LOCATION ESTIMATION IN WIRELESS NETWORKS

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

Yiming Ji

Certificate of Approval:

Kai H. Chang
Professor
Department of Computer Science and Software Engineering

Saâd Biaz, Chair
Assistant Professor
Department of Computer Science and Software Engineering

Min-Te Sun
Assistant Professor
Department of Computer Science and Software Engineering

Stephen L. McFarland
Acting Dean
Graduate School

LOCATION ESTIMATION IN WIRELESS NETWORKS

Yiming Ji

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
May 11, 2006

LOCATION ESTIMATION IN WIRELESS NETWORKS

Yiming Ji

Signature of Author

Date of Graduation

DISSERTATION ABSTRACT

LOCATION ESTIMATION IN WIRELESS NETWORKS

Yiming Ji

Doctor of Philosophy, May 11, 2006
(M.S., Computer Science, Southern Polytechnic State University, 2002)
(M.S., Aerodynamics, Nanjing University of Aeronautics and Astronautics , 1995)
(B.S., Aerodynamics, Nanjing University of Aeronautics and Astronautics , 1992)

169 Typed Pages

Directed by Dr. Saâd Biaz

Wireless location determination has attracted much attention lately due to its many applications in mobile (sensor) networking including target tracking and network intrusion detection. However, it is challenging due to the complexities of the wireless radio propagation characteristics exacerbated by the mobility of the mobile. In this dissertation, we propose realistic localization mechanisms for both indoor and outdoor environments.

For the indoor localization, a common practice is to mechanically generate a table showing the radio signal strength at different known locations in the building. A mobile user's location at an arbitrary point in the building is determined by measuring the signal strength at the location in question and determining the location by referring to the above table using a LMSE (least mean square error) criterion. Obviously, this is a very tedious and time consuming task. This dissertation proposes a novel and automated location determination method called ARIADNE. Using a two dimensional construction floor plan and only a single actual signal strength measurement, ARIADNE dynamically generates an estimated signal strength map comparable to those generated manually by actual measurements. Given the signal measurements for a mobile, a proposed clustering algorithm

searches that signal strength map to determine the current mobile's location. Extensive experiments with various deployment strategies have been carried out to evaluate the ARIADNE system at two different buildings. The results indicate that the ARIADNE system outperforms all other existing indoor localization schemes.

For outdoor wireless sensor networks, this dissertation proposes two reliable and precise distributed localization algorithms: iterative multidimensional scaling (IT- MDS) and simulated annealing multidimensional scaling (SA- MDS). It uses only radio communication constraints to infer node distances, and adapts the multidimensional scaling algorithm (MDS) in the localization research. The research analytically establishes the upper- bound on the estimation error. The proposed techniques can estimate all node positions even with limited and imprecise network knowledge. Analysis and test runs show that the proposed methods are independent of the topology randomness and the range measurement errors. Simulation results for the proposed methods yield an average estimation error of about 25% of radio transmission range.

Besides the detailed research on localization, recently there has been an increasing interest in exploring wireless communications, prototypes and measurements on real test- beds. In order to find a realistic tuning of simulation models, this research investigates one of the critical fundamentals - the realistic radio range irregularity (RRI) model - according to the measurements made under various settings. Using the RRI model, a set of representative localization algorithms are evaluated and compared. Through detailed analysis and extensive simulations, the dissertation points out how the localization performance is affected by the use of simplistic models. The RRI model reflects and highlights the weaknesses of those algorithms and allows the design of countermeasures. This dissertation also introduces a *constrained-greedy forwarding radio propagation method* to remedy the negative effects encountered under actual operating environments.

ACKNOWLEDGMENTS

I have had an very exciting time at Auburn University. The learning and research experience were the best I have ever had. I appreciate all the excellent professors from whom I gained help and instructions through my study.

First, I would like to express my great appreciation to Dr. Kai Chang. He has always been patient and quick to response to my needs with cordial advice. It is under his suggestion and recommendation that I can have the fortune to become a student of Dr. Saâd Biaz. Also, I am grateful for the detailed suggestions he contributed to this work.

