Summer 7-12-2012

A WLAN Fingerprinting Based Indoor Localization Technique

Landu Jiang
University of Nebraska-Lincoln, landueric@gmail.com

A WLAN FINGERPRINTING BASED INDOOR LOCALIZATION TECHNIQUE

by

Landu Jiang

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Computer Science

Under the Supervision of Professor Lisong Xu and Professor Zhigang Shen

Lincoln, Nebraska

July, 2012

A WLAN FINGERPRINTING BASED INDOOR LOCALIZATION TECHNIQUE

Landu Jiang, M. S.

University of Nebraska, 2012

Advisors: Lisong Xu and Zhigang Shen

Satellite-based Global Positioning Systems (GPS) have enabled a variety of location-based services such as navigation systems, and become increasingly popular and important in our everyday life. However, GPS does not work well in indoor environments where walls, floors and other construction objects greatly attenuate satellite signals. In this paper, we propose an Indoor Positioning System (IPS) based on widely deployed indoor WiFi systems. Our system uses not only the Received Signal Strength (RSS) values measured at the current location but also the previous location information to determine the current location of a mobile user. We have conducted a large number of experiments in the Schorr Center of the University of Nebraska-Lincoln, and our experiment results show that our proposed system outperforms all other WiFi-based RSS IPSs in the comparison, and is 5% more accurate on average than others.

ACKNOWLEDGMENTS

Firstly, I would like to express my heartfelt gratitude to my advisor and committee chair, Professor Lisong Xu and the co-advisor Professor Zhigang Shen for their constant encouragement and guidance throughout the course of my master's study and all the stages of the writing of this thesis. Without their consistent and illuminating instruction, this thesis work could not have reached its present form. Their technical and editorial advice and infinite patience were essential for the completion of this thesis. I feel privileged to have had the opportunity to study under them. I thank Professor Ziguo Zhong and Professor Mehmet Vuran for serving on my Master's Thesis defense committee, and their involvement has greatly improved and clarified this work. I specially thank Prof Ziguo Zhong again, since his support has always been very generous in both time and research resources.

I thank all the CSE staff and friends, for their friendship and for all the memorable times in UNL. I would like to thank everyone who has helped me along the way.

At last, I give my deepest thanks go to my parents for their self-giving love and support throughout my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Location Based Service

As the development of the communication networks and mobile computing, location based service (LBS) becomes very popular in recent years. The location-based service refers to the applications that rely on a user's location to provide services such as construction real-time locating, safety and health care [1], indoor navigating guidance, etc. The core of LBS is the positioning technique. The Global Positioning Systems (GPS) [2] are the earliest widely used modern systems for civilian positioning service, and can offer an accuracy close to 10 meters. However, GPS cannot provide good accuracy in indoor environments since the satellite signals are blocked by building obstructions.

An increasing number of indoor positioning systems have been proposed such as Cellular-network [3], Ultrasound [4], Computer Vision [5], Infrared Ray [6], Radio signal [7], Bluetooth technique [8], PHY information [9], etc. Most of these systems are able to provide accurate results, however, they rely on additional hardware or large-scale infrastructures. Thus, such systems are hard to be widely deployed due to significant cost, energy consumption and specific environment range limitations.

The indoor positioning techniques rely on different types of measurements involving Time-of-arrival (TOA), Time-difference-of-arrival (TDOA), Angle-of-arrival (AOA) [10, 37], Wireless Local Area Network (WLAN) Received Signal Strength (RSS), etc. As the IEEE 802.11 principle has become the industry standard, the WLAN RSS techniques draw great attention and enables a new layer of the indoor positioning approaches. Unlike other measurements that need additional hardware and synchronization schemes in indoor environments, the WLAN RSS techniques are economical solutions due to the wide deployment of wireless network infrastructures [11-13, 36, 39]. In addition, personal laptops, tablet computers or other mobile devices equipped with WLAN capability such as smart phones and powerful PADs are easy to be connected to WLAN systems. Hence, the WLAN RSS positioning systems take advantages on stability, flexibility and mobility.

1.2 WLAN Received Signal Strength Techniques

There are three fundamental methods using WLAN RSS measurement: the Strongest Base Station, the Propagation Model and Fingerprinting based method. The strongest base station method is the most simple solution in WLAN RSS techniques. The user's location is estimated as the position of the nearest data communication access point (AP), and this method has no computational issues and is applicable in most networks. However, the strongest base station method could not achieve good accuracy because of the complexity of indoor WLAN environments and limitation of the AP coverage.

In the propagation model method, the received signal strength information - signal path loss is taken into account to estimate the location of a mobile user. The system commonly uses a theoretically-calculated propagation model to convert the RSS path loss value to the physical distance from the base station side to the user. The coordinates of a user can be determined by the propagation model using the geometry techniques such as trilateration

[31] and triangulation [10, 39]. In the trilateration technique, at least three base stations with coordinates information are required to draw circles using the distances r_i from the user to base stations, and then we can locate the mobile user by the intersection of the three circles. In the triangulation technique, there could be fewer or more base stations that measure the orientation of the signal to estimate a user's location. The propagation model approach is relatively simple and efficient when the accuracy requirement is not very high. However, accurately measuring the distance based on signal attenuation is still difficult due to the noise of wireless signals and the interference of indoor obstructions such as multi-story floors, doors and walls [35].

Recently, WLAN RSS Fingerprinting becomes one of the most exploited techniques in indoor localization [14]. Compared to the strongest base station and the propagation model, the fingerprinting method is easy to deploy and is tolerant to wireless signal noise, and thus can achieve the highest accuracy. The fingerprinting systems normally consist of two phases: the offline training phase and the online determination phase. In the training phase, the goal is to build an empirical training database for each reference position by sampling the WLAN signal strength from several wireless access points [15, 16]. Then in the determination phase, the mobile user with a given RSS sample is estimated as the best matching location record in the training database. We will present a detailed discussion of WLAN RSS fingerprinting methods in later chapters.

1.3 Challenges in Fingerprinting Technique

In this paper, we aim at providing an accurate and efficient indoor positioning system based on IEEE 802.11 wireless technique. The system uses a fingerprinting method which creates a probability distribution map of the WLAN received signal strength (RSS) collected at known coordinates to estimate the location of a mobile user.

There are some challenges in designing an RSS fingerprinting-based indoor positioning system [17, 18].

Firstly, the IEEE 802.11 WLAN frequency range is in the 2.4 GHz public band which is also used by mobile phones, microwave ovens and other wireless signal transmitters. In the determination phase, any other devices in this public band can cause the irregular WLAN RSS patterns to mobile users as the source of the interference.

Secondly, M. Ghaddar et al. [19] and J. Ryckaert et al. [20] have observed that the blocking effect of human body on various frequencies and even indoor wireless signal quality. Thus a human user could weak the WLAN RSS value on the straight line between the mobile device and an AP.

Furthermore, the accuracy of the fingerprinting method relies on the long term WLAN RSS sampling. Any changes in the environment such as AP replacing and facilities upgrading can lead to a poor system performance [37]. Thus, a large amount of sampling work is required to maintain the training database which brings a heavy burden.

As the WLAN RSS fingerprinting indoor positioning performance is largely limited by the challenges mentioned above, our purpose is to provide effective solutions to overcome these challenges.

1.4 Contribution of Thesis

In this thesis, we present an accurate and efficient WLAN RSS fingerprinting indoor positioning system. We firstly provide a probabilistic framework using K most likely neighbors (KMLN) to determine the location of a mobile user, and then propose a novel tracking algorithm employing the shortest path scheme to enhance the estimation accuracy. To reduce the propagation error, we provide a study on integrating previous (historical) WLAN RSS observations in the tracking algorithm. We also analyze the human body orientation interference and the long term WLAN signal characteristics in indoor environments, providing the basis for our proposed positioning system. We conduct our experiments on the first floor of the Schorr Center which is used by the Department of Computer Science and Engineering. In our experiment, the proposed system shows very promising results and achieves better accuracy than other fingerprinting methods.

1.5 Outline of Thesis

This thesis is organized as follows. Chapter 2 presents a brief introduction of the fingerprinting techniques and its related approaches. Chapter 3 describes the methodology and framework of our proposed indoor positioning system. Chapter 4 describes the experiment setup and the analysis of WLAN RSS propagation characteristics, and a detailed algorithm performance comparison is also presented. Chapter 5 provides the conclusion of this thesis and discusses the possible future work.

Chapter 2

Related Work

Recently, WLAN Receive Signal Strength (RSS) Fingerprinting has become the most promising indoor positioning technique because of its easier deployment and lower cost compare to other methods. In this chapter, we firstly present a background study of the RSS Fingerprinting techniques, then an overview of the related work on employing fingerprinting methods is provided.

