it may happen at the beginning of the system operation with only few samples. We proposed a method to detect it. If the distance between the positioned location and its nearest sample is larger than a specific value (we give 5 meters in our system), it can be used as positioning result but not for the new sample. That is, only the positioned case whose distance between positioned location and nearest sample is less than 5 meter can be added into Location Tables as a new sample.

3.4.5 Algorithms on Server Site

Definition:

Prd_Adv: MN's predicted advise message;
Prd_Adv_ACK: ACK for Prd_Inf;
Lo_Query: MN's location query message;
Lo_Reply; Reply for Lo_Query;
SimiTh: Similarity Threshold;
Pot_Lo: Potential locations;
Pot_Lo_SminTh: Similarity Threshold of each Pot_Lo;
Pot_Lo_APSS: AP Signal Strength Set of each Pot_Lo;
APSS: AP Signal Strength Set;
CAPSS: MN's current AP Signal Strength Set;
Prd_Lo: Location predicted by MN;
Simi(APSS1,APSS2): Similarity between APSS1 and APSS2;
LoTable: Location Tables;
NB_2m: The neighbors with 2m distance;
NB_in2m: The neighbors within 2m distance;
num: The number of …
Evaluate_by_APSS( ): Subroutine to Evaluate MN Location by APSS;
Eval_Lo: Location evaluated by Evaluate_by_APSS( );

Input: MN message;
Output: Pos_Lo, Pot_Lo, Pot_Lo_SminTh, Pot_Lo_APSS;

```
Begin:
        receive message from MN;
        if (message == Prd_Adv)                    //MN Predicted Informing message arrived
        {
                update CAPSS to Location Table for Prd_Lo;
                re-calculate SimiTh(Prd_Lo, NB_2m);
                send Prd_Adv_ACK(Pot_Lo, Pot_Lo_SminTh, Pot_Lo_APSS);
                                 // ACK Pot_Lo, Pot_Lo_SminTh and Pot_Lo_APSS for Prd_Lo
        }

        else if (message == Lo_Query (Prd_Lo, CAPSS))
                                         //MN Location Query message with Prd_Lo arrived
        {
                if (num[Simi(CAPSS,APSS(NB_in2m))≧SimiTh(NB_in2m)] ≧
                        0.5*num[NB_in2m])    // Confirming Succeed by within 2m neighbors
                {
                        update LoTable(Prd_Lo) by CAPSS and SimiTh(Prd_Lo);
                        re-calculate and update SimiTh(NB_2m);
                        send Lo_Reply(Prd_Lo, Pot_Lo, Pot_Lo_SminTh, Pot_Lo_APSS);
                              // reply with Prd_Lo and its Pot_Lo, Pot_Lo_SminTh and Pot_Lo_APSS
                }
                else                                                // Confirming Failed
                {
                        Eval_Lo = Evaluate_by_APSS(CAPSS);
                        send Lo_Reply(Eval_Lo, Pot_Lo, Pot_Lo_SminTh, Pot_Lo_APSS);
                             // reply with Eval_Lo and its Pot_Lo, Pot_Lo_SminTh and Pot_Lo_APSS
                }
        }

        else if (message == Lo_Query (Prd_Lo, CAPSS))
                                      //MN Location Query message without Prd_Lo arrived
        {
                        Eval_Lo = Evaluate_by_APSS(CAPSS);
                        send Lo_Reply(Eval_Lo, Pot_Lo, Pot_Lo_SminTh, Pot_Lo_APSS);
                             // reply with Eval_Lo and its Pot_Lo, Pot_Lo_SminTh and Pot_Lo_APSS
        }
End
```

Figure 14: Algorithm on Location Server

```
// Evaluate_by_APSS( )
Definition:
        APSS: AP Signal Strength;
        LoTable: Location Table;
        EvalBase: The previous sample which this sample was evaluated from;
        SimiTh: Similarity Threshold;
        Pot_Lo: Potential Locations of the sample;
        Sample: Samples recorded in LoTable. With data of EvalFrom, APSS,
                SimiTh, and Pot_Lo;
        NSmp: Nearest sample in LoTable;
        Simi(APSS1,APSS2): Similarity between APSS1 and APSS2;
        Dist(a, b): the distance from a node to b;
        dbm(N): the selected AP signal strength at N node;
        AP_Lo: AP Location;
        Trig((AP_Lo & DistAP)*3): Subroutine to evaluate MN location by
                                  Trigonometric;
        Eval_Lo: The location evaluated by Location Server

Input: CAPSS;
Output: Eval_Lo;
Begin:
      NSmp = Max[Simi(CAPSS, all sample APSS)];        //Discover the nearest sample
      for (1st ~ 3rd strongest AP signal duplicated in NSmp(APSS))
      {
            Dist(APi, MN) = Dist(APi, NSmp) * √(10^(dbm(NSmp)/10) / 10^(dbm(MN)/10))
                                                   //Calculate DistAPi(MN) by Equation (8)
            i ++;
            next AP;
      }

      Eval_Lo = Trig (AP1_Lo, AP2_Lo, AP3_Lo, Dist(AP1), Dist(AP2), Dist(AP3));
                                                   //Evaluate MN Location by Trigonometric
      if (any Dist(APi, Eval_Lo) > 40m)             //Evaluated location is impossible
      {
            remove NSmp from LoTable;              //NSmp cannot be trusted as sample
            remove Samples(EvalBase =NSmp) from LoTable;
```

```
                                    //Those samples evaluated from NSmp also cannot be trusted
            perform Evaluate_by_APSS( ) again;
        }
        else
        {
            if (Dist(Eval_Lo, NSmp) ≦ 5)        //Checking if can be trusted for sample
            {
                update LoTable(Eval_Lo) by CAPSS and SimiTh(Eval_Lo);
                re-calculate and update SimiTh(NB_2m);
            }
            return (Eval_Lo)
        }
End
```

Figure 15: Evaluation_by_APSS( ): Subroutine for Evaluating Location by APSS

Chapter 4 Experiment and Numerical Results

4.1 Experiment Setup – Simulation Model

To evaluate our proposed system, we built a simulator to verify its accuracy and capability. In our simulation, we define an area as the positioned area first. Then we setup several APs in this area. Finally, we give some initial samples. At each position simulation, it will do the following:

1. Randomly select a location in the defined area (Actual Location).
2. At this location, calculate the distance to each AP, simulate the signal strength of each AP through Equation (9), and add ±5% error rate to the result. If the signal strength is less than -94 dBm, consider the AP as out of awareness at this location, as we can see in captured packets (Signal dbm= -35 ~ -94).

$$\textbf{Signal dBm} = \mathbf{-35 + 20 * \log_{10}(1/d^2)} \qquad \textbf{(9)}$$

3. Insert the AP Signal Strength Set (simulated at Step.2) into the proposed system, and evaluate the location.
4. Calculate the distance between Actual Location and Evaluated Location as Error Distance. The smaller the Error Distance, the higher accuracy the positioning.

There are 4 simulation models in our experiment to estimate our system. A 20x20 m^2 small area, a 50x50 m^2 big area, a 20x100 m^2 rectangular area which is separated into 4 sub areas, and an extended positioning area case. For simulation model 1~3, we add two cases as the contrast cases.

- **Sample Every Location:** Before system initialization, every location already been measured to obtain the AP Signal Strength. The location with maximum similarity will be the positioned result.

- **Without Machine Learning:** This is for positioning only. The evaluated locations will not be added into Location Table as samples.

4.1.1 20x20 m^2 with three Access Points

This is the basic positioning experiment, and the testing cases include:

- **Sample Every Location Case**
- **Without Machine Learning Case**
- **One Initial Sample:** Only one initial sample before the system starts.
- **One Initial Sample for each AP:** To measure one initial sample near each AP before the system starts.

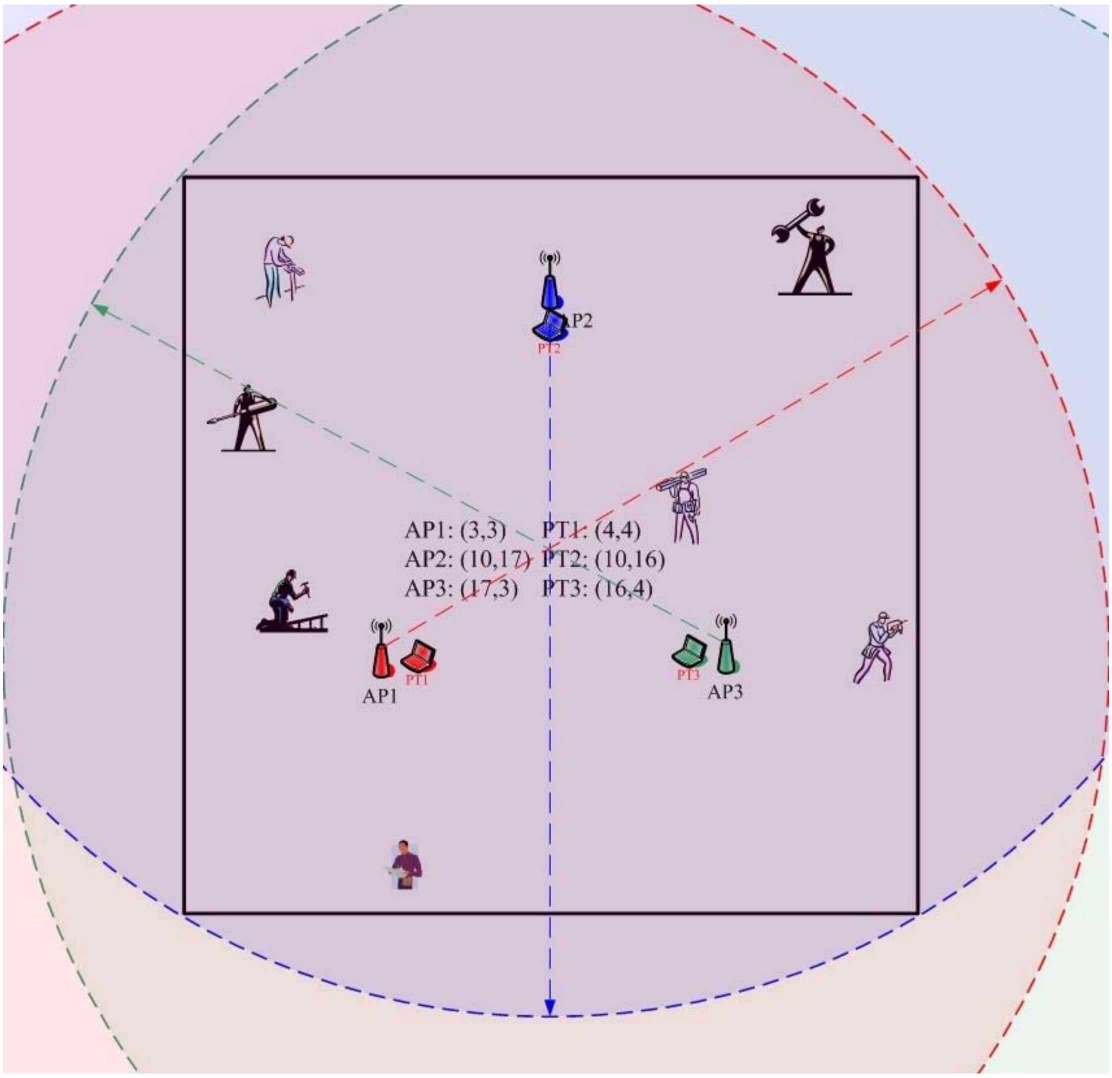

Figure 16: Environment for simulation model in 20x20 m^2 area.