I thank Dr. Saâd Biaz for taking me under his tutelage. Working with him, my dream of becoming a professor was reignited. I was always impressed by his instructions and sharp insights to technical issues. He introduced me to the field of wireless engineering and greatly broadened my understanding in this field. I would never forget his inspiration and encouragement that helped me throughout my PhD program. It is a great pleasure to work with Dr. Biaz, and what I have learned from him will definitely guide all my life. I highly appreciate his significant efforts in helping and guiding me!

I thank Dr. Min-Te Sun for his excellent advice on my personal development. He is reliable and always kind to me. I thank him for the great contribution to this dissertation.

I thank Dr. Stanley Reeves from the department of Electrical and Computer Engineering, I thank him for the reading and detailed comments for this work. It is under his tutelage that I entered the area of the *digital image* and *multimedia*. Although I chose the wireless engineering as my major, the knowledge from Dr. Reeves will definitely help me in the future.

Style manual or journal used Journal of Approximation Theory (together with the style known as "auphd"). Bibliograpy follows van Leunen's *A Handbook for Scholars*.

Computer software used The document preparation package TEX (specifically LATEX) together with the departmental style-file `auphd.sty`.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Recently, the demand for wireless communications has grown tremendously. The increasing market for information *anywhere anytime* has been a driving force for the increasing advances in mobile wireless communication. To meet the *anywhere anytime* challenge, many issues remain to be addressed. Location management and mobility management are two critical issues to be analyzed in order to provide seamless and ubiquitous computing environment for mobile users. Consequently, many localization systems have recently been proposed (Bulusu, Heidemann and Estrin, 2000 [1]; Bahl and Padmanabhan, 2000 [2]; Chen and Kobayashi, 2002 [3]; Chincholle, Eriksson and Burden, 2002 [4]; Haeberlen et al., 2004 [5]; Niculescu and Nath, 2004 [6]).

To determine the position of a user, three basic methods are available:

1. Range based lateration method, which requires at least three distinct estimates of the distance from the user to known fixed locations (Savarese, 2002 [7]; Savvides, Park and. Srivastava, 2004 [8]; Ji and Zha, 2004 [9]; Hu and Evans, 2004 [10]);

2. Direction based angulation method that involves the direction or angle of arrival of at least two distinct signals from known locations (Niculescu and Nath, 2004 [11]; Niculescu and Nath, 2004 [6]);

3. Location fingerprinting method where location dependent signal characteristics (i.e. fingerprints like signal strength, image, sound, or other unique information) are pre-collected and stored in a database table. To locate a user, the current fingerprint of the mobile is measured

and used to search the database to find a matching position (Hills, Schlegel and Jenkins., 2004 [12]; Krishnan et al. 2004 [13]; Hatami and Pahlavan; 2004 [14]).

The range based *lateration* and direction based *angulation* methods are effective for line-of-sight radio signals, and thus they are mainly used for outdoor localization systems. Due to the multipath propagation properties of the indoor environment, the method of location fingerprinting is generally used in indoor location systems.

This dissertation focuses on the design, analysis, and evaluation of both indoor and outdoor localization systems. The objective is to develop precise and automatic localization systems that do not depend on special hardware support and incur minimum deployment and maintenance cost.

For both indoor and outdoor environments, this dissertation exploits the wireless technology between communication nodes. For example, the received signal strength indicator (RSSI) from off-the-shelf wireless cards is exploited for the localization implementation. In this dissertation, RSSI is referred as signal strength. This dissertation considers the popular widely deployed 802.11 systems operating in the ISM band (2.4 GHz). However, the developed methodologies may well be implemented over any RF (radio frequency) technology.

This next two sections respectively introduce the research for the indoor and outdoor localization systems. Section 1.1 presents the research for the indoor localization system, and section 1.2 describes the outdoor system.

1.1 Indoor Localization Systems Research

Due to the multipath radio propagation characteristics, the indoor localization research remain to this day challenging tasks for the research community. If the radio propagation signal strength is

tightly correlated with the distance between the transmitter and the receiver, then the location determination is simple in that only three quadratic equations need to be solved for two unknowns (2-D environment). Unfortunately, the relationship between the radio signal strength and the distance is not straightforward for the indoor radio propagation. Additionally, due to the dynamic property of the indoor radio propagation, the signal strength at the same location varies over time. These difficulties make a solution of indoor localization quite complex and elusive.