2.1 The Background Study of WLAN RSS Fingerprinting

A WLAN RSS fingerprinting system normally consists of two phases: the offline training phase and the online determination phase. In the training phase, the goal is to build an empirical training database for each reference location by sampling the WLAN signal strength from several wireless access points [15, 16]. Then in the determination phase, the mobile user with a given RSS sample is estimated as the best matching location record in the training database.

The basic process architecture of a WLAN RSS fingerprinting system is shown in

Figure 2.1. In the offline training phase (upper block), the RSS fingerprints are carefully sampled at each reference position (RP(i)(x,y)) as a vector $(RSS_{l_i1}, RSS_{l_i2}, ..., RSS_{l_in})$ from a number of access points (APs) (n is the number of APs and l_i is the ith Reference Position ID, and the signal strength range is from 0 dBm to 100 dBm) to build a WLAN RSS training database. In wireless networks, APs are usually fixed transmitters such as communication base stations. A reference position is a location in a WLAN indoor environment, and the signal fingerprints at each reference position are recorded in the training database. The received signal strength (RSS) is the measure of the signal power from an AP to a receiver which can be easily sampled in WLAN environments without additional requirements.

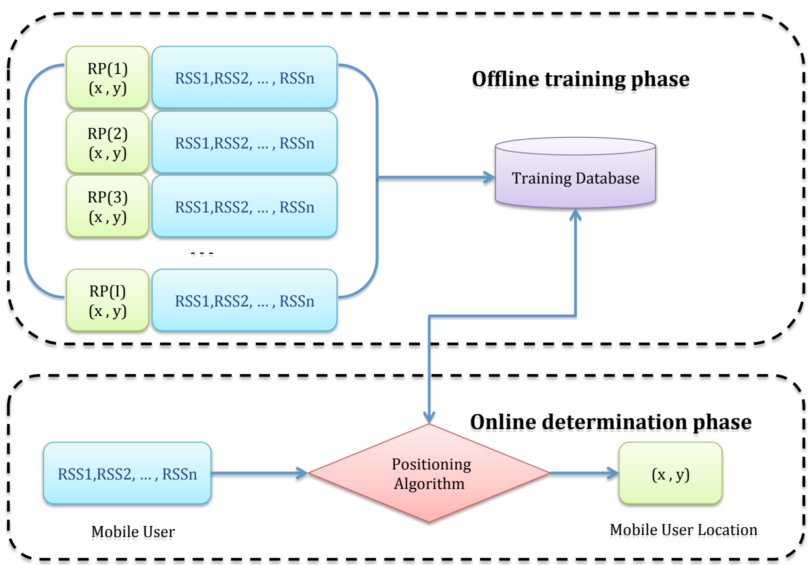

Figure 2.1: The diagram of the fingerprinting method

In the online determination phase (bottom block), a mobile user measures a vector of RSS values at an unknown location, then compares the wireless RSS vector records in the training database using a positioning algorithm, and finally calculates the most likely location of a mobile user. There are two basic positioning algorithms in WLAN RSS fingerprinting approaches [23] - Euclidean distance determination and Bayes rule determination.

The Euclidean distance determination [24] is a simple choice for RSS fingerprinting that measures the distance between an online RSS value and the offline training database RSS records. K Nearest Neighbors (KNN) and the weighted K nearest neighbor method (WKNN) are the basic schemes that are generally used for mobile user indoor positioning estimation:

$$EuDis = \sqrt{\sum_{i=1}^{n}(RSS_i - RSS_{i'})^2} \qquad 2.1$$

In Equation 2.1, n is the number of APs, RSS_i is the ith AP's signal strength received in the online phase and $\overline{RSS_i}$ is the average RSS value in the training database. The location of the mobile user is estimated by averaging the coordinates of the K neighbors with the minimum Euclidean distance. The value of Kcan influence the result accuracy, and if $K = 1$ the algorithm calculates the nearest neighbor.

In a complex indoor environment, the variation (Euclidean distance) of the RSS measured at each reference position could be very large. Therefore, the Bayes rule determination is proposed to achieve a more accurate estimation [18, 25]. The Bayes rule determination uses the probabilistic method to find the most possible location l_i out of the reference positions set given the observation RSS vector that maximizes the conditional probability $p(l_i/RSS)$, Following the Bayes rule, the $p(l_i/RSS)$ can be calculated as follows:

$$p(l_i/RSS) = \frac{p(RSS/l_i)p(l_i)}{p(RSS)} \qquad 2.2$$

where $p(RSS/l_i)$ is the conditional probability of obtaining the RSS at the ith location l_i, which can be approximated by the number of times that RSS signal strength vector $(RSS_{l_i1}, RSS_{l_i2}, ..., RSS_{l_in})$ appears at location l_i according to training database records. In [26], the conditional probability is calculated as the marginal probability: $p(RSS/l_i) = p(RSS_1/l_i)p(RSS_2/l_i)p(RSS_3/l_i) ... p(RSS_n/l_i)$.

$p(l_i)$ is the prior probability of being at position l_i, and this brings a new way to use prior position information that enables the mobile user tracking algorithms. In addition, $p(RSS)$ does not depend on location l_i and is often regarded as the normalizing constant.

2.2 An Overview of Related Approaches

RADAR [24] is an early approach using WLAN RSS to establish an indoor positioning system, which combines the empirical fingerprinting method and the theoretical propagation model to locate and track a mobile user. Since not all of the K nearest neighbors contribute to the positioning result, RADAR experiments both KNN and weighted KNN (WKNN) schemes to estimate the location of a mobile user. Kaemarungsi et al. [27] present two weighing schemes in WKNN: one scheme is based on the number of sampling points and the other one uses standard deviation of RSS samples as the neighbor weight. The Cluster Filtered KNN (CFK) approach [34] uses the clustering technique on K nearest neighbors determination to achieve a better estimation of the user location. Fang et al. [28] uses a Neural Network based model to determine the position of a mobile user inside a working area, and the performance of their system is close to WKNN algorithms.

P. Castro et al. [29] estimate the location of a mobile user using Bayes rule. Myllymaki et al. [30] formulate it as a machine learning problem using a probabilistic framework to estimate the indoor mobile user location. Youssef et al. [26, 33] propose the location system using joint probability distribution and location-clustering method called joint clustering technique. They firstly take the positioning computational burden into account and achieve higher accuracy than previous work.

The above algorithms provide a number of solutions in indoor positioning determination using Euclidean distance and Bayes rule. However, they do not take the historical information such as the topology knowledge of a mobile user's prior positions and the previous WLAN RSS data information into account.

Altintas et. al [21] present a short term memory scheme using previous (historical) WLAN RSS observations to smooth the error distance during the online determination phase. IBM researchers [22] consider the prior probability $P(l_i/L)$, and they suppose that a moving user should follow the basic topology rules such as the user has a limited moving speed. The definition of the priori probability $P(l_i)$ in their tracking assistant algorithm is presented as follows:

$$P(l_i/L_k^P, L_{k-1}^P, ..., L_1^P) = \frac{1}{k \times D} \sum_{j=1}^{k} (e^{-(j-1)} \times dist^{-1}(l_i, L_j^P)) \qquad 2.3$$

$L_k^P, L_{k-1}^P, ..., L_1^P$ are the k determined positions prior to the new location l_i. D is a constant that normalizes the tracking probability. $dist^{-1}(l_i, L_j^P)$ is the tracking probability reversely proportion to the distance between the current position l_i and the priori location L_j^P. The shorter distance to the priori position set, the higher probability of the current position.

Their tracking assistant algorithm uses the conditional probability by considering the topology knowledge, however, the system may have low positioning accuracy in some cases. For instance, a wrong location may get a much higher probability $p(RSS/l_i)$(in experiments it can be 10 times higher) than the actual position, if the actual position's probability $p(l_i)$ is not high enough.

Chapter 3

Problem Setting and Our Approach

In this chapter we firstly describe the problem, and then present our proposed indoor positioning system.

3.1 The Indoor Positioning Problem

There are two types of users in an indoor positioning system: stationary users and mobile users. A stationary user stays at a location forever (or for a very long time period), and a mobile user moves within a building. It is relatively easier to determine the location of a stationary user than the location of a mobile user. This is because we can collect the RSS samples of a stationary user at the same location for as many time as needed to improve the accuracy, however, very few RSS samples of a mobile user can be collected at one location because the user is moving. In this thesis, we consider how to determine the location of a mobile user in a building using the indoor WiFi system.

3.2 Our Approach

3.2.1 Overview

Our approach follows the diagram of a general fingerprinting method as illustrated in Figure 2.1, and consists of two phases.

- 1) Offline training phase which collects RSS samples at reference positions and builds a training database,
- 2) Online determination phase which determines the location of a mobile user by comparing the measured RSS values with the training database.

Our offline training phase is very similar to the general offline training phase described in Chapter 2, and more details will be given in Chapter 4. Below, we focus on the online determination phase.