4.1.2 50x50 m^2 with 13 Access Points

This is basic positioning experiment, too. The testing cases include:

- **Sample Every Location Case**
- **Without Machine Learning Case**
- **4 Initial Sample:** Only 4 initial samples before system starts.
- **One Initial Sample for each AP:** To measure one initial sample near each AP before system starts.
- **33 Initial Sample:** To speedup the training time, we add another 20 initial samples randomly in addition to that for each AP before the system starts.
- **50 Initial Sample:** To speedup training time much quickly and get higher accuracy earlier, we set initial samples to 50 before the system starts.

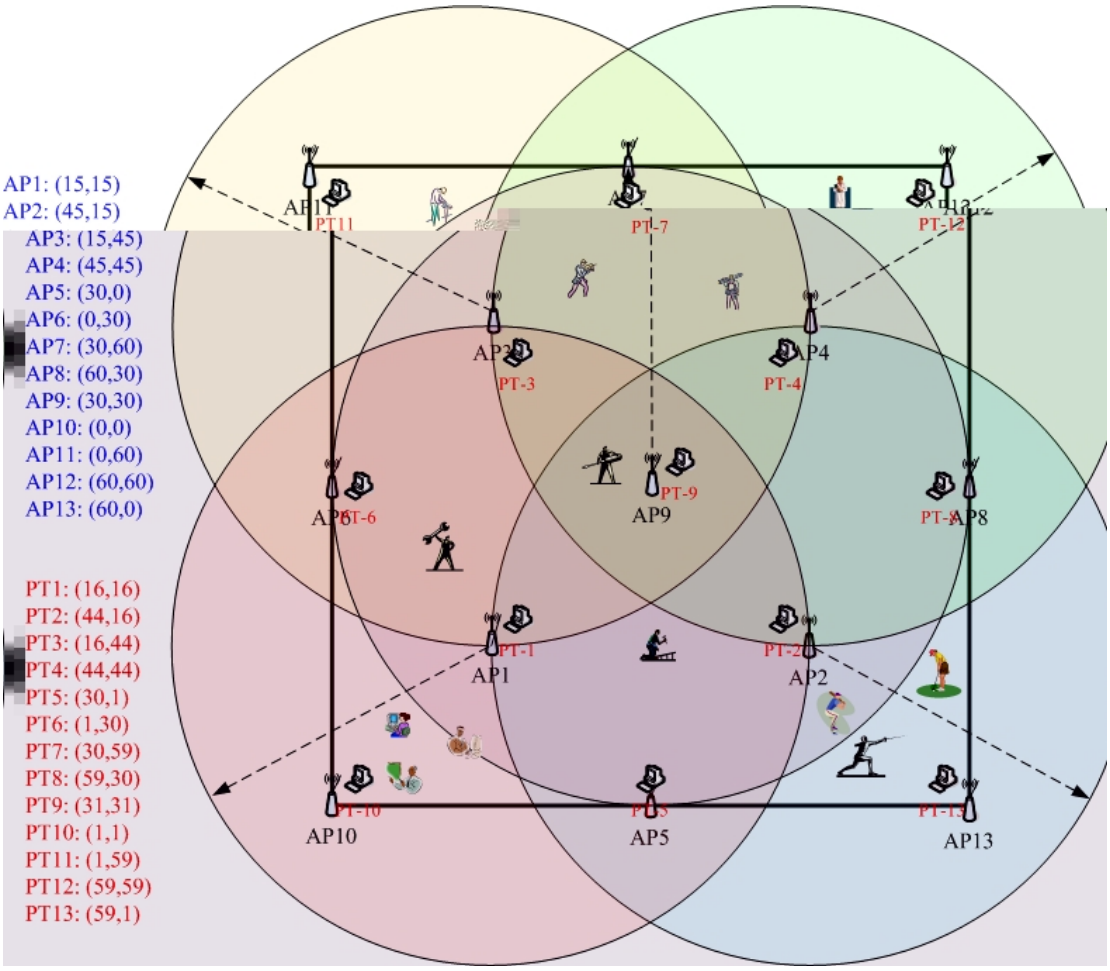

Figure 17: Environment for simulation model in 50x50 m^2 area.

4.1.3 100x20 m^2 with 12 Access Points

It is another basic positioning experiment. The testing cases include:

- **Sample Every Location Case**
- **Without Machine Learning Case**
- **4 Initial Sample:** Only 4 initial samples before the system starts.
- **One Initial Sample for each AP:** To measure one initial sample near each AP before the system starts.
- **32 Initial Samples:** To speedup training time, we add another 20 initial samples randomly in addition to that each AP before the system starts.
- **50 Initial Samples:** To speedup training time much quickly and get higher accuracy earlier. We set initial samples to 50 before the system starts.

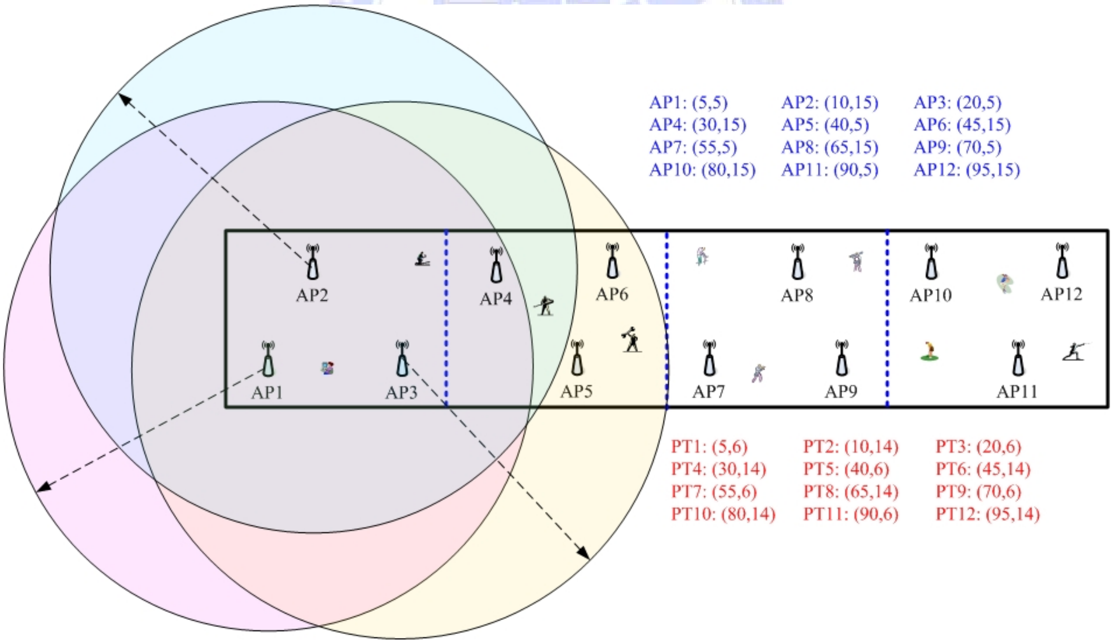

Figure 18: Environment for simulation model in 100x20 m^2 area.

4.1.4 20x20 m^2 Extended Case:

Only one case in this simulation model, the scenario is:

1. Use 1000 positioning for training the existing position area (Original Area).
2. Add another 3 AP to extend the position area (Extended Area).
3. In Extended Area, run the positioning test.

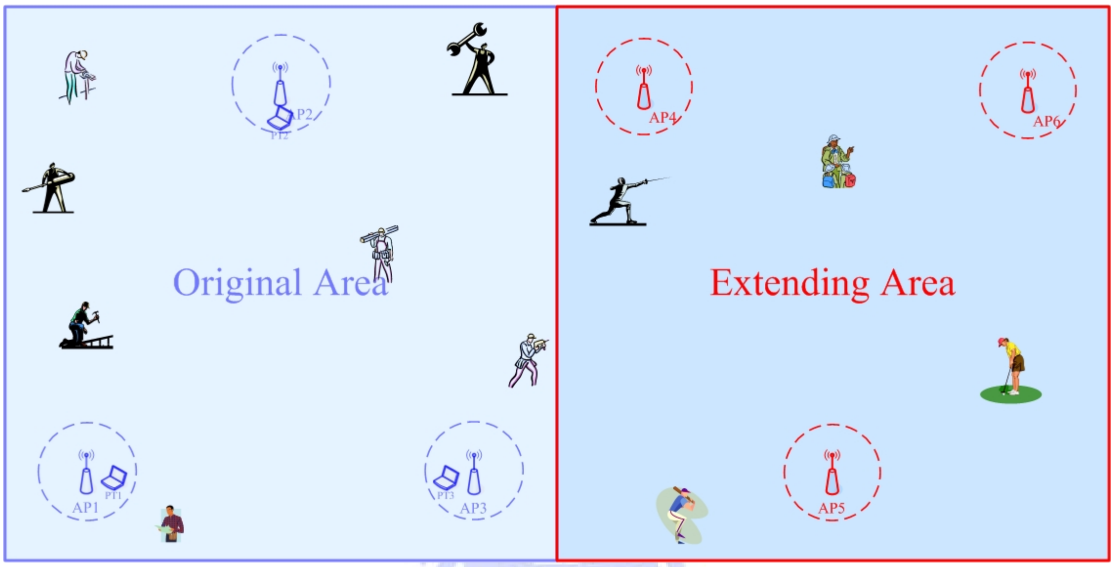

Figure 19: Environment for extended simulation model in20x20 m^2 area.

4.2 Experiment Result

In this section, we present our experiment result. We show it by separating the "Basic Positioning Scenario" and "Extended Area Scenario".

4.2.1 Basic Positioning Scenario

Before showing the result of our proposed system, we present the contrast models in our simulator first. The first one is "**Sample Every Location**" case which was shown in Figure 20~22. As we can observe, the average distance error for 20x20 m^2 area is 1.44 meter with 1000 positions. It is 1.54 meter for 50x50 m^2 area and 1.77 meter for 100x20 m^2 area. Although this method has highest accuracy in our simulator, but it is a time consuming work to measure samples at every location. System construction and maintenance will be a large burden. As a consequence, this method doesn't have extending ability.

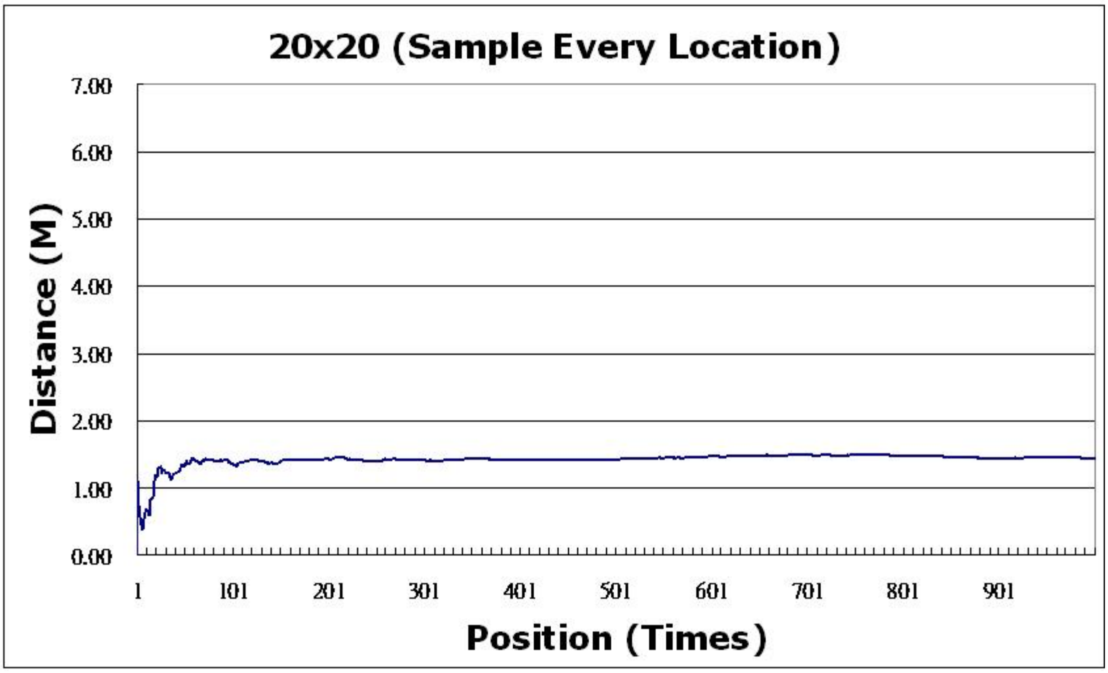

Figure 20: Average distance error on "Sample Every Location" Case in 20x20 m^2 area.