Therefore, the location fingerprinting method is generally more applicable for the indoor environment. In this method, the reference signal strength of a mobile user is collected, in advance, at many locations throughout the site. A database table comprising of the location information along with the reference signal strengths associated with that location is then built. In this dissertation, we call the database table as Signal Strength Map table (SS-MAP). If there are total n sniffers, a typical record in the signal strength map table is in the form of: $< location_{ID}, SS_1, \ldots, SS_n >$, where $Location_{ID}$ is the tag associated with the location or coordinates on the floor plan, SS_j $(1 \leq j \leq n)$ is the signal strength sensed by j^{th} sniffer when the user is present at the location corresponding to $Location_{ID}$. To locate the position of a mobile user, sniffers measure the signal strength of the received packet and compare with the records in signal strength map table. The position corresponding to the best match is used to represent the position of the mobile user [2, 3].

For the location fingerprinting method used in the indoor localization research, previous works either rely on specialized hardware, or require extensive manually measurements. For example, Krishnan *et al* [13] presents an adaptable infrastructure-based system that uses the *'stationary emitters'* to periodical collection of signal strength at specific locations on the site. And Haeberlen *et al* [5] took 28 man-hours to build a signal strength map that covers an area of 12,000 sq. meters. The hardware approach usually arouses high cost associated with devices, installation, and maintenance.

And the brute force measurement does not take into account the time-varying nature of the indoor radio signal strength and would necessitate remeasurement to keep track of changes due to variation in propagation environment like human occupancy, furniture and partition updates. These existing methods are not practical for general large-scale applications.

Therefore, the goal for this dissertation is to develop an automated and dynamic indoor localization method. The research is interested in a general indoor localization environment where no special hardware is required thus incurring lower deployment and maintenance cost.

In the dissertation, the research considers the localization of a mobile user on the same floor where the properties of the indoor partitions are not significantly different; or in other words, the construction materials and the thickness of all walls are similar. This research focuses on the network-based localization system where the location of the user is determined by the network. In this system, the user's signal strength must be sensed by three or more sniffers which are deployed at known positions.

Much research has been done in this area and many good solutions have been proposed for the indoor localization system. However, four common problems remain to be solved for most indoor systems:

1. None of existing systems consider the **dynamic property** of the indoor radio propagation;
2. No systems **systematically evaluate** the whole indoor system;
3. All systems require **extensive measurement** to either manually construct signal strength map or to estimate the site specific attenuation parameters;
4. If **special hardware** is to be used, detailed deployment and maintenance are required.

Consequently, most existing indoor systems fail to address the cost and complexity of the system deployment and maintenance, and they are unsuitable for large-scale deployment. Therefore, an automated and scalable indoor localization system is highly desirable.

1.2 Outdoor Localization Systems Research

Recently, location-based services have gained considerable attention from both academia and industry. Current reported applications and services include coverage analysis [15, 16], location-aware applications [17, 18, 4, 19, 20, 21], environmental monitoring [22, 23], target tracking [24, 25], battlefield surveillance [26], and intrusion detection [27]. In addition, location management and efficient routing protocols also require location knowledge (Perlman and Haas, 1999 [28]; Ko and Vaidya, 2000 [29]).

The mobile's location at outdoor environment can easily be obtained by GPS enabled systems with a good precision. However, GPS system is inappropriate for large scale deployment of ad hoc sensor networks because of constraints in volume, power consumption and cost.

An alternative to GPS is the use of distributed localization algorithms that enable nodes to determine their relative positions. If absolute positions are necessary, a limited number of nodes (called **anchors**) must have the capability of determining their absolute positions. Anchors can get their absolute position using manual configuration or techniques such as GPS. Much localization research has been carried out recently using this approach [30, 8, 9, 31, 10].

Based on the reliance on the hardware support, localization algorithms can be classified into two main categories: **range-based** algorithms and **range-free** algorithms. Range-based algorithms rely more on hardware support by applying either one or a combination of TOA, TDOA, AOA,

or RSSI technologies. On the contrary, range-free algorithms require less or no hardware support [31, 32, 33, 34, 35].