Our online determination phase uses two algorithms to determine the location of a mobile user.

- 1) K Most Likely Neighbor (KMLN) Algorithm which determines the K most likely locations of a mobile user.
- 2) Shortest-Path-Based Tracking Algorithm which determines the location of a mobile user by using the current and past location information of the user.

3.2.2 K Most Likely Neighbor (KMLN) Algorithm

We propose K Most Likely Neighbor (KMLN) algorithm to determine the K most likely locations of a user. Among these K locations, we finally select one location using the shortest-path-based tracking algorithm which is described in the next subsection.

Recall that Chapter 2 introduces two types of determination algorithms to determine the location of a stationary user.

- Type 1: Euclidean Distance Determination which selects the K most nearest neighbors based on the Euclidean distance, and then returns the average of these K locations as an estimate of the current location of a user.

- Type 2: Bayes Rule Determination which selects the most likely location using Bayes Rule.

The above two types of algorithms are used to determine the location of a stationary user, and thus they finally return only one location.

Our proposed KMLN is inspired by and combines the above two types of determination algorithms. The pseudocode of KMLN is shown in Algorithm 1. Specifically, KMLN selects the K most likely locations using Bayes Rule. The reason that we use Bayes rule instead of Euclidean distance is that Bayes rule is more robust and can achieve higher accuracy in cases of poor WiFi signals with noises which are very common in indoor environments. The reason that we select the K most likely locations instead of the most likely location is that our experiments show that sometimes the actual location may not be the most likely location. We use KMLN only to select the K most likely locations, and then use the shortest-path-based tracking algorithm to finally select one location.

3.2.3 Shortest-Path-Based Tracking Algorithm

We propose Shortest-Path-Based Tracking algorithm to determine the current location of a mobile user. The proposed algorithm is based on one important assumption: a mobile user is walking at a relatively slow speed (i.e., not running at a relatively fast speed) inside a building. This assumption has the following two implications.

Algorithm 1 The KMLN algorithm

The $Tagnumber$ = the number of the APs n.
for each RSS value from AP_j RSS_i **do**
  for each AP propagation at Location j $P_j(i)$ **do**
    if $RSS_i > 0$ **then**
      $P_j(i)$ = the probability of RSS_i value in the Histogram Distribution.
    else
      $P_j(i) = 1$.
      $Tagnumber = Tagnumber$ -1.
    end
  end
end
for each $P_j(i)$ in stack **do**
  $P_{KMLN}(j) = (P_j(1) * P_j(2) * P_j(3) * ... * P_j(n))^{(1/Tagnumber)}$
end
$KMLNStack$ = Sorting the array P_{KMLN} in the descending order.
return $KMLNStack(1:K)$

First implication: in a short time period, such as less than one second, a mobile user with moving locations can be considered as a stationary user with a fixed location. Therefore, the RSS values continuously measured within a short time period at slightly different locations can be considered as RSS values measured at the same location, and then are used to determine the location of the user.

Second implication: within two or three consecutive time periods, the locations of a mobile user are not too far away from one another. Our proposed shortest-path-based tracking algorithm is inspired by this implication. Let L_i denote the set of K most likely locations selected by KMLN in time period i. For each location in L_i, we calculate the physical distance between it and each location in sets L_{i-1} and L_{i-2}. Finally, we select the location with the shortest distance as an estimate of the user location in time period i. In cases of ties where multiple locations with the same shortest distance, we use the average of these locations as an estimate of the user location in time period i. The pseudocode of the algorithm is shown in Algorithm 2.

Algorithm 2 The Shortest Path based Tracking Algorithm

for each RP_j in current step i KMLNStack location L_i^j **do**
  for each RP_k in previous step $i-1$ and step $i-2$ KMLNStack location L_{i-1}^k and L_{i-2}^k**do**
    if $i \geq 3$ **then**
      $TrackingA(j,k)$ = Distance of (L_i^j - L_{i-1}^k).
      $TrackingB(j,k)$ = Distance of (L_i^j - L_{i-2}^k).
    end
  end
end
      $TrackingAmin$ = min $TrackingA(j,k)$.
      $TrackingBmin$ = min $TrackingB(j,k)$.
      SPA = Find L_i^j in KMLNStack where $TrackingA(j,k)$ equals$TrackingAmin$
      SPB = Find L_i^j in KMLNStack where $TrackingB(j,k)$ equals$TrackingBmin$
      $SPStack = SPA \cup SPB$
return The average coordinates of the set SPStack

In order to avoid cumulative errors [22] when using the past location information, we monitor the distance between the locations in two consecutive time periods. If the distance form the current position to the previous position is longer than a threshold d_0, that means the location estimated by the shortest-path-based tracking algorithm is possibly too far away from the actual location. In this case, we use the average of the K most likely locations selected by KMLN (KMLN state in the figure) as the current step estimation. In addition, if the average of K most likely locations has a shorter distance than shortest-path-based tracking estimation to the previous position, we also choose the average of KMLM as an estimate of the location of a mobile user.

3.2.4 Algorithms Using the Median RSS Values

We also study another slightly different algorithm in which KMLN selects the K most likely locations using the median RSS value of each AP in time periods i, $i-1$, and $i-2$, instead of the RSS value of each AP in time period i. The advantage is that we can filter out some

RSS noises using the median RSS values [21]. The disadvantage is that a mobile user is more likely at three different locations in these three time periods, and thus this may sometimes filter out the actual RSS values.

3.2.5 Missing Data Handling

A special study is required for handling the missing values associated with the cases in which the signal of some access points are not observed at all. In our work, if the signal fingerprint at location l_i has sampled any RSS information from a specific AP_j, the obvious choice is to set that distribution to $p(RSS_j/l_i) = 1$. When there is an unknown AP to the online sampling vector and the training database, we simply ignore the AP RSS value.

Chapter 4

Experiments and Evaluation

In this chapter, we describe how we evaluate our proposed approach using real-world experiments, and also discuss the human body orientation effect and long-term RSS characteristics.

4.1 Experiment Test Bed Setup and Data Collection

4.1.1 Experiment Setup

We conduct our experiments on the first floor of the Schorr Center at the University of Nebraska-Lincoln Figure 4.1. This is a two-story building used by the faculty and students of the Department of Computer Science and Engineering. The first floor of the building is covered by several wireless APs, and we do not know the physical locations and transmission ranges of these APs.

As shown in Figure 4.2, We have selected 60 reference positions on the first floor of the building, including the lobby (area A), hallway (area B), and research lab rooms (area C). We have also chosen 10 tracking lines (i.e., paths) for evaluating our approach. Each tracking line has 12 points (2 pre-sampling points for tracking reference and 10 points for positioning test), thus there are a total of 100 testing points for the evaluation.

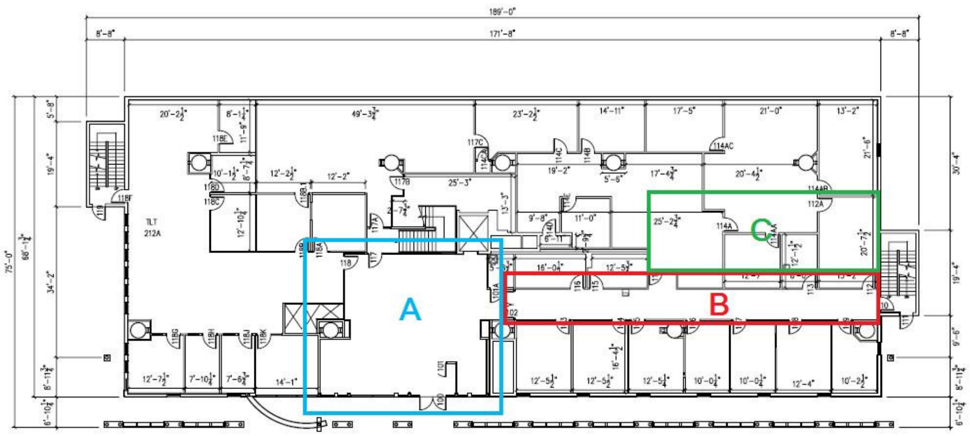

Figure 4.1: The map of the target building

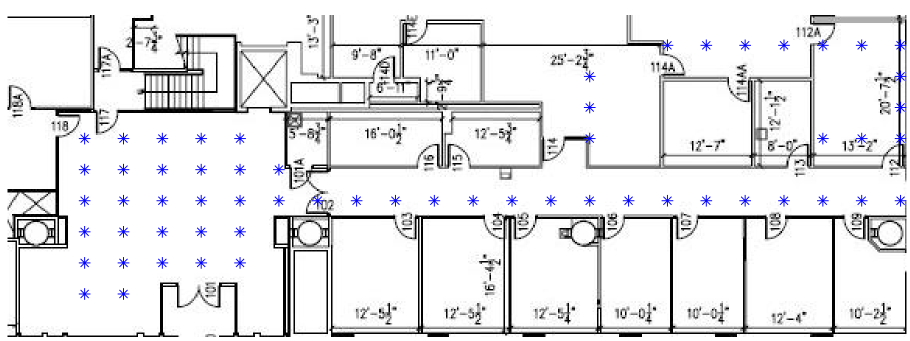

Figure 4.2: The 60 Reference Positions in the building

Table 4.1: AP MAC address list

AP No	MAC Address
AP01	00:17:df:ab:98:f1
AP02	00:17:df:aa:fa:21
AP03	00:17:df:ab:99:51
AP04	00:17:df:aa:fb:11
AP05	00:19:a9:b5:15:e0
AP06	00:1c:0f:82:b7:b0
AP07	1c:aa:07:c7:86:d1
AP08	00:27:0d:0b:4d:41

4.1.2 Data Collection

To capture the RSS data, we use a Sony personal laptop with a normal wireless network interface card (NIC). We also developed a WiFi RSS sampling application using Matlab 2010 and Windows Network Shell (netsh) command-line scripting utility. The application retrieves the basic information of each wireless AP detected by the NIC, such as the MAC address and RSS values of each AP.