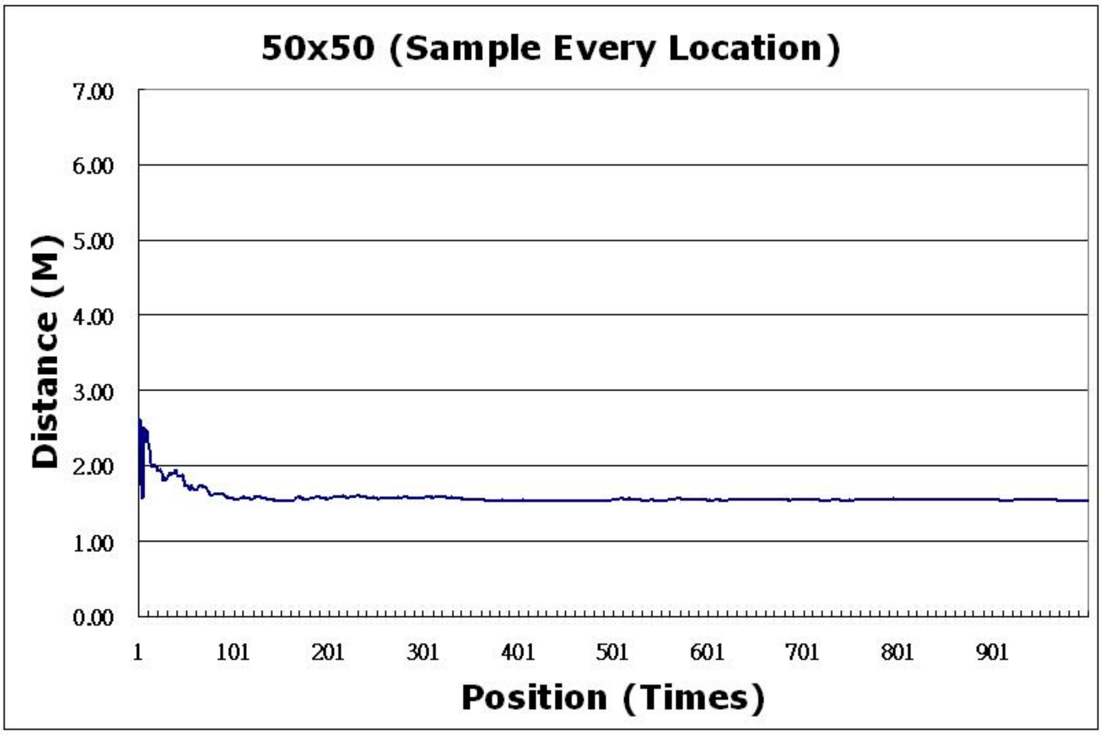

Figure 21: Average distance error on "Sample Every Location" Case in 50x50 m^2 area.

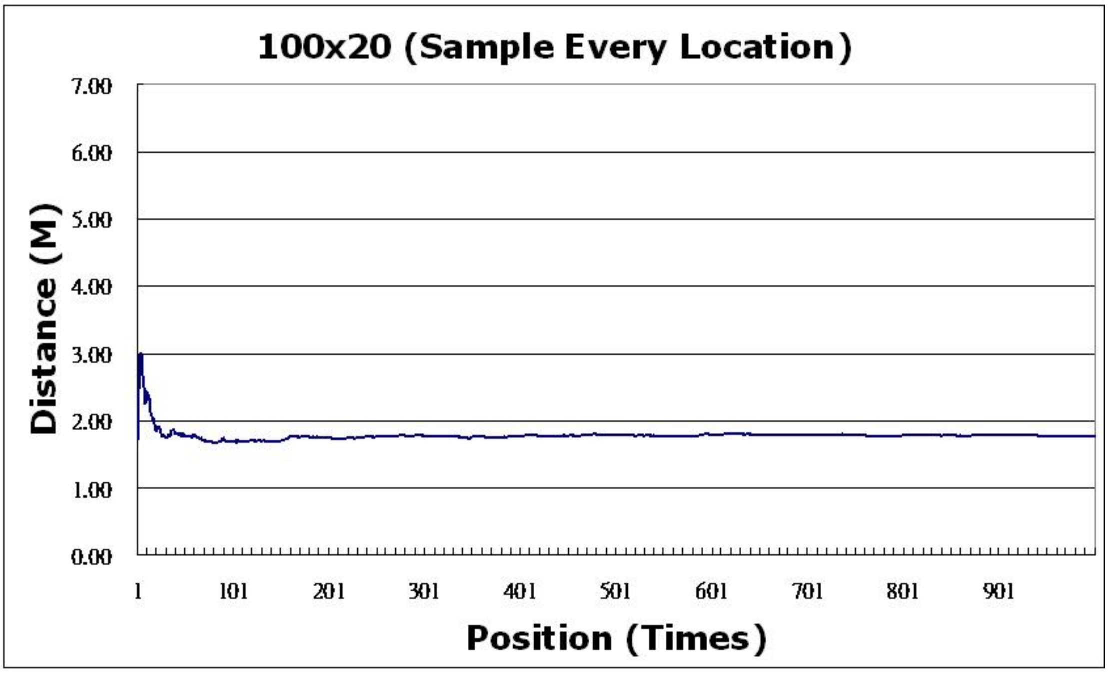

Figure 22: Average distance error on "Sample Every Location" Case in 100x20 m^2 area.

The second contrast model is "**Without Machine Learning**". Figure 23~25 show the average distance error for each simulation model is 5.10 m, 4.79 m and 4.70

m, respectively. Besides low accuracy, it doesn't have extending ability.

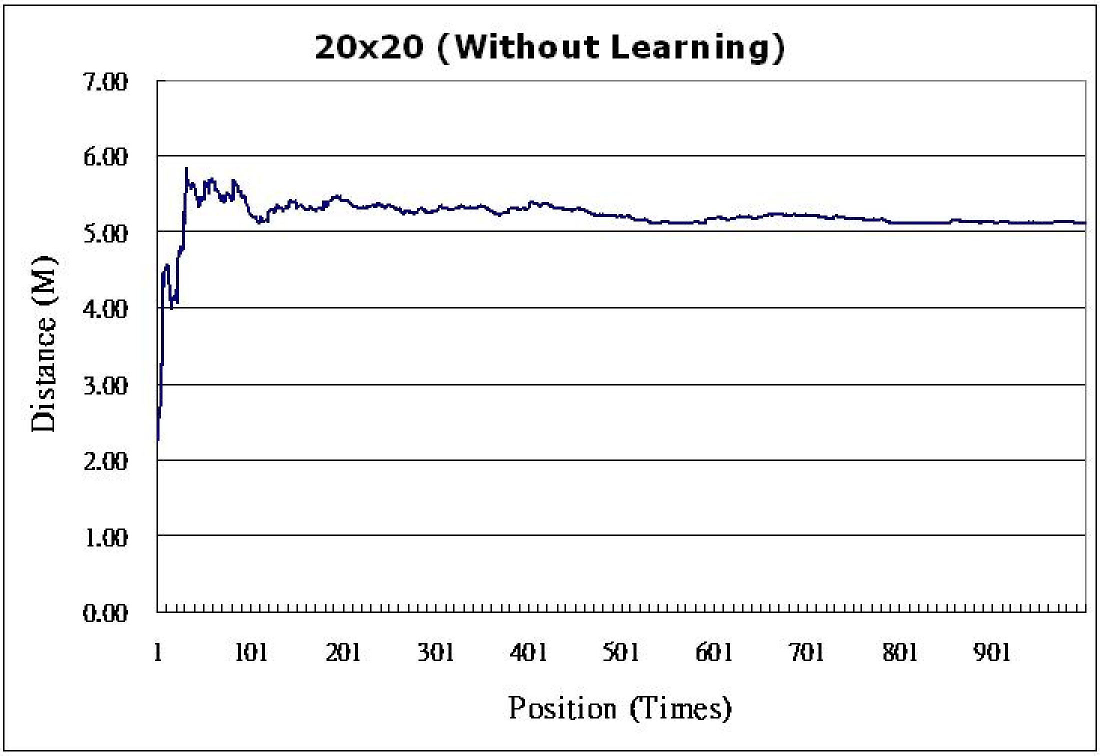

Figure 23: Average distance error on "Without Learning" Case in 20x20 m^2 area.

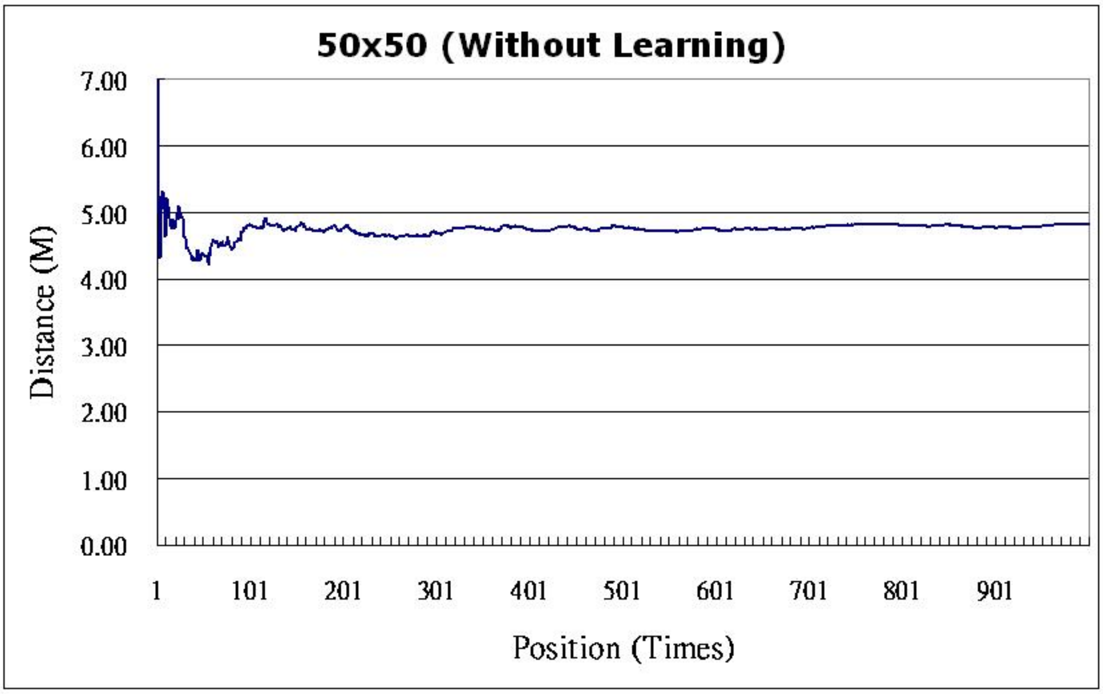

Figure 24: Average distance error on "Without Learning" Case in 50x50 m^2 area.