Absolute positions are necessary for many applications. But some applications require only the knowledge of nodes' relative position (without actual physical or absolute position information). This means that, at least for some applications, the position information can be relative instead of absolute [36, 37]. Consequently, the localization algorithms can also be categorized as absolute and relative.

For the traditional outdoor localization research, the research community usually exploits simple radio propagation models in localization simulations to investigate and evaluate the localization algorithms and protocols. This simple model assumes *spherical* radio transmission pattern of fixed range at all wireless nodes [11, 16, 38]. However, results on the field widely differ from simulation results. The research community become aware that simplistic models may lead flawed protocols. A key reason is that radio propagation is irregular [39, 33]. Pawlikovski *et al.* [40] surveyed 2200 published network simulation results and identified that the simple circular model is popularly adopted by most researchers and simulators. From experiments, Pawlikowski showed that the ideal radio model yields results radically different from reality. Kotz *et al.* [41] reviewed six of the most used simplistic assumptions and showed their individual impact on the obtained results. Thus, the credibility of applications/protocols based simplistic model is dubious.

Thus, most existing localization algorithms contain four major problems:

1. Most existing algorithms work only under **specific and restrictive network conditions**. When such conditions are not met, the confidence in location estimates is not satisfactory and makes key decisions hard and risky;

2. Most algorithms require **excessive network communications** and complex refinement procedures with multiple computing rounds that are energy inefficient for small devices in ad hoc networks;

3. Precision evaluation of these algorithms is conducted using **different metrics and criteria** making uneasy the comparison of algorithms.

4. **Bi-directional communication links** are commonly used by most algorithms, where perfect circular radio pattern (in a 2D environment) and fixed radio range are assumed at all nodes. For a general sensor network, radio transmission range at different nodes is not fixed, and the radio shape is irregular in nature. Consequently, a practical radio model is highly needed, and the impact of the irregular radio pattern on the localization performance has to be analyzed.

The objective of the outdoor localization research in this dissertation is to develop practical localization algorithms that work on the most general environment, and the radio transmission shape could be non-circular. The interest of the research is on wireless sensor networks, where nodes are randomly deployed at interested environment,

The research assumes that:

1. Some nodes, called anchor nodes, know their physical positions in advance by special deployment or by the use of GPS. The others are called normal nodes.

2. Each node can measure, through the received signal strength indicator (RSSI), the approximate range to its immediate neighbors.

3. Certain routing optimization mechanisms are employed such that each node only maintain the least hop counts to a sender.

Consequently, the goal of the outdoor localization research in this dissertation is to: 1) develop a practical localization algorithm that works on the most general environments; and 2) construct a real test-bed for the wireless research where the radio communication is based on a realistic non-circular radio pattern model in which the radio range is adjustable according to different operating environment.

The rest of the dissertation is organized as follows: Chapter 2 presents a literature review for both indoor and outdoor localization systems. Chapter 3 introduces proposed methodology for the indoor localization systems, and Chapter 4 describes the detailed research for the outdoor localization systems. Chapter 5 discusses applied results for the proposed indoor localization system. Chapter 6 gives the applied results and comparison outcomes for the outdoor localization algorithms, and Chapter 7 concludes the dissertation.

CHAPTER 2

LITERATURE REVIEW

This chapter surveys the related research for both the indoor and outdoor localization systems. In Section 2.1, we first introduce various range measurement techniques. Sections 2.2 and 2.3 are respectively dedicated to the introduction of indoor and outdoor localization systems.

2.1 Distance Measurement Technologies

This section reviews existing distance measurement technologies. Based on the applied hardware and the operating environment, the distance measurement technologies can be categorized into the following two main categories:

1. Satellite positioning technologies, which include Global Position System (GPS), Differential GPS (DGPS), Assisted GPS (A-GPS), Wide Area Augmentation System (WAAS), and Galileo system (European version of GPS);

2. Network-based technologies, which include Time of arrival (TOA), Time difference of arrival (TDOA), Angle of arrival (AOA), Enhanced observed time difference (E-OTD), and Received signal strength indicator (RSSI).

The following sections briefly review these measurement techniques.