We use the RSS sampling application to collect the AP information at the predetermined 60 reference positions. The map of these 60 positions is shown in Figure 4.2. The distance between two horizontally or vertically adjacent positions is about 1.5 meters (or 5 feet). For each position, we collect RSS values at different times of a day and at different days of a week in order to create a more comprehensive AP RSS value database. We have detected a total of 8 APs, and Table 4.1 shows the the MAC addresses of these 8 APs. Table 4.2 shows part of the AP RSS value database. Among these 8 APs, we select the top 4 APs (AP1, AP2, AP3, AP4 in the table)with the strongest RSS values to be used in our experiments.

Table 4.2: WLAN RSS Sample training database profile

RP Number (1-60)	AP 1	AP 2	AP 3	AP 4	AP 5	AP 6	AP 7	AP 8
RP 15	77	0	0	46	0	0	0	0
RP 15	77	99	88	58	0	0	0	0
RP 15	75	99	88	58	0	0	0	0
RP 15	79	99	93	0	0	0	0	0
...	...	...	...	...	...	...	...	...
RP 15	79	98	0	68	23	0	0	0

4.2 Signal Propagation Analysis

4.2.1 Impact of Human Body Orientation

In this subsection, we discuss the impact of human body orientation on the accuracy of indoor AP localization systems. Zhang et al. [38] shows that the body of a user could be an obstruction blocking a portion of WiFi signals. Specifically, WiFi signals are strong at the line of sight (LOS) propagation from an AP to a user, and is weak when the user is at the opposite orientation and blocks the signal. Based on this fact, they develop an outdoor AP localization system which determines the location of an AP by rotating the body of a user.

To study the impact of the human body orientation, we select 4 reference positions. At each of these 4 reference positions, we measure RSS values at different rotational angles ranging from 0 degree to 315 degree. The results are shown in Table 4.3. From the results, we can see that the impact of human body orientation in our indoor environment is very small. Therefore, we will not consider the human body orientation problem in our later experiments.

Table 4.3: Human factor in orientation of the RSS value (in dBm)

Rotational Angles	0°	45°	90°	135°	180°	225°	270°	315°
Position 1 AP 01	88	88	88	88	88	88	88	88
Position 1 AP 02	99	99	99	99	99	99	99	99
Position 1 AP 03	99	99	99	99	99	99	99	99
Position 1 AP 04	40	66	66	66	66	66	66	66
Position 2 AP 01	88	88	88	88	88	88	88	88
Position 2 AP 02	99	99	99	99	99	99	99	99
Position 2 AP 03	99	99	99	99	99	99	99	99
Position 2 AP 04	40	66	66	66	66	66	66	66
Position 3 AP 01	88	88	88	88	88	88	88	88
Position 3 AP 02	99	99	99	99	99	99	99	99
Position 3 AP 03	99	99	99	99	99	99	99	99
Position 3 AP 04	40	66	66	66	66	66	66	66
Position 4 AP 01	88	88	88	88	88	88	88	88
Position 4 AP 02	99	99	99	99	99	99	99	99
Position 4 AP 03	99	99	99	99	99	99	99	99
Position 4 AP 04	40	66	66	66	66	66	66	66

4.2.2 Long Term Signal Propagation Analysis

In this subsection, we study the long term signal propagation and discuss its impact on indoor AP localization systems. We randomly pick 10 reference positions, and at each position we measure $800 - 1000$ RSS values from 9 AM to 5 PM for three weeks using the same laptop.

Figure 4.3 to Figure 4.6 show part of the measurement results of four reference positions. We have the following observations. 1) The RSS value of some AP at some position changes frequently. For example, AP2 at position 1, AP1 at position 2, and AP1 at position 4. 2) The RSS value of some AP at some position is stable in most of the time, but has sharp changes occasionally. For example, AP3 at position 2. 3) The RSS value of the same AP has different patterns at different positions. For example, AP1 is relatively stable at position 1 but changes frequently at position 2. Some possible reasons for these observations are