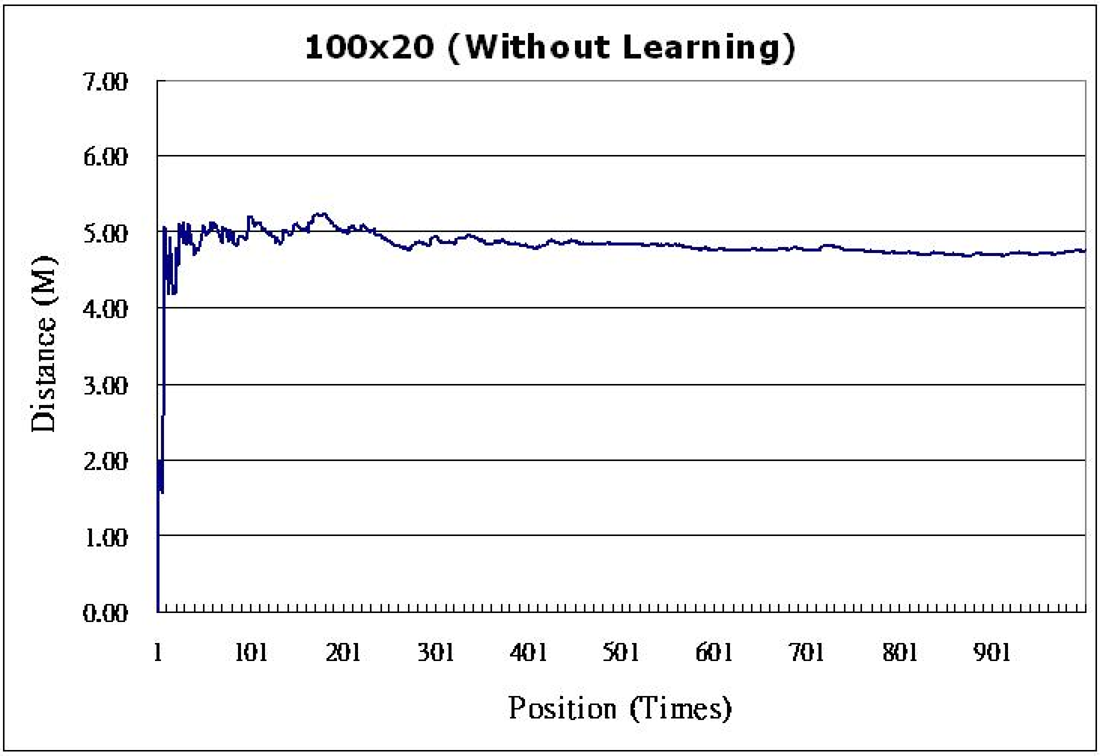

Figure 25: Average distance error on "Without Learning" Case in 100x20 m^2 area.

Now, let us show the result of our system simulation. Figure 26~28 show experiment results with few Initial Samples for 20x20 m^2, 50x50 m^2 and 100x20 m^2 respectively. Figure 26 indicates that the average distance error is down to less than 3 meters after 900 positioning. In Fig. 27, for 50x50 m^2 area case, it also has the same result after 6400 positioning. Figure 28 shows100x20 m^2 area case, it reaches the same result after 10000 positioning. It proves that our system can get significant result after certain period of training.

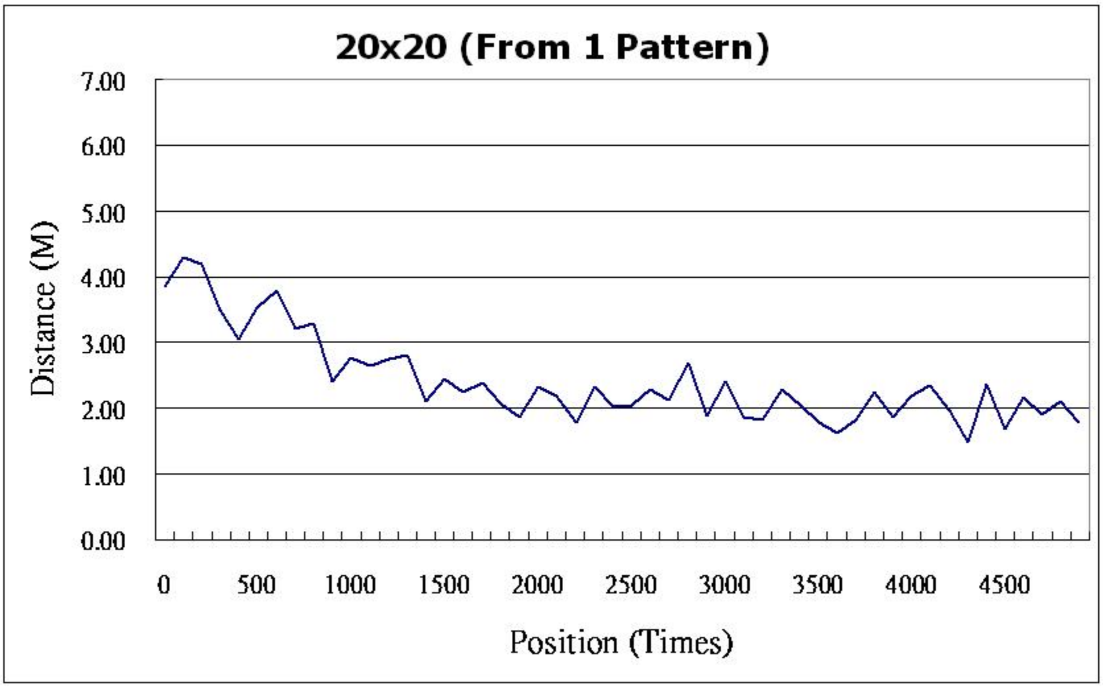

Figure 26: Average distance error on the case with few Initial Samples in 20x20 m^2 area (1).

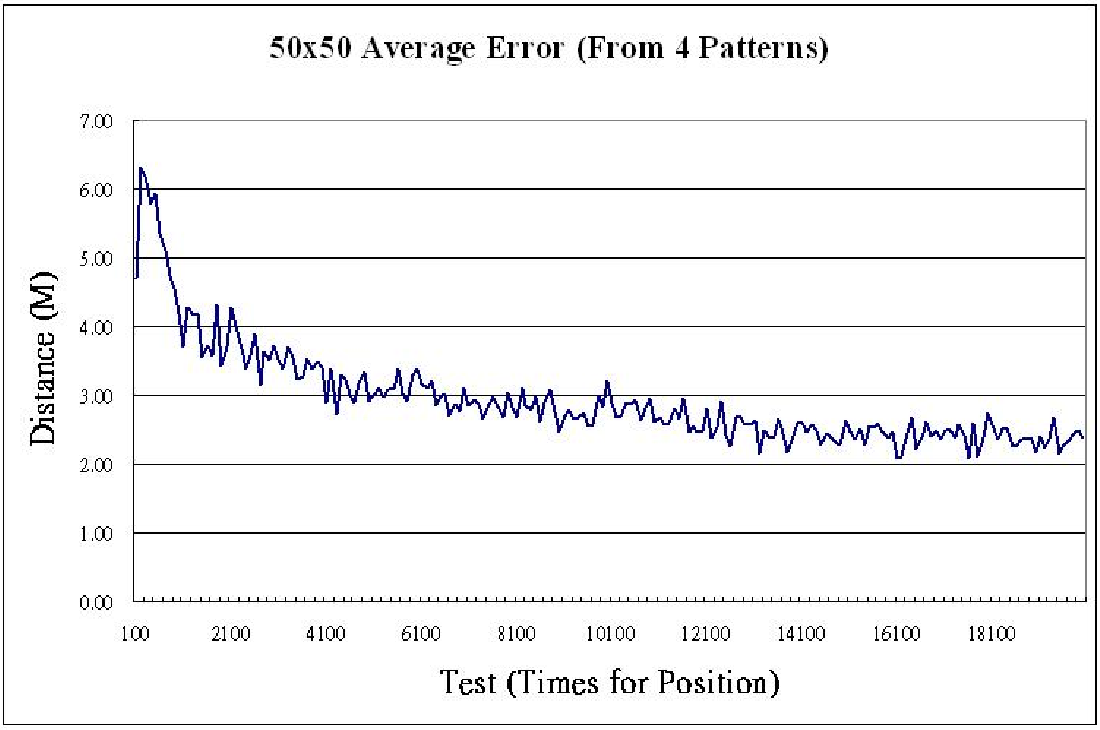

Figure 27: Average Distance Error on the case with few Initial Samples in 50x50 m^2 area(4).

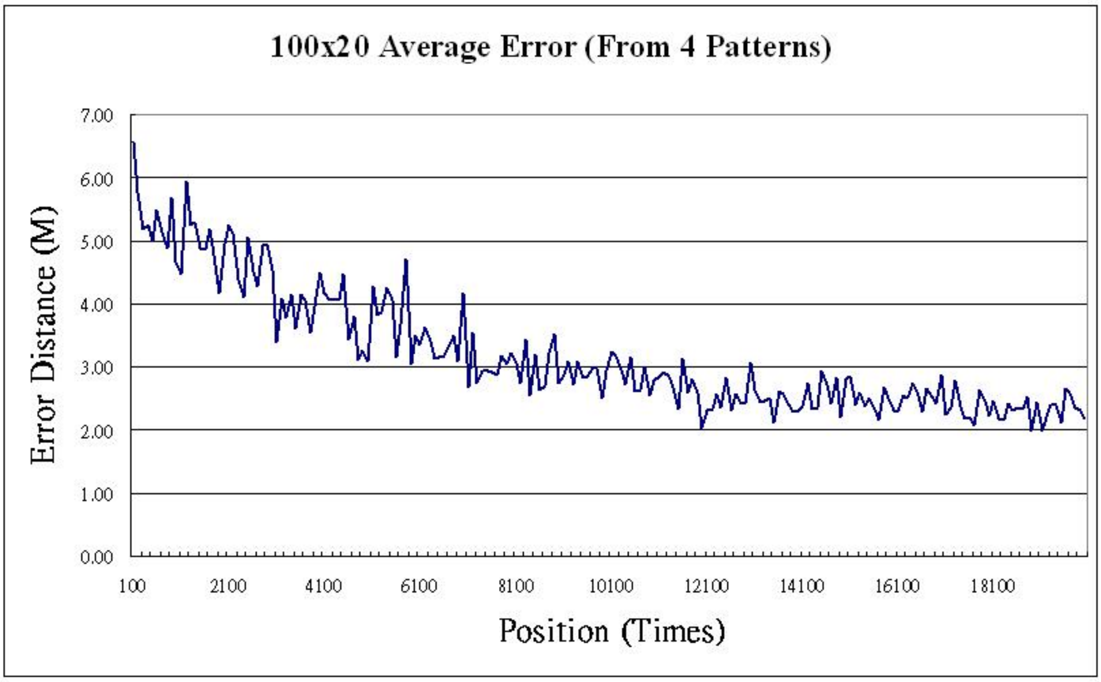

Figure 28: Average distance error on the case with few Initial Samples in 100x20 m^2 area(4).

Somebody may have judgment for the above result that it spends too long time to train the system for significant positioning. This issue can be improved by adding few Initial Samples manually. In the following figures, we show the result of each model with different Initial Samples.