2.1.1 Satellite Positioning Technologies

- **Global Positioning System (GPS)**, GPS consists of three components: satellites, control stations, and GPS receivers [42].

The GPS system consists of 24 orbiting operational satellites, which can transmit very low power radio signals that allow any GPS receiver to determine its position on the Earth. Among the 24 satellites, 21 of them are active at any time, and the rest three are backups.

GPS control system monitors the satellites and providing them with correct orbital and clock data. There are five control stations around the world.

In order to determine the position, the receiver needs to know at least four satellites' location and the distance to them. The approximate location of the satellites is obtained from the satellites directly. The approximation can be adjusted by using data from control stations.

The distance from the receiver to the satellite is calculated as the product of radio speed and the radio travel time between them. The travel time is estimated from the difference of the 'pseudo-random' code generated at the same time by the satellite and the receiver. With the information of four or more satellites, the position of the receiver can be determined.

There are two levels of service provided by GPS. The more-accurate one is Precise Positioning Service (PPS), which is only available to authorized users and is intended for military usage. The less-precise service is Standard Positioning Service (SPS), which is available for all civil users worldwide without charge or restrictions. According to [43], the position estimation error for PPS is between 5-10m. The reported SPS horizontal accuracy by U.S. Department of Defense is of $\pm 100m$ at 99.5% confidence level and $\pm 300m$ at 99.9%. In [44], an error of $\pm 185m$ in latitude and $\pm 216m$ in longitude was reported.

- **Differential GPS (DGPS)**, **Assisted GPS (A-GPS)** and **Wide area augmentation system (WAAS)**

DGPS is an extension of the GPS system intended to improve the accuracy of the GPS system. DGPS uses stations to broadcast position correction beacons. With these correction messages, GPS receivers can correlate them with received satellite signals. This technology effectively reduces the effect of selective availability (SA) and propagation delay. It is reported that DGPS provides an estimation with 2 to 10 meters accuracy [45, 46].

A-GPS is a technique that improves the functionality and performance of GPS using only GPS satellite signals. It works by integrating the classic GPS information with sophisticated geographic software and mobile/cellular network information. A-GPS is currently widely used in many systems, and the accuracy in system *U-Map* is within 5 meters [46].

WAAS is similar to DGPS and it is initially intended to be used for precision flight control by the Federal Aviation Administration (FAA) and the Department of Transportation. WAAS system consists of a network of approximately 25 ground reference stations in the USA. And it provides a very large service area where the DGPS source may not be available. The reported accuracy is approximately seven meters.

- **Galileo system**

  Galileo is the European version of satellite navigation system, it is a system that both competes and complements with the US GPS system [47].

  The fully deployed Galileo system will consist of 30 satellites, with 27 of them operational and the other three as active reserves. With dual frequencies, Galileo delivers real-time positioning accuracy of meter range.

2.1.2 Network-based Technologies

- **Time of Arrival (TOA)**: This technology works by measuring the arrival time of a known signal sent from a (mobile) node received at three or more measurement units. Synchronization of the measurement units is essential. Therefore, this method requires additional measurement hardware unit in the network at the geographical vicinity of the (mobile) node so as to accurately measure the TOA of the signal bursts. It is reported that the accuracy of TOA is about $100 \sim 200m$ [48].

- **Time Difference of Arrival (TDOA)**: This technology works similar as TOA in that both technologies utilize signal propagation time. In TDOA, however, two types of signals are selected so that their propagation speeds are significantly different. When a transmitter sends two types of signals simultaneously, the receiver can easily detect the difference in the time of arrival between the two types of signals. The time difference can then be used to compute the distance between the communication pairs. To deal with multipath problems, TDOA systems has been historically based on wide-band radio technology. The known accuracy is about $100 \sim 200m$.

- **Enhanced Observed Time Difference (E-OTD)**:In this method, it is the (mobile) node that performs the time measurement of beacon signals from nearby base stations. This method does not require synchronization in the network. The reported accuracy for E-OTD is $50 \sim 200m$.

- **Angle of Arrival (AOA)**: This method requires special antenna arrays at the base station to determine the angle of the arriving signal from the mobile node. The angles from two or more

base stations can determine the position of the node by the intersection of the signal arrival directions. The accuracy is $100 \sim 200m$.