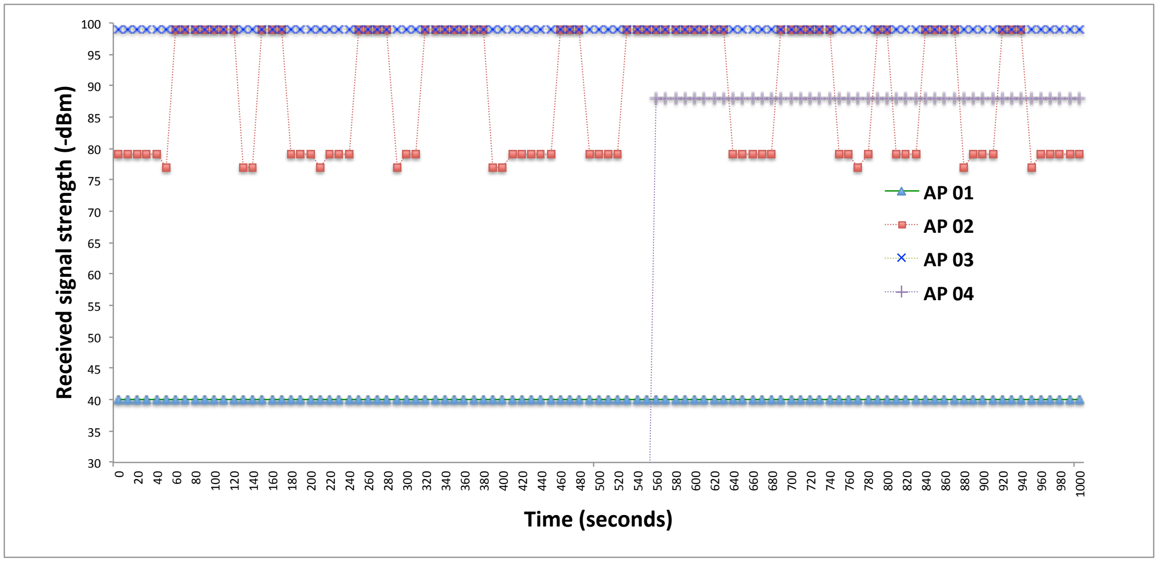

Figure 4.3: The long term WLAN signal propagation at Position 01

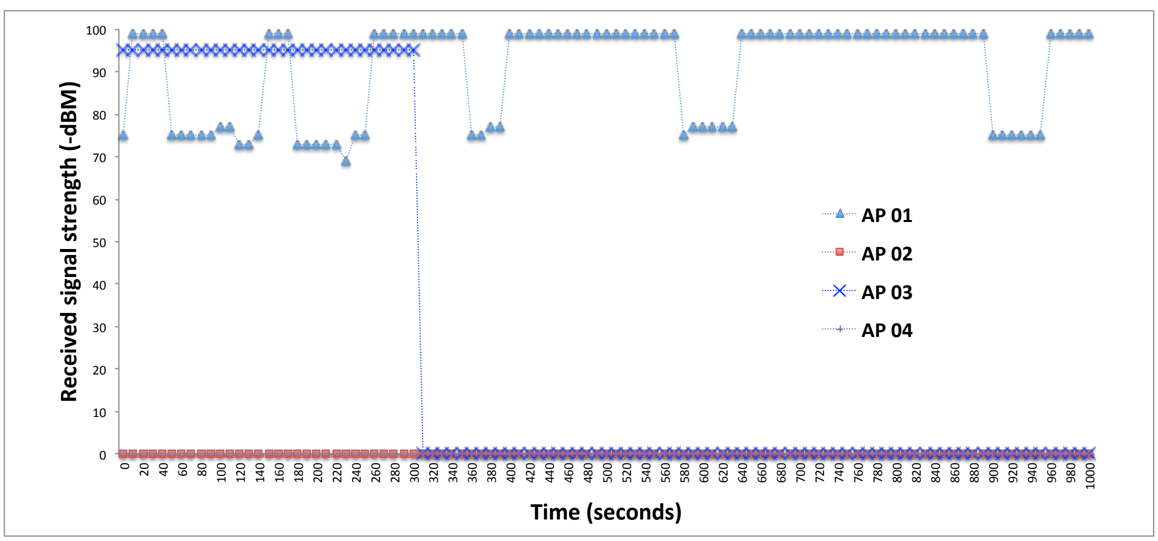

Figure 4.4: The long term WLAN signal propagation at Position 02

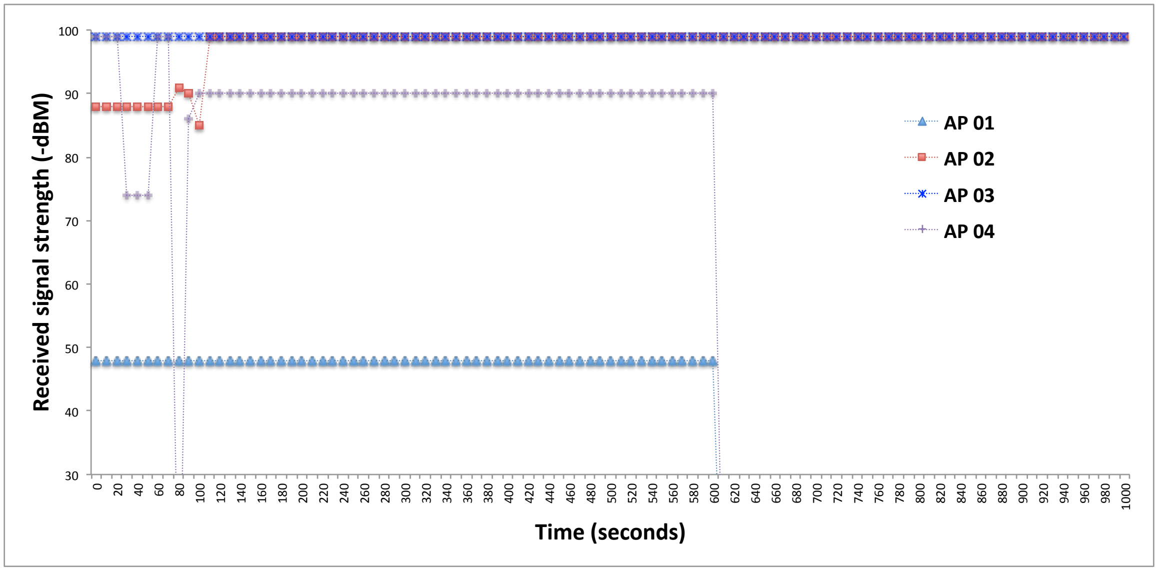

Figure 4.5: The long term WLAN signal propagation at Position 03

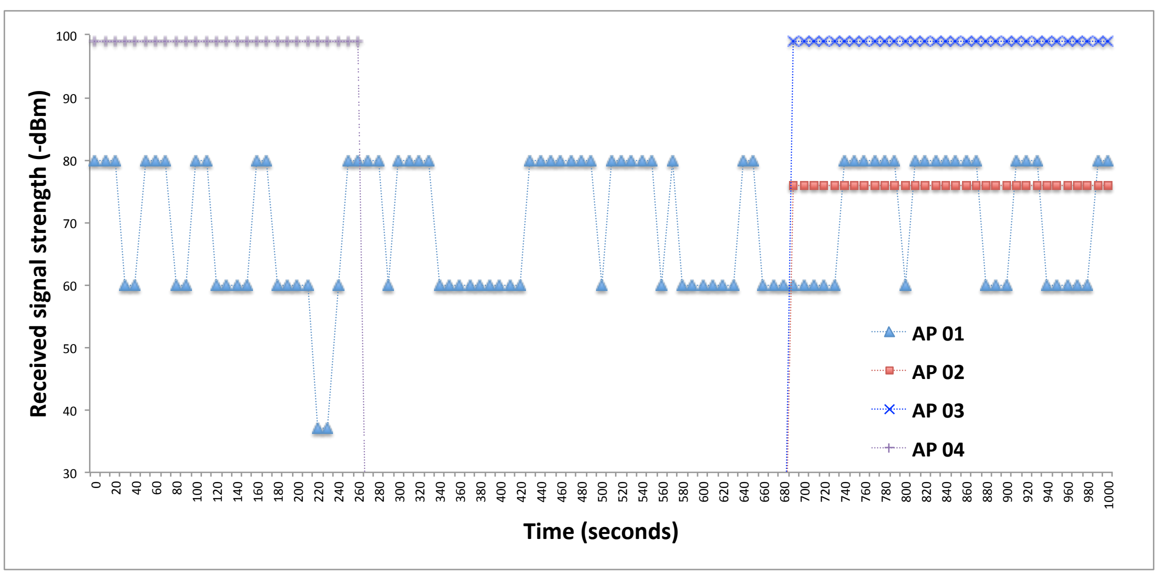

Figure 4.6: The long term WLAN signal propagation at Position 04

human activities, microwave oven, door opening and closing, and AP status change (e.g., from active to be non-active).

Based on these measurement results, we believe that a simple signal propagation model (e.g., with a single RSS value for an AP) is not sufficient in a complicated indoor environment. Therefore, we measure and manage a reasonable number of RSS values for each AP. We also use a signal-distribution shaping filter mentioned in [22] which enables a relative smaller number of measurements in training database to construct a long-term characteristics of the WLAN RSS propagation. The distribution shaping filter well handles the training phase workload and improves the performance of the positioning system at the same time.

The shaping filter is shown as follows, M is the number of the scanning operations in training phase, O_j^m is the RSS observation of AP j in the mth scanning operation, and s is the RSS value from 1 dBm to 100 dBm.

$$P_i^j(s) = \sum_{m=1}^{M} e^{-(\left|s-o_j^m\right|)} / E$$

where

$$E = \sum_{s=0}^{100} \sum_{m=1}^{M} e^{-(\left|s-o_j^m\right|)}$$

4.3 Algorithm Evaluation

In this section, we evaluate our proposed algorithms by conducting experiments on the first floor of Schorr center.

We evaluate the following algorithms.

- Group 1: Algorithms based only on the RSS values at the current location. This group of algorithms are proposed to determine the location of a stationary user. We use them as reference algorithms.
  - Algorithm 1: Generic Probabilistic Distribution (GPD) which directly uses Equation 2.2 described in Chapter 2.
  - Algorithm 2: K most likely neighbors (KMLN) which is described in Chapter 3. We use the average of the K most likely locations as an estimate of the current location of a mobile user. Note that, GPD is a special case of KMLN with $K = 1$.
- Group 2: Algorithms using the previous location information. This group of algorithms use not only the RSS values at the current location but also the previous location information.
  - Algorithm 3: Topology-based Tracking Algorithm which is similar to the algorithm proposed by IBM[34]. The difference is that this topology-based tracking algorithm uses KMLN instead of the joint probability method.
  - Algorithm 4: Our proposed Shortest-Path-based Tracking Algorithm which is described in Chapter 4. The time period is set to 2 seconds, and the distance threshold d_0 is set to 7 meters.
- Group 3: Algorithms using the median RSS values. This group of algorithms are very similar to the above four algorithms, and the only difference is that they use the median values of the RSS values measured at the most recent three locations instead of the RSS values at the current location.
  - Algorithm 5: GPD using Median RSS
  - Algorithm 6: KMLN using Median RSS

Table 4.4: The comparison result

Localization Technique	1.5 m	2 m	2.5 m	3 m	4 m	4.5 m	5 m	6 m
GPD	10%	14%	15%	27%	35%	46%	48%	58%
KMLN	20%	28%	33%	47%	53%	61%	66%	75%
Topology Tracking	20%	25%	36%	43%	52%	61%	63%	73%
Shortest-Path Tracking	24%	29%	37%	52%	57%	65%	69%	78%
GPD (Median)	6%	10%	10%	20%	30%	38%	41%	54%
KMLN (Median)	17%	24%	30%	43%	50%	56%	61%	71%
Topology Tracking (Median)	16%	24%	28%	38%	48%	52%	56%	67%
Shortest-Path Tracking (Median)	18%	23%	30%	40%	49%	55%	61%	72%