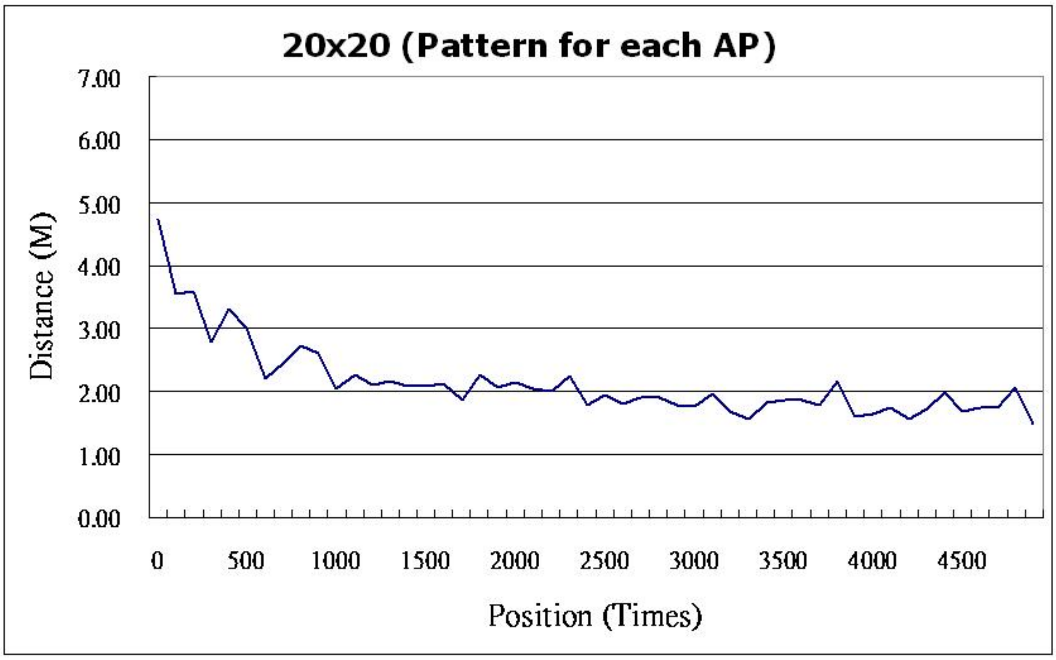

Figure 29: Average distance error on 20x20 m^2 area (One Initial Samples per AP).

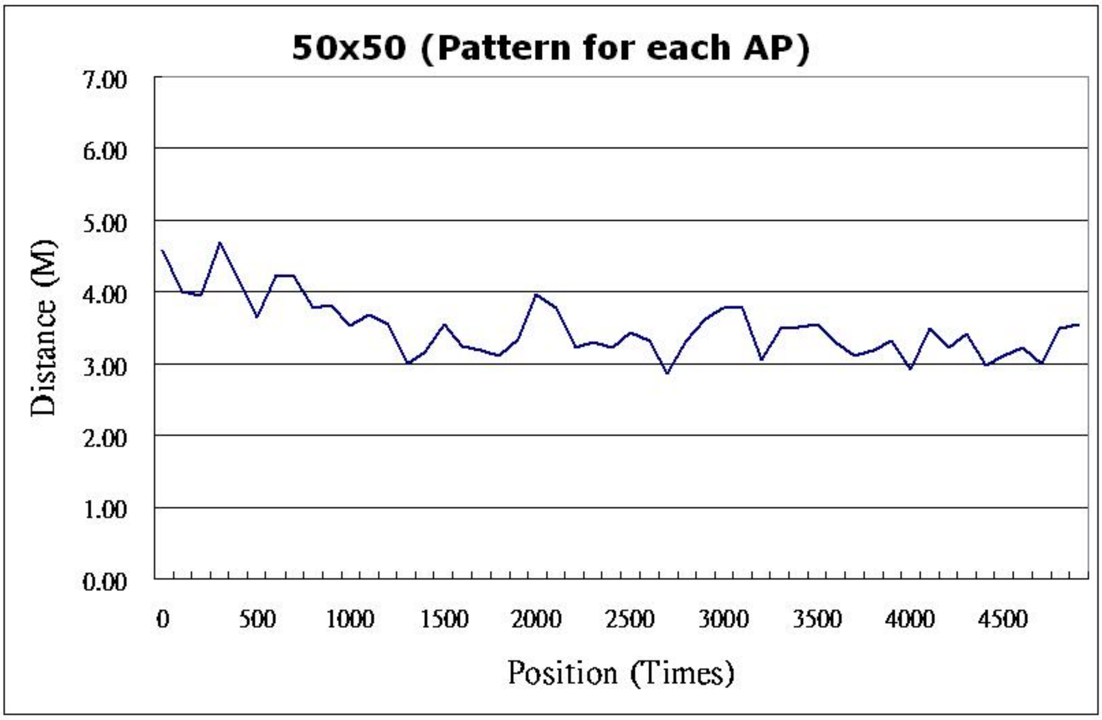

Figure 30: Average distance error on 50x 50 m^2 area (One Initial Samples per AP).

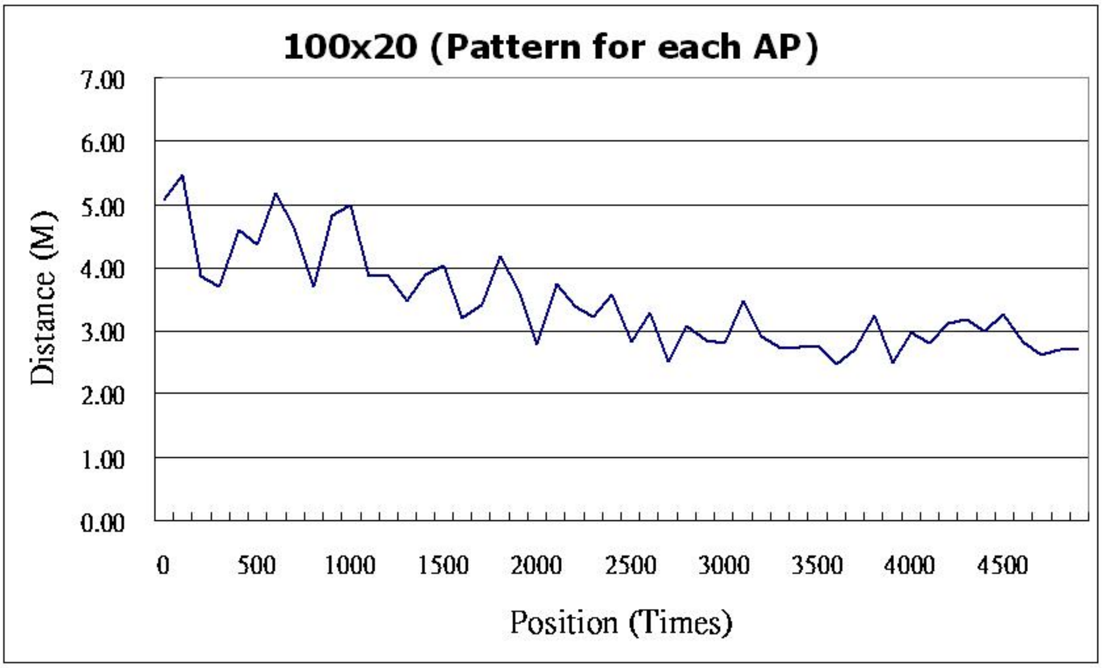

Figure 31: Average distance error on 100x20 m^2 area (One Initial Samples per AP).

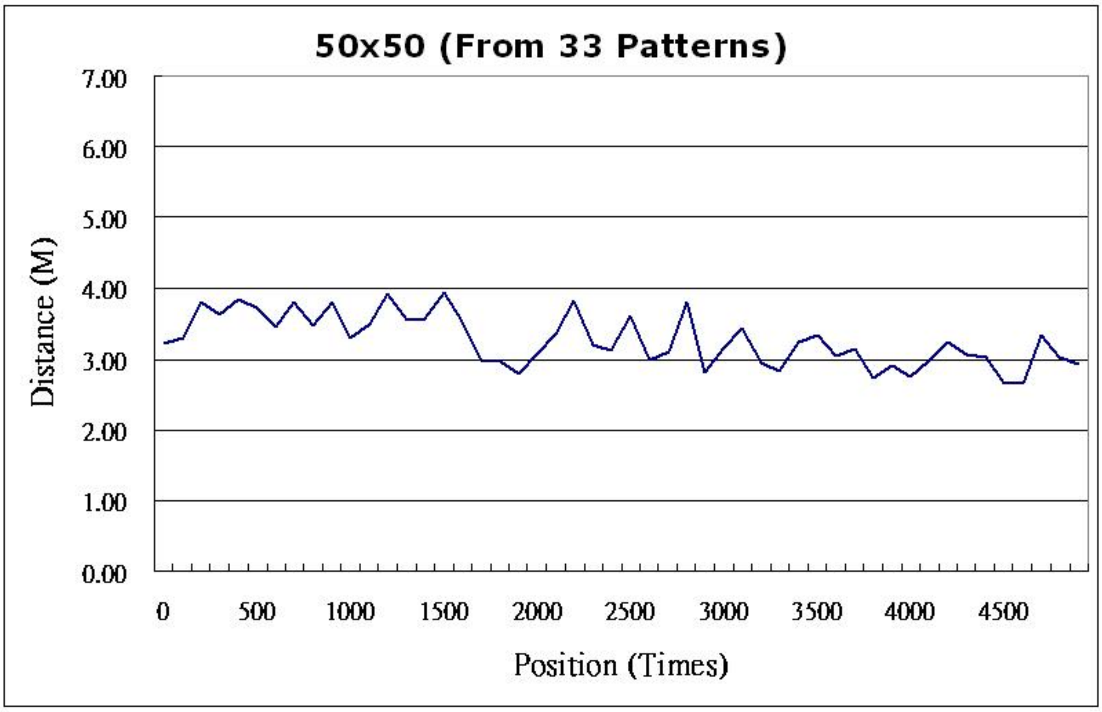

Figure 32: Average distance error on 50x50 m^2 area (33 Initial Samples).

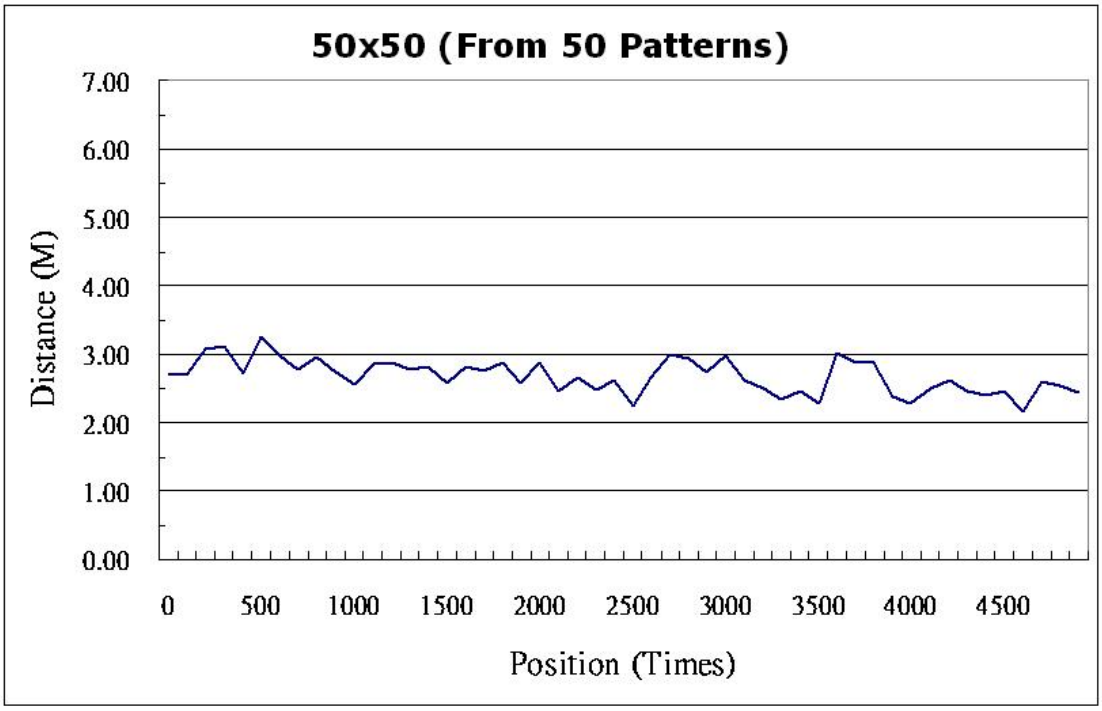

Figure 33: Average distance error on 50x50 m^2 area (50 Initial Samples).