- **Received Signal Strength Indicator (RSSI)**, this method measures the strength of the received signal in order to deduce the possible range the signal has propagated from the sender to the receiver. It is applicable if the transmission power is constant or known in advance. RSSI is a low accuracy localization method. However, it is simple and doesn't require extra equipment. Consequently, it is widely supported by most current wireless devices.

2.2 Literature Review for Indoor Localization Systems

Due to multipath propagation in indoor environments, indoor localization systems generally require specialized hardware to be deployed. Depending on hardware technologies, indoor localization systems may be broadly classified as non-RF and RF based systems.

Non-RF based location systems were designed using a specific technology independent from data communication networks. Such location systems exploit **1)** infrared (IR) (*Active Badge* , Want *et al.* (1992), [49]); **2)** ultrasound, Ward, Jones, and Hopper, (1997) [50, 51, 52]; **3)** magnetic field [53]; or **4)** light (cameras), krumm *et al.*, (2000) [54]. Such early location systems require specialized hardware used only for the location determination and incur in general a high deployment and maintenance cost.

In the recent years, the popular success and widespread deployment of RF 802.11 wireless networks enticed many researchers to explore existing RF 802.11 wireless network infrastructure to build location systems. Consequently, **RF** based indoor positioning systems generally exploit *RF signal strength* sensed at reference sniffers or base stations.

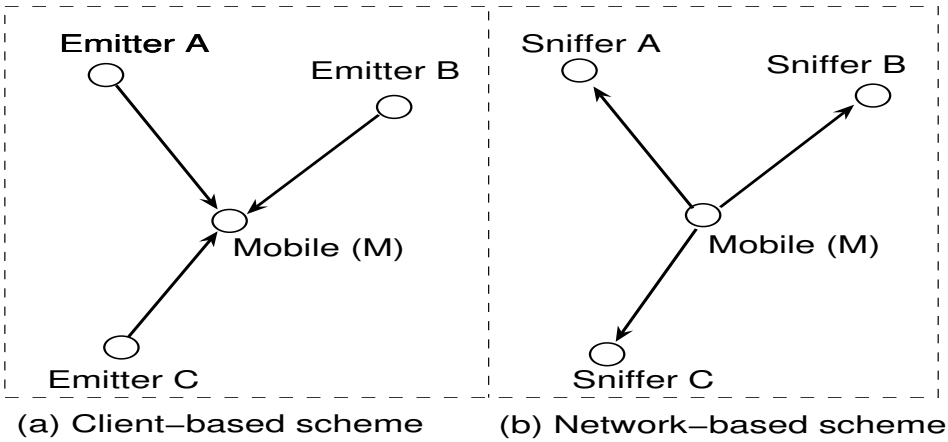

Figure 2.1: Signal Strength Measurement

Oversimplifying, if the signal strength of the radio propagation is tightly correlated with the distance between emitter and receiver, then location determination would be a trivial problem that could be solved by one of the following two approaches as illustrated in Figure 2.1.

Figure 2.1(a) illustrates a *client-based scheme* where three emitters A, B, and C are at known positions. A mobile (client) would listen successively to the three emitters and would measure the signal strength from them. If the measured signal strength yields the distance from each emitter to the mobile, the location of the mobile reduces to the solution of a simple system of three quadratic equations with two unknowns (assuming the mobile moves in a plan). Note that in this scheme, the client has an active part in the location process: it measures the signals and infers its location.

A dual approach, the *network-based scheme*, is illustrated in Figure 2.1(b): three sniffers at known positions listen to the mobile and measure the signal strength of received packets. It suffices to collect the signal strength measurements (over the network!) from the three sniffers to determine the location using basic calculations. Unfortunately, the relationship between signal strength and distance is not straightforward and is dynamic in nature: even if a mobile does not move, the sniffers

in Figure 2.1(b) will measure the signal strength that varies over time. Moreover, two mobiles that are quite close may generate signals of significantly different strength at the same sniffer. These difficulties make a solution to location determination quite elusive.