– Algorithm 7: Topology-based Tracking Algorithm with Median RSS

– Algorithm 8: Shortest-Path-based Tracking Algorithm with Median RSS

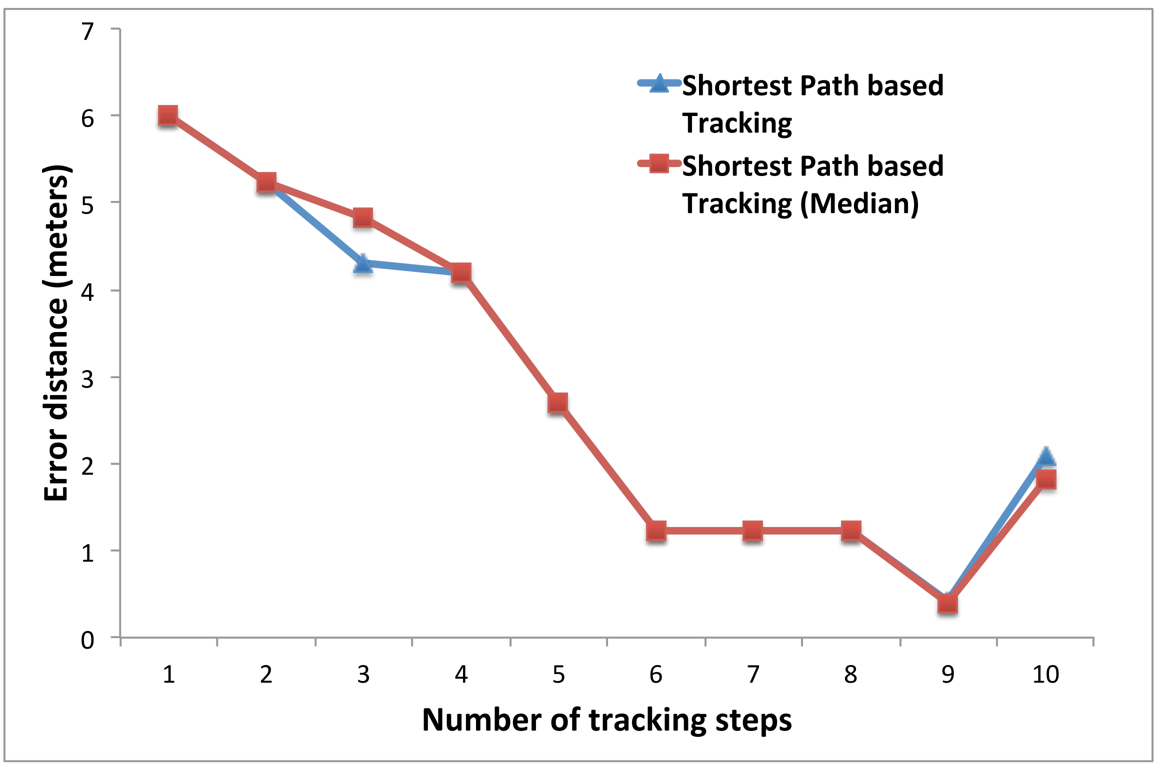

Figure 4.7: The Shortest-Path-based tracking algorithm performance

Figure 4.7 shows the error distance of shortest-path-based tracking algorithm without and

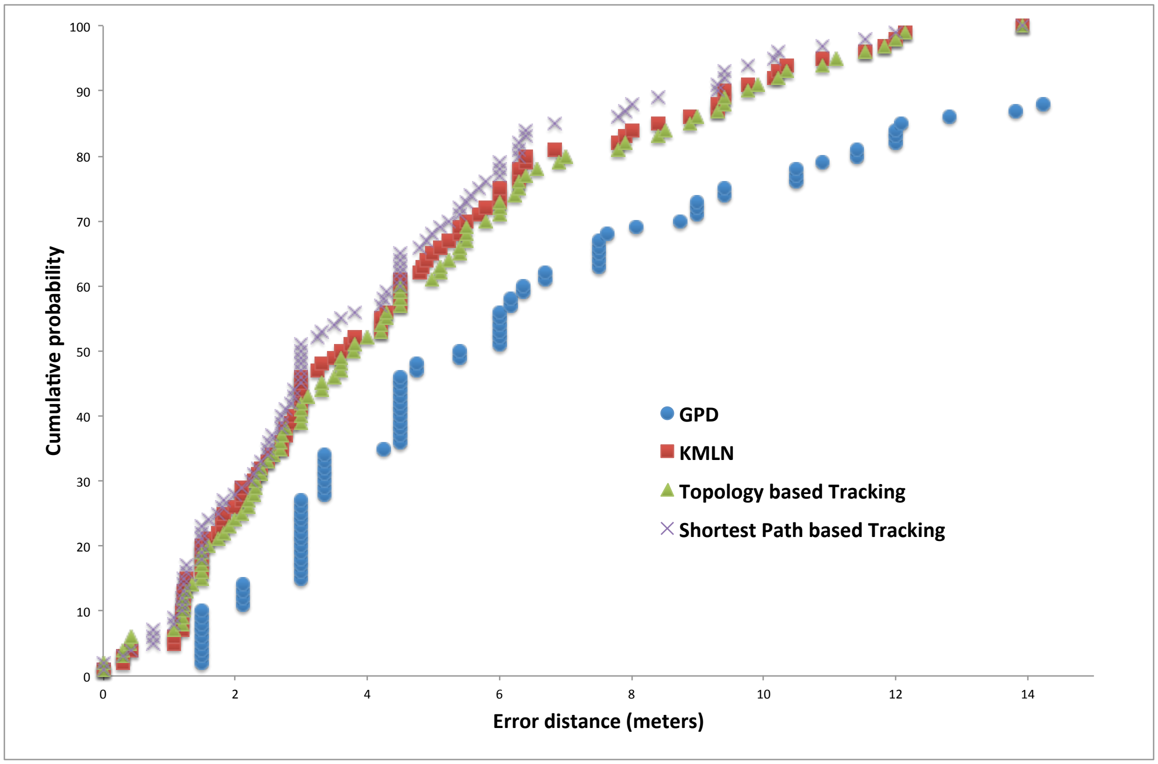

Figure 4.8: The algorithm performance without median RSSl

with median RSS for one tracking line. The error distance is the distance between the actual location and the estimated location. Recall that each tracking line has a total of 12 points. We only show the results for the last 10 points, and this is because that shortest-path-based algorithm requires the information of the previous two positions. We can see that initially at point 3 the error distances are very big (about 6 meters or 30 feet). The error distances become smaller and smaller, finally are around 1 or 2 meters. We also notice that there is no big difference between shortest-path-based tracking algorithm without and with median RSS.

Figure 4.8 and Figure 4.9 show the cumulative probability of the error distance of all eight algorithms. The cumulative probability is calculated by considering the error distance of a total of 100 points of all 10 tracking lines. Again, for each tracking line, we do not

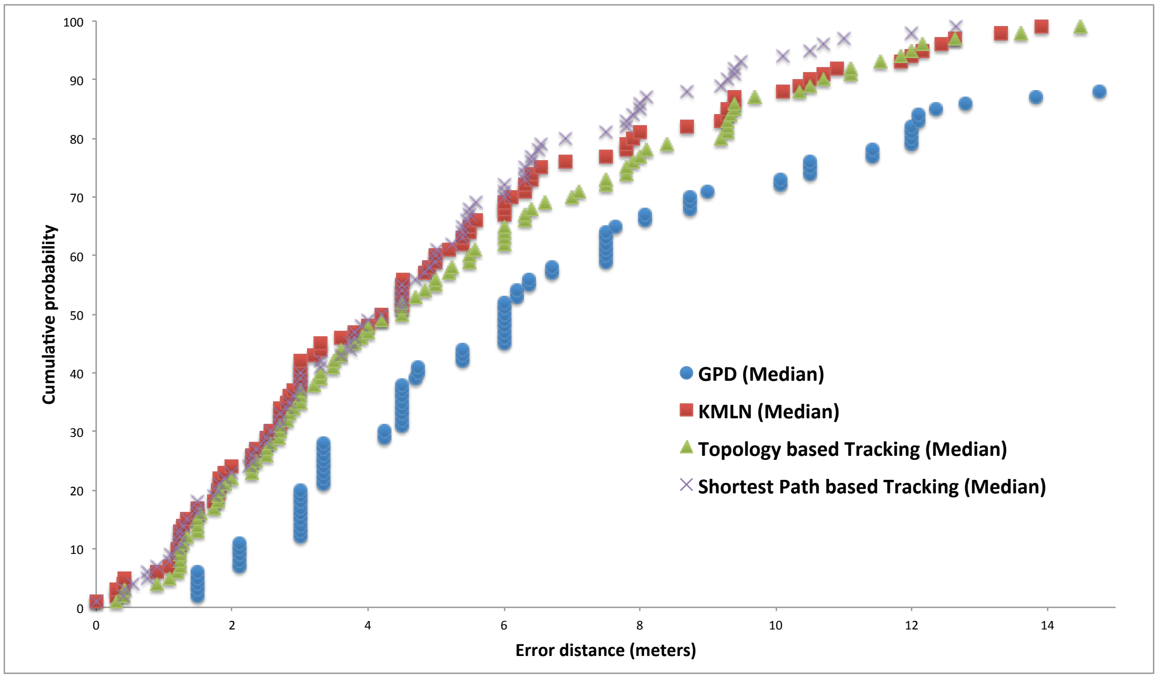

Figure 4.9: The algorithm performance using median RSS

consider the first 2 points, and only consider the remaining 10 points. To help you read the two figures, we also show the same cumulative probability information using Table 4.4, in which, the top row is the error distance. We can see that for both without and with median RSS, shortest-path-based tracking algorithm is more accurate than (5% higher) the other three algorithms: GPD, KMLN, and topology-based tracking algorithm. In addition, as demonstrated in Figure 4.7 , the average error distance of shortest-path-based tracking algorithm becomes smaller for longer tracking lines. Thus, we expect that shortest-path-based tracking algorithm would achieve even better average accuracy in Figures 4.8 and 4.9 if we have longer tracking lines.