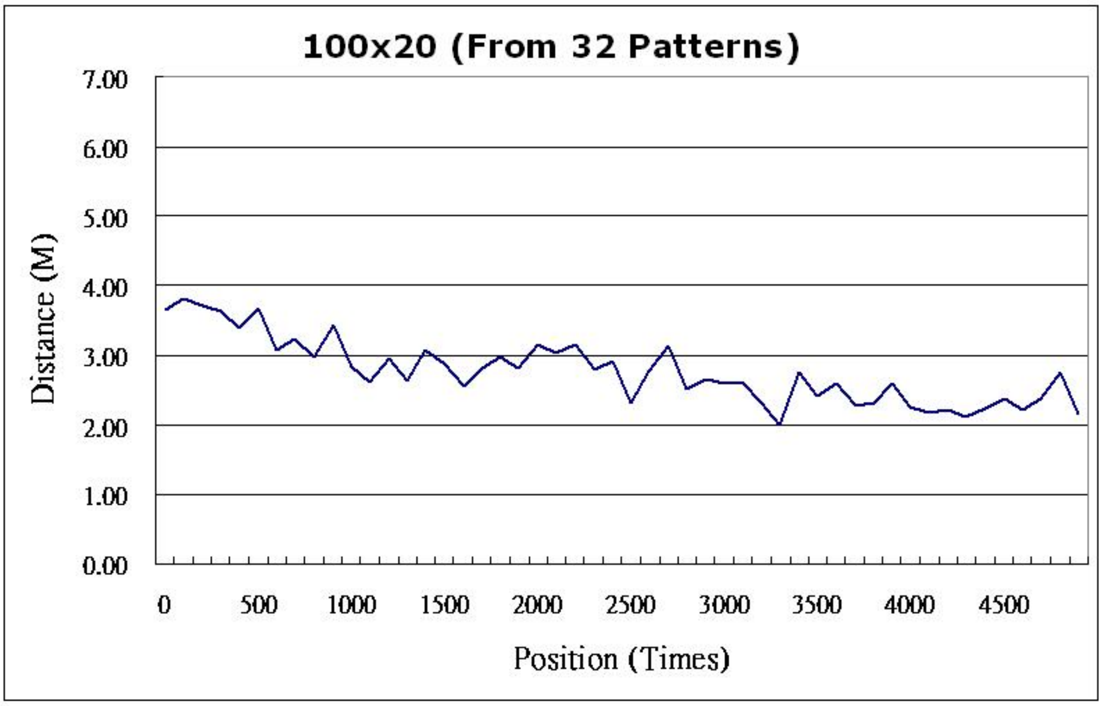

Figure 34: Average distance error on 100x20 m^2 area (32 Initial Samples).

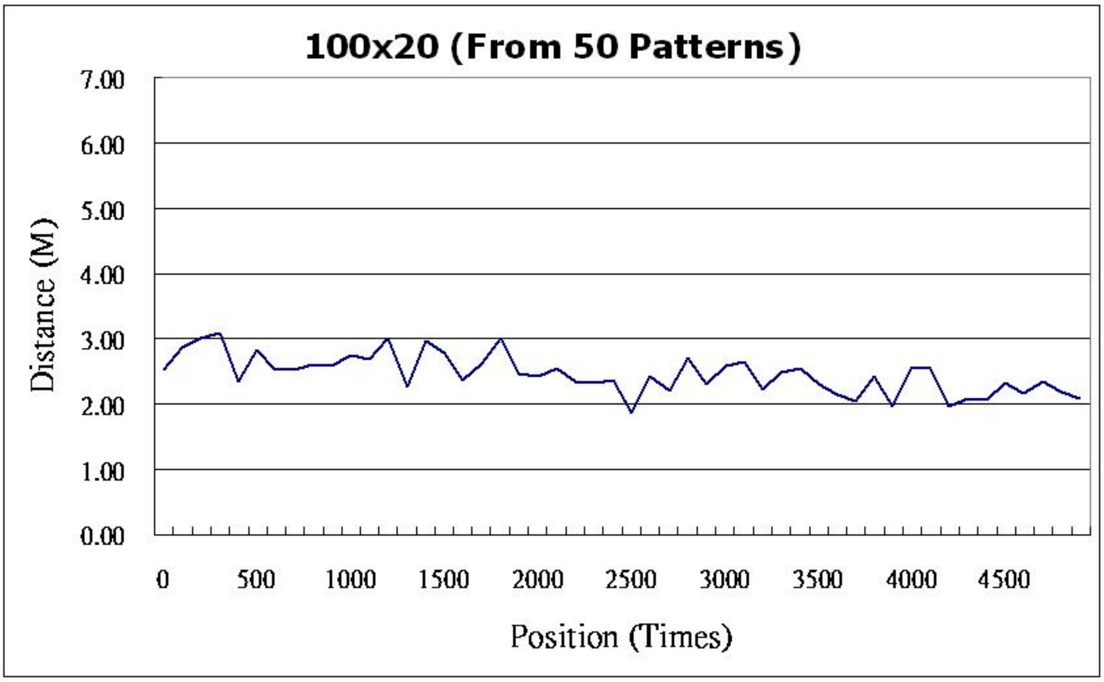

Figure 35: Average distance error on 100x20 m^2 area (50 Initial Samples).

Finally, we compare all test cases in precision for each model. For 20x20 m^2 model, the comparison is shown in Figure 36.

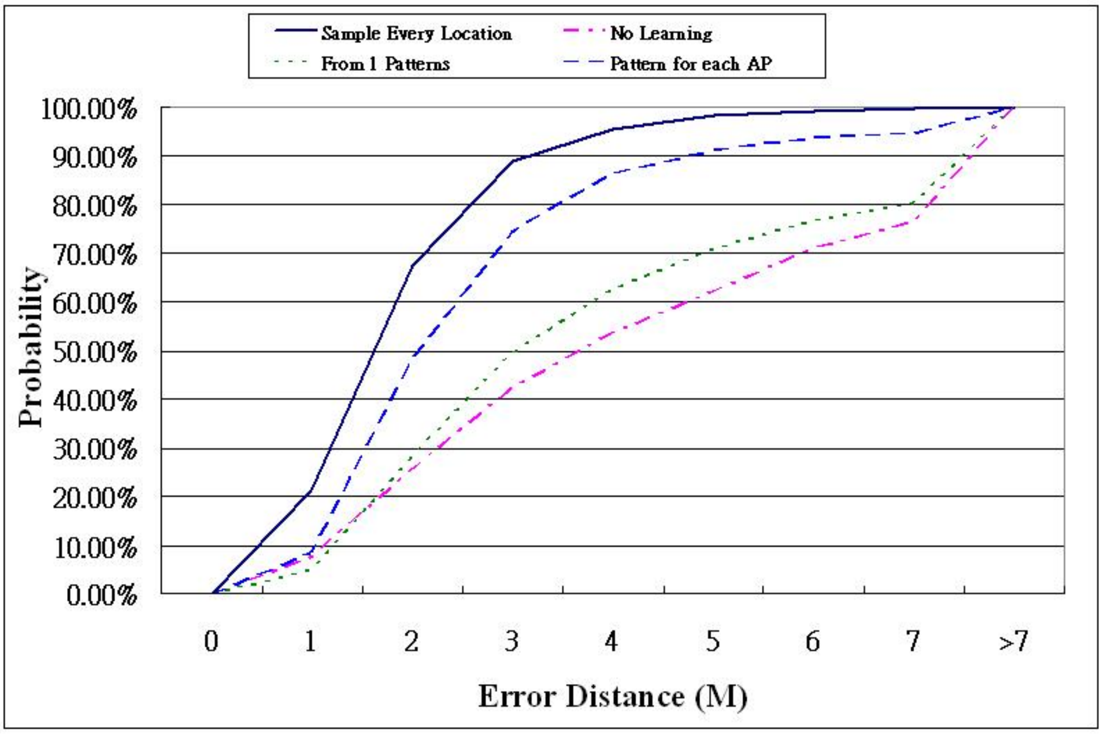

Figure 36: Comparison for the case on 20x20 m^2 area (all positions).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	21	67	89	95	98	99	99.5
No Learning	7	26	42	54	62	71	76
1 Initial Sample	**5**	**28**	**49**	**63**	**71**	**77**	**80**
Sample Per AP	**9**	**48**	**74**	**86**	**91**	**93**	**94**

Table 2: Precision (%) table in 20x20 m^2 area (all positions).

Above are the figure and table of comparison data for all positions on 20x20 m^2 model. It includes the data during system beginning.

In the following figure and table, we show the data which was run after 4000 position trainings. As we can see in Figure 37, for “Pattern for each AP, it is very close to Sample Every Location. According to Table.3, 82% of positions are within 3 meter positioning accuracy.

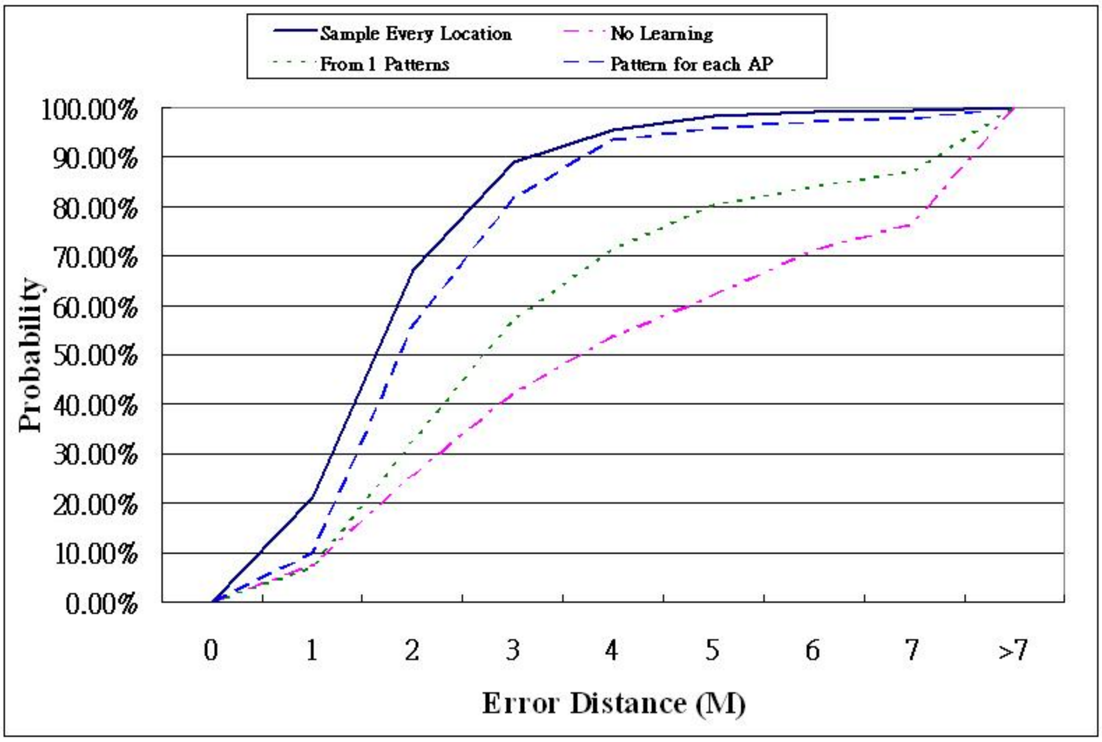

Figure 37: Comparison on 20x20 m^2 area (after training).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	21	67	89	95	98	99	99.5
No Learning	7	26	42	54	62	71	76
1 Initial Sample	**7**	**32**	**57**	**71**	**80**	**84**	**87**
Sample Per AP	**10**	**56**	**82**	**93**	**95**	**97**	**98**

Table 3: Precision (%) table on 20x20 m^2 area (after training).