In order to address this problem, researchers proposed a *two-step solution*: **First**, establish a signal strength map *SS-MAP* where the signal strength at known and predetermined locations is either manually measured or theoretically estimated, and **Second**, measure signal strength for a mobile at a given location and *SEARCH* the signal strength map *SS-MAP* for the "closest" location that with the best signal strength measurement match.

The RADAR [2] system proposed by Bahl and Padmanabhan is exemplary of such an approach: the authors adopt a client-based scheme to collect the radio signal strength received from three base stations at a mobile (method in Figure 2.1(a)) at a selected set of predetermined and known locations. The scheme records the collected signal strength as a function of location, and such records constitute what we call a signal strength map *SS-MAP*. This *measured* signal strength map was used by the authors in two different strategies: (1) the first strategy (they dubbed "empirical method") consists of the mobile sensing the signal strength from the three base stations and *searching* for a record in the measured *SS-MAP* for the best signal strength measurements match; and (2) the second strategy consists of using a simple propagation model to construct an *estimated SS-MAP* that is validated using the *measured SS-MAP*.

Estimating is more convenient than measuring a *SS-MAP* especially for a large building. The *estimated SS-MAP* is used the same way as the measured *SS-MAP* in strategy 1. Unfortunately, the authors [2] report that the first strategy (i.e., the "empirical method") outperformed the second strategy that uses the estimated *SS-MAP*. The key weakness of the second strategy is that the radio propagation model results in an estimated *SS-MAP* does not fit well the measured one.

This dissertation proposes a convenient and scalable network-based location system, and defines two modules corresponding to the above *two-step solution*. The first *MAP GENERATION* module estimates a signal strength map SS-MAP, and the second *SEARCH* module determines the location of the mobile when given the input the current measured signal strength of the mobile. We introduce related research for the two modules in Section 2.2.1 and Section 2.2.2, respectively; and Section 2.2.3 reviews the similar indoor localization systems. In Section 2.2.4, we separately introduce the sniffers deployment in previous research.

2.2.1 Map generation module

A signal strength map (SS-MAP) usually contains a dense grid of locations together with the measured signal strength at those positions. Considering a total of n sniffers in a network-based indoor localization system, a typical record in the signal strength map table is in the form of: $< location_{ID},\ SS_1,\ \ldots,\ SS_n >$, where $location_{ID}$ is the tag associated with the location or coordinates on the floor plan, and $SS_j,\ (1 \leq j \leq n)$ is the signal strength sensed by j^{th} sniffer when the user is present at the location corresponding to j.

In this research, three sniffers[1] (A, B and C) are used to monitor the signal strength of the mobile user, and the signal strength measurements made by the three sniffers is called *one signal strength measurement triplet*. To locate the position of a mobile user, sniffers measure the signal strength of the received packet and compare the signal strength measurement triplet with the records in SS-MAP. The position corresponding to the best match is used to represent the position of the mobile user (Chen and kobayashi, 2002 [3]; Bahl and Padmanabhan, 2000 [2]).

[1] In Chapter 3, we also consider four or five sniffers in order to derive an optimal sniffers deployment strategy.

To build the signal strength map SS-MAP, most indoor localization systems take the *manual measurement* method, which is carried out by manually measurements of signal strength at short intervals within the building (*empirical method*). The manual measurement is time-consuming. Moreover, the manually measurements cannot take into account the time-varying nature of the signal strength, and would necessitate remeasurement to keep track of changes due to the variation in propagation environment (ex. Human occupancy, furniture and partition updates).

To address the dynamic property of the radio propagation, Krishnan *et al.* (2004) [13] presents an adaptable infrastructure-based system. The system suggests the deployment of 'stationary emitters' to collect signal strength at specific locations on the site, and thus it is called *'stationary emitters'* method. Apart from the cost associated with the purchase of such devices, installation issues such as placement and power supply need to be considered for large scale deployment.

Therefore, in order to construct a signal strength map SS-MAP, the *empirical method* and the *'stationary emitters' method* fail to address the cost and complexity of the system deployment and maintenance. Hence, many research tries to find a indoor radio propagation model that estimates the signal strength at reference locations. The next section will separately survey related work for the radio propagation model.

Indoor Radio Propagation