We notice that in Figure 4.8 and Figure 4.9, shortest-path-based tracking algorithm achieves slightly better accuracy than shortest-path-based algorithm with median RSS. To more clearly show their difference, Figure 4.10 shows the cumulative probability of shortest-

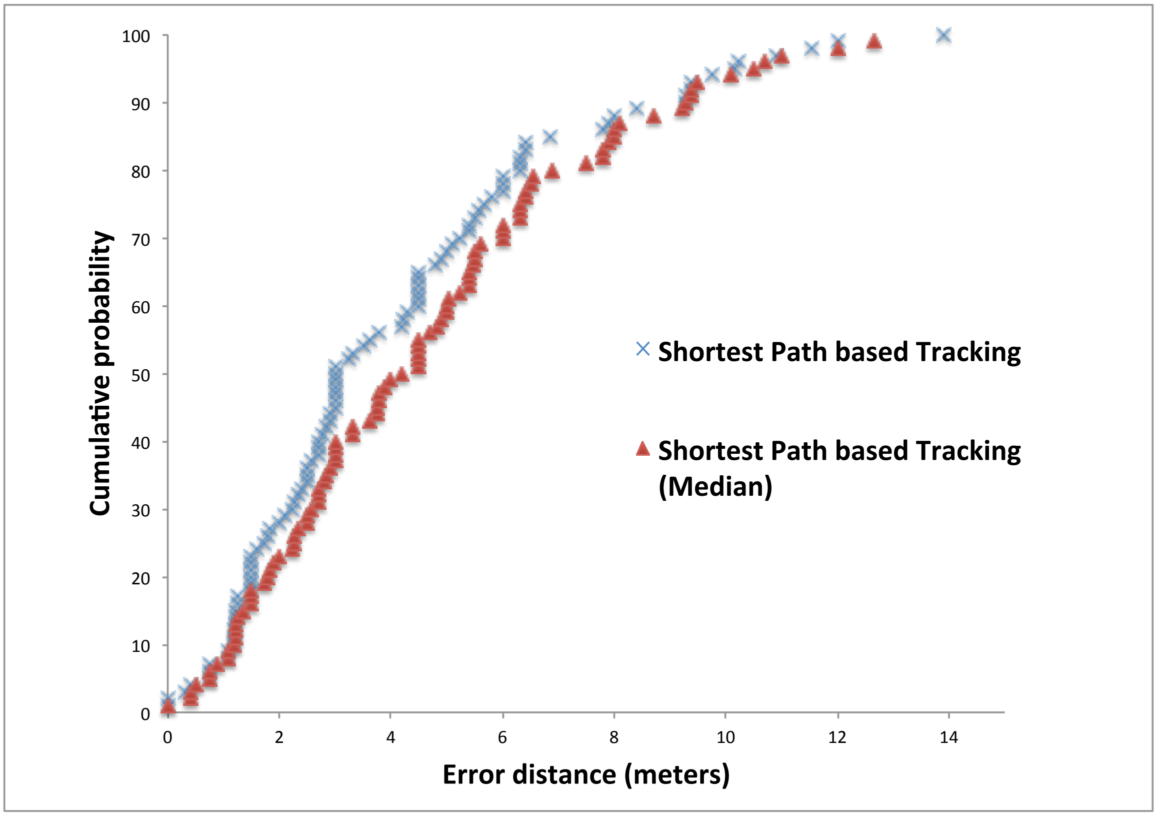

Figure 4.10: The comparison of Shortest-Path-based tracking algorithm performance without/with the median RSS

path-based algorithm without and with median RSS algorithms only. But we also notice that for some tracking lines, shortest-path-based algorithm achieves slightly worse accuracy than shortest-path-based algorithm with median RSS. For example, Figure 4.11 shows the error distances of one tracking line. We plan to have a more detailed study of the impact of median RSS in the future.

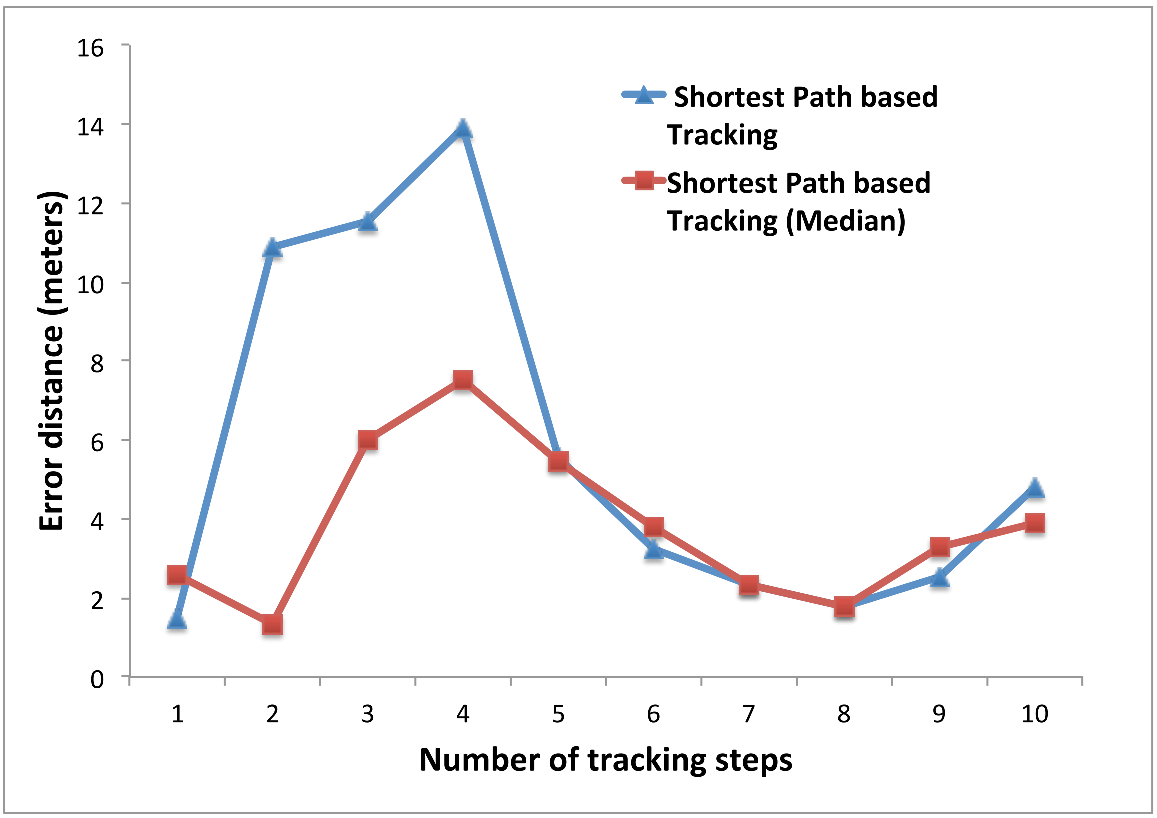

Figure 4.11: Better performance of the Shortest-Path-based tracking algorithm using median RSS

Chapter 5

Conclusions and Future Work

5.1 Conclusions

In this thesis, we first introduced the WLAN indoor location determination problem, and then propose an RSS fingerprinting indoor positioning system. To enhance the accuracy, a well-designed K most likely neighbors scheme and a tracking algorithm considering the previous position geometry are proposed. A discussion on human body interference and the analysis of the indoor WLAN signal characteristics are briefly demonstrated. In our evaluation, we setup a test bed in our CSE department building with 60 RPs and 4 APs. We examined the performances of the Generic Probabilistic Distribution, KMLN Probabilistic method, Topology-based tracking scheme and our proposed Shortest-Path-based Tracking algorithm. Our algorithm performs superiorly in both without and with median WLAN RSS case comparisons.