The following figures and tables are comparison data for 50x50 m^2 area, with display for “all positions” and “after training”, respectively.

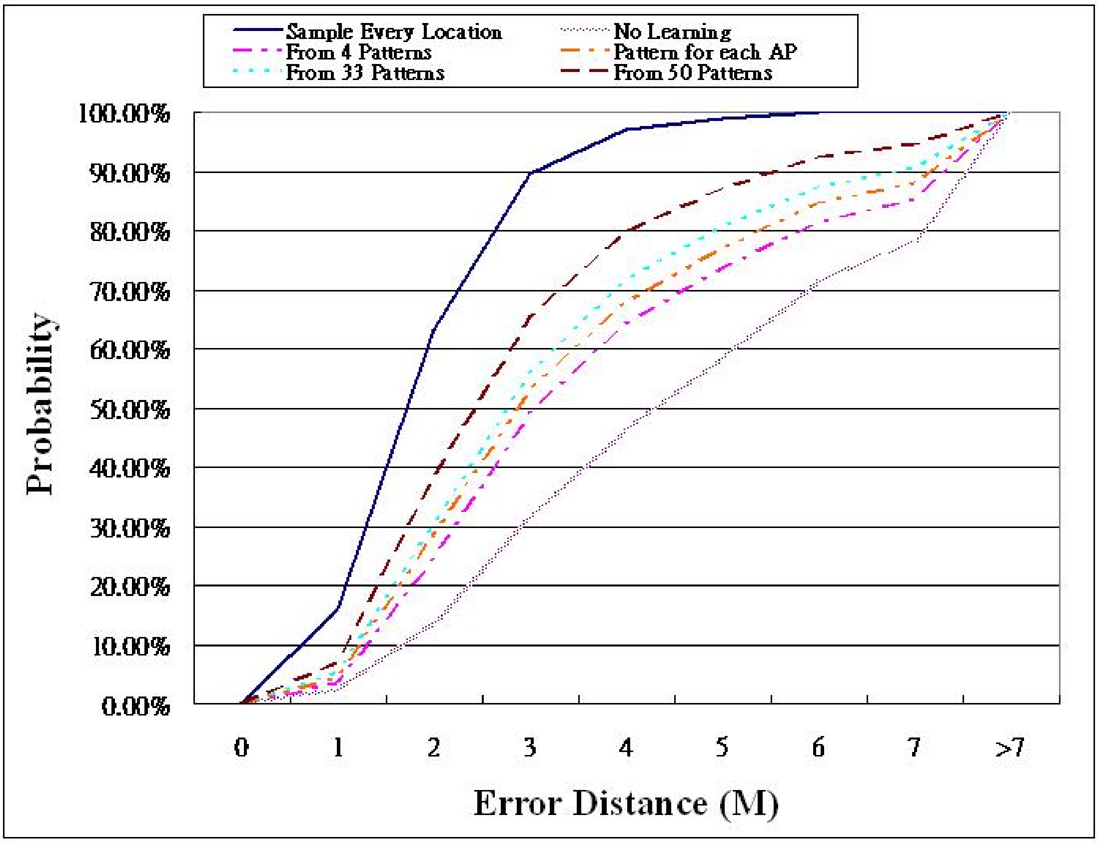

Figure 38: Comparison on 50x50 m^2 area (all positions).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	16	63	90	97	98	99	100
No Learning	2	14	31	46	58	72	78
4 Initial Samples	**4**	**25**	**49**	**64**	**74**	**81**	**85**
Sample Per AP	**4**	**29**	**53**	**68**	**77**	**85**	**88**
33 Initial Samples	**5**	**30**	**56**	**72**	**81**	**87**	**90**
50 Initial Samples	**7**	**39**	**65**	**80**	**87**	**92**	**94**

Table 4: Precision (%) table on 50x50 m^2 area (all positions).

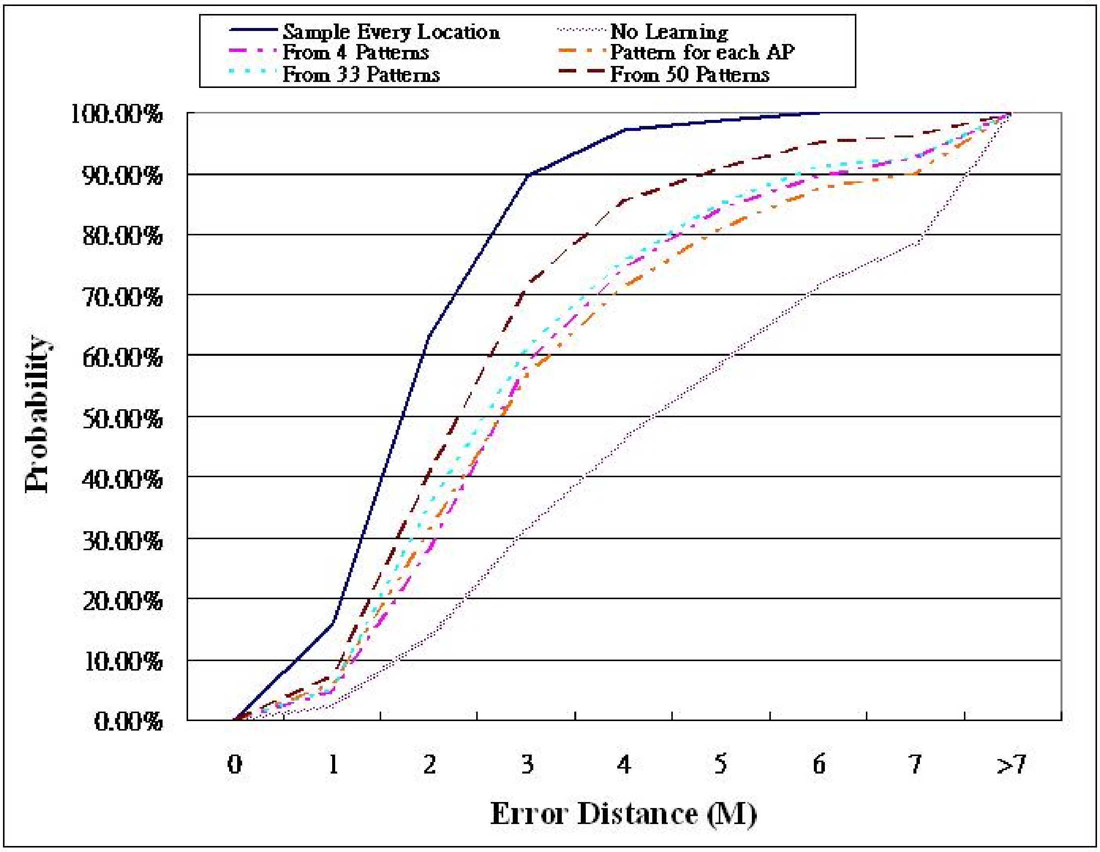

Figure 39: Comparison on 50x50 m^2 area (after training).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	16	63	90	97	98	99	100
No Learning	2	14	31	46	58	72	78
4 Initial Samples	**5**	**28**	**59**	**64**	**74**	**81**	**85**
Sample Per AP	**6**	**31**	**57**	**71**	**81**	**88**	**90**
33 Initial Samples	**5**	**35**	**61**	**76**	**85**	**91**	**93**
50 Initial Samples	**7**	**41**	**72**	**85**	**91**	**95**	**96**

Table 5: Precision (%) table on 50x50 m^2 area (after training).

The following figures and tables are comparison data for 100x20 m^2 area, with display for “all positions” and “after training” respectively.

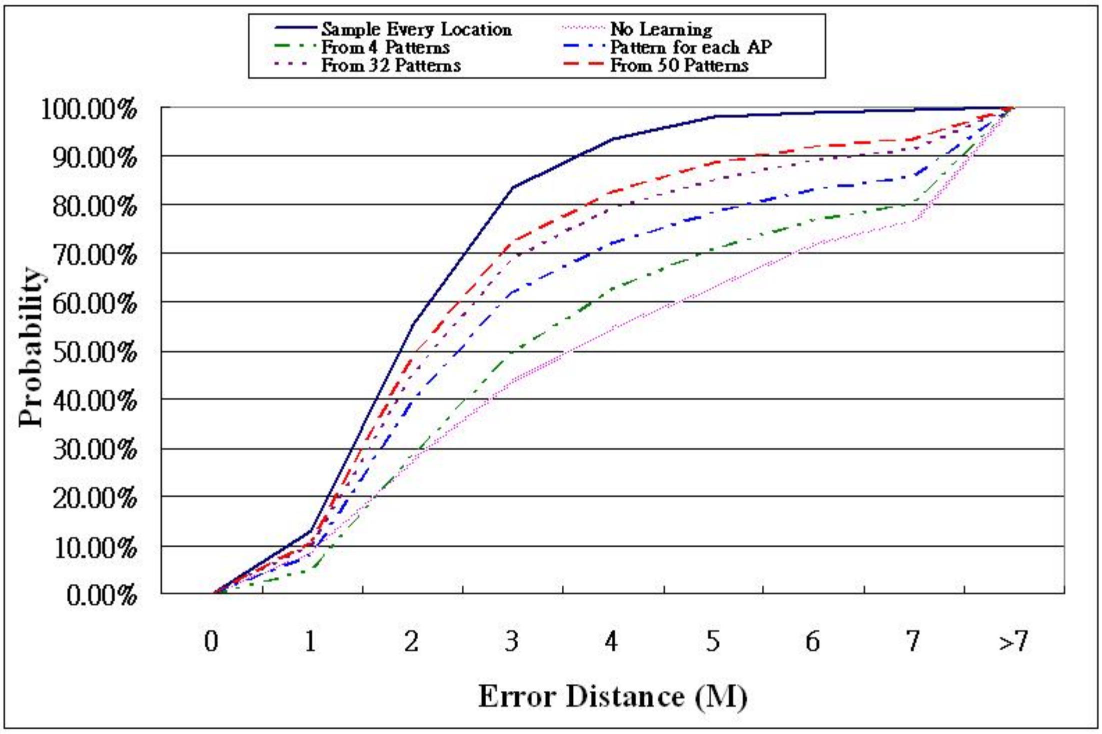

Figure 40: Comparison on 100x20 m^2 area (all positions).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	13	55	84	93	97	98	99
No Learning	8	27	44	55	63	72	77
4 Initial Samples	**5**	**28**	**49**	**63**	**71**	**77**	**80**
Sample Per AP	**8**	**39**	**62**	**72**	**78**	**83**	**86**
32 Initial Samples	**10**	**45**	**69**	**79**	**85**	**89**	**91**
50 Initial Samples	**10**	**49**	**72**	**83**	**88**	**92**	**94**

Table 6: Precision (%) tables on 100x20 m^2 area (all positions).

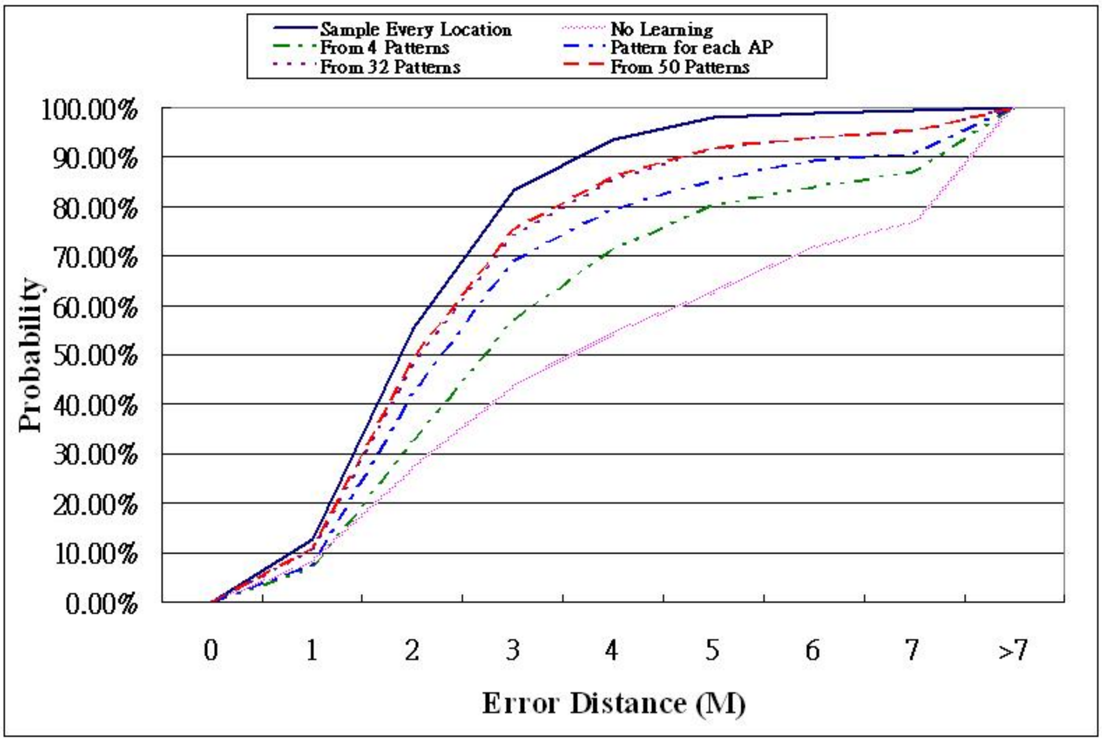

Figure 41: Comparison on 100x20 m^2 area (after training).

	< 1m	< 2m	< 3m	< 4m	< 5m	< 6m	< 7m
Sample Every Location	13	55	84	93	97	98	99
No Learning	8	27	44	55	63	72	77
4 Initial Sample	**7**	**32**	**57**	**71**	**80**	**84**	**86**
Sample Per AP	**8**	**42**	**69**	**80**	**85**	**89**	**91**
32 Initial Sample	**11**	**48**	**74**	**85**	**92**	**94**	**95**
50 Initial Sample	**11**	**49**	**75**	**86**	**92**	**94**	**95**

Table 7: Precision (%) table on 100x20 m^2 area (after training).

4.2.2 Extended Area Scenario

As described in Section 4.1.4, we setup an extended area scenario to verify our system in auto learning ability. First, we perform positioning in the original area (Simulation Model in Section 4.1.1 with one sample for each AP). Then we add another 3 AP in the extended area and perform positioning in it. Figure 42 shows our result.

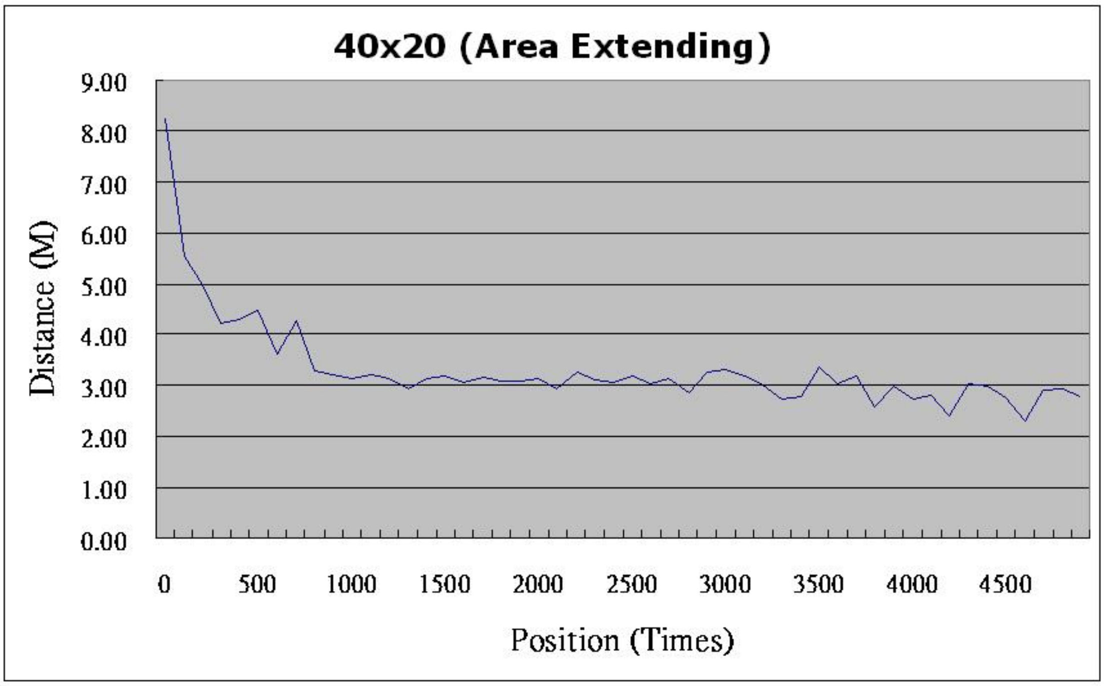

Figure 42: Average distance error on Extended Area Scenario.

Compared with Figure 28 (positioning in original area), although it spends more time for training the system to reach significant accuracy, it really proves that our system has the ability to do positioning in the extended area by just adding the AP and recording its locations into system table.

Chapter 5 Conclusions

In recent years, location awareness issue becomes interesting research topic in network communications area, especially with the advance of wireless network technology. As we know, GPS is the widely used and most popular location awareness service which can be seen everywhere. But GPS has its limitations as we discussed in the beginning of this thesis. Since WiFi LAN (Wireless LAN) is so popular now, implementing position service based on WiFi LAN should be feasible and affordable. It can provide positioning and communication service at the same time without any other system assistance.

In this thesis, we studied some measurements and introduced several proposed positioning systems. Then we proposed an intelligent positioning system which was built on application layer based on WiFi LAN. In our scheme, we adopt empirical and machine learning concept to provide the ability of self adjustment for environment change. This ability is very helpful for system maintenance and positioned area extension. Our experiment shows that the proposed scheme has significant accuracy after self training. And based on the result of Extended Position Area testing, it also has good effect on extended area by setup AP and registering its location only. It has similar accuracy with the area we built by measuring initial samples manually.

Our work is not considered as completed yet, we will investigate the following topics as future works:

- Implement the system on real-world environment.
- Combine more moving behavior or perform more efficient algorithm to improve the positioning accuracy.
- Reduce the self training time.

- Resolve user dependent issues.

Nowadays, more and more location awareness services have been proposed or developed. But most of them are not as popular as GPS for driving and traveling. If the position system is built on WiFi LAN with high precision, friendliness, flexibility, and low cost, much more useful services can be implemented. We believe there will be more epochal services to be devised in the near future.

Bibliography

[1] Stuart A. Golden and Steve S. Bateman, “Sensor Measurements for Wi-Fi Location with Emphasis on Time-of-Arrival Ranging”, IEEE Transactions On Mobile Computing, Vol. 6, No. 10, October 2007

[2] P. Bahl and V.N. Padmanabhan, “RADAR: An In-Building RFBased User Location and Tracking System,” Proc. INFOCOM, 2000

[3] J. Barnes, C. Rizos, J. Wang, D. Small, G. Voigt, and N. Gambale. (2003). Locata: The positioning technology of the future? presented at 6th Int. Symp. Satellite Navig. Technol. Incl. Mobile Positioning Location Services, Melbourne, Australia [Online]. pp. 49–62. Available: http://www.gmat.unsw.edu.au/snap/snap.htm

[4] J. Hightower, R. Want, and G. Borriello, “SpotON: An indoor 3D location sensing technology based on RF signal strength,” Univ. Washington, Seattle, Tech. Rep. UW CSE 2000–02-02, Feb. 2000.

[5] L. M. Ni,Y. Liu,Y. C. Lau, and A. P. Patil, “LANDMARC: Indoor location sensing using active RFID,” Wireless Netw., vol. 10, no. 6, pp. 701–710, Nov. 2004.

[6] J. J. Caffery and G. L. Stuber, “Overview of radiolocation in CDMA cellular system,” IEEE Commun. Mag., vol. 36, no. 4, pp. 38–45, Apr. 1998.

[7] V. Otsason, A. Varshavsky, A. LaMarca, and E. de Lara, “Accurate GSM indoor localization,” UbiComp 2005, Lecture Notes Computer Science, Springer-Varlag, vol. 3660, pp. 141–158, 2005.

[8] P. Prasithsangaree, P. Krishnamurthi, and P. K. Chrysanthis, “On indoor position with wireless LANs,” in Proc. IEEE Int. Symp. Pers. Indoor, Mobile Radio Commun., Sep. 2002, vol. 2, pp. 720–724.

[9] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen, “Experiments on local positioning with Bluetooth,” in Proc. IEEE Int. Conf. Inf. Technol.: Comput. Commun., Apr. 2003, pp. 297–303.

[10] J. Hallberg, M. Nilsson, and K. Synnes, “Positioning with Bluetooth,” in Proc. IEEE 10th Int. Conf. Telecommun., Mar. 2003, vol. 2, pp. 954–958.

[11] M. Youssef, A. Agrawala, and A. Udaya Shankar, “WLAN location determination via clustering and probability distributions,” IEEE Int. Conf. Pervasive Comput. Commun., Mar. 2003, pp. 143–151

[12] M. Youssef and A. K. Agrawala, “Handling samples correlation in the Horus system,” IEEE INFOCOM 2004, Hong Kong, vol. 2, pp. 1023–1031, Mar. 2004

[13] T. Roos, P.Myllymaki, H. Tirri, P.Misikangas, and J. Sievanan, “A proba- bilistic approach toWLAN user location estimation,” Int. J. Wireless Inf. Netw., vol. 9, no. 3, pp. 155–164, Jul. 2002.

[14] P. Castro, P. Chiu, T. Kremenek, and R. R. Muntz, “A probabilistic room location service for wireless networked environments,” in Proc. 3rd Int. Conf. Ubiquitous Comput., Atlanta, GA, Sep. 2001, pp. 18–34.