5.2 Future Work

There are several tasks can be extended in the future work. Firstly, the KMLN scheme could be improved by utilizing the clustering method to filter out some of the most likely neighbors, and such selective preprocessing techniques have been proposed in [32, 34]. In addition, it is worthwhile to explore the big variations in historical RSS information caused by irregular RSS patterns to enhance the system estimation accuracy. Furthermore, as the more measurements in RSS sampling the better system performance achieves, we will continue our study on the tradeoff between the positioning accuracy and the training phase workload.

Bibliography

[1] K. Takata, J. Ma, and B. O. Apduhan, “A dangerous location aware system for assisting kids safety care,” in *Proceedings of ACM 20th International Conference on Advanced Information Networking and Applications*, pp. 657–662, Vienna, Austria, April 2006.

[2] P. K. Enge, “The global positioning system: Signals, measurements, and performance,” *International Journal of Wireless Information Networks*, vol. 1, pp. 83–105, April 1994.

[3] S. Tekinay, “Wireless geolocation systems and services,” *IEEE Communications Magazine*, vol. 36, no. 4, pp. 28–29, April 1998.

[4] N. B. Priyantha, A. K. L. Miu, H. Balakrishnan, and S. J. Teller, “The cricket compass for context-aware mobile applications,” in *Proceedings of ACM MOBICOM*, pp. 1–14, Rome, Italy, July 2001.

[5] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, “Multi-camera multi-person tracking for easy living,” in *Proceedings of IEEE International Workshop on Visual Surveillance*, pp. 3–10, Washington, DC, July 2000.

[6] R. Azuma, “Tracking requirements for augmented reality,” *Communications of ACM*, vol. 36, no. 7, pp. 50–51, July 1993.

[7] N. Patwari, I. Hero, A.O., M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, August 2003.

[8] R. Bruno and F. Delmastro, "Design and analysis of a bluetooth-based indoor localization system," in *Proceedings of IEEE International Conference on Personal Wireless Communications*, vol. 27, pp. 711–725, Venive, Italy, September 2003.

[9] S. Sen, R. R. Choudhury, B. Radunovic, and T. Minka, "Precise indoor localization using PHY layer information," in *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*, Cambridge, Massachusetts, July 2011.

[10] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 37, pp. 1067–1080, November 2007.

[11] M. Ciurana, F. Barceló-Arroyo, and F. Izquierdo, "A ranging method with IEEE 802.11 data frames for indoor localization," in *Proceedings of IEEE Wireless Communications and Networking Conference*, pp. 2092–2096, Hong Kong, March 2007.

[12] A. Ladd, K. Bekris, G. Marceau, A. Rudys, D. Wallach, and L. Kavraki, "Using wireless *e*thernet for localization," in *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pp. 402–408, EPFL, Switzerland, September 2002.

[13] A. Narzullaev, Y. Park, and H. Jung, "Accurate signal strength prediction based positioning for indoor WLAN systems," in *Proceedings of IEEE on Position, Location and Navigation Symposium*, pp. 685–688, Monterey, CA, May 2008.

[14] X. Luo, W. J. O'Brien, and C. Julien, "Comparative evaluation of received signal-strength index (RSSI) based indoor localization techniques for construction jobsites," *Advanced Engineering Informatics*, vol. 25, pp. 355–363, April 2011.

[15] Y. Ji, S. Biaz, S. Pandey, and P. Agrawal, "ARIADNE: a dynamic indoor signal map construction and localization system," in *Proceedings of ACM MobiSys*, pp. 151–164, Uppsala, Sweden, June 2006.

[16] W. Yeung and J. Ng, "An enhanced wireless LAN positioning algorithm based on the fingerprint approach," in *Proceedings of IEEE TENCON*, pp. 1–4, Hong Kong, November 2006.

[17] A. Dempster, B. Li, and I. Quader, "Errors in determinstic wireless fingerprinting systems for localisation," in *Proceedings of 3rd International Symposium on Wireless Pervasive Computing*, pp. 111–115, Santorini, Greece, May 2008.

[18] V. Savic, A. Població andn, S. Zazo, and M. Garcí anda, "An experimental study of RSS-based indoor localization using nonparametric belief propagation based on spanning trees," in *Proceedings of Fourth International Conference on SENSORCOMM*, pp. 238–243, Venice/Mestre, Italy, July 2010.

[19] M. Ghaddar, L. Talbi, and T. Denidni, "Human body modelling for prediction of effect of people on indoor propagation channel," *Electronics Letters*, vol. 40, pp. 1592–1594, December 2004.

[20] J. Ryckaert, P. De Doncker, R. Meys, A. de Le Hoye, and S. Donnay, "Channel model for wireless communication around human body," *Electronics Letters*, vol. 40, pp. 543–544, April 2004.

[21] B. Altintas and T. Serif, “Indoor location detection with a RSS-based short term memory technique (KNN-STM),” in *Proceedings of IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 794–798, Lugano, Switzerland, March 2012.

[22] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao, “A wireless LAN-based indoor positioning technology,” *IBM Journal of Research and Development*, vol. 48, pp. 617–626, September 2004.

[23] B. Dawes and K.-W. Chin, “A comparison of deterministic and probabilistic methods for indoor localization,” *Journal of Systems and Software*, vol. 84, no. 3, pp. 442–451, 2011.

[24] P. Bahl and V. Padmanabhan, “RADAR: an in-building RF-based user location and tracking system,” in *Proceedings of IEEE INFOCOM*, pp. 775–784, Tel Aviv, Israel, March 2000.

[25] N. Swangmuang and P. Krishnamurthy, “Location fingerprint analyses toward efficient indoor positioning,” in *Proceedings of Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pp. 100–109, Hong Kong, March 2008.

[26] M. Youssef, A. Agrawala, and A. Udaya Shankar, “WLAN location determination via clustering and probability distributions,” in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pp. 143–150, March 2003.

[27] K. Kaemarungsi and P. Krishnamurthy, “Modeling of indoor positioning systems based on location fingerprinting,” in *Proceedings of IEEE INFOCOM*, pp. 1012–1022, Hong Kong, March 2004.

[28] S.-H. Fang and T.-N. Lin, “Indoor location system based on discriminant-adaptive neural network in IEEE 802.11 environments,” *IEEE Transactions on Neural Networks*, vol. 19, pp. 1973–1978, November 2008.

[29] P. Castro, P. Chiu, T. Kremenek, and R. R. Muntz, “A probabilistic room location service for wireless networked environments,” in *Proceedings of ACM Ubicomp*, pp. 18–34, Atlanta, Georgia, October 2001.

[30] T. Roos., P. Myllymäki., H. Tirri., P. Misikangas., and J. Sievänen., “A probabilistic approach to WLAN user location estimation,” *International Journal of Wireless Information Networks*, vol. 9, pp. 155–164, July 2002.

[31] S. Mazuelas, A. Bahillo, R. Lorenzo, P. Fernandez, F. Lago, E. Garcia, J. Blas, and E. Abril, “Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, pp. 821–831, October 2009.

[32] L. Mengual, O. Marb, and S. Eibe, “Clustering-based location in wireless networks,” *Expert Systems with Applications*, vol. 37, no. 9, pp. 6165–6175, 2010.

[33] M. Youssef and A. Agrawala, “The horus WLAN location determination system,” in *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pp. 205–218, Seattle, Washington, June 2005.

[34] J. Ma, X. Li, X. Tao, and J. Lu, “Cluster filtered KNN: A WLAN-based indoor positioning scheme,” in *Proceedings of International Symposium on World of Wireless, Mobile and Multimedia Networks*, pp. 1–8, June 2008.

[35] K. Curran, E. Furey, T. Lunney, J. Santos, D. Woods, and A. McCaughey, "An evaluation of indoor location determination technologies," *J. Location Based Services*, vol. 5, no. 2, pp. 61–78, 2011.

[36] J. Zheng, C. Wu, H. Chu, and P. Ji, "Localization algorithm based on RSSI and distance geometry constrain for wireless sensor network," in *Proceedings of International Conference on Electrical and Control Engineering (ICECE)*, pp. 2836–2839, June 2010.

[37] A. Roxin, J. Gaber, M. Wack, and A. Nait-Sidi-Moh, "Survey of wireless geolocation techniques," in *Proceedings of IEEE Globecom Workshops*, pp. 1–9, November 2007.

[38] Z. Zhang, X. Zhou, W. Zhang, Y. Zhang, G. Wang, B. Y. Zhao, and H. Zheng, "I am the antenna: accurate outdoor AP location using smartphones," in *Proceedings of ACM MOBICOM*, pp. 109–120, 2011.

[39] M. Robinson and I. Psaromiligkos, "Received signal strength based location estimation of a wireless LAN client," in *Proceedings of IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 2350–2354, New Orleans, LA, March 2005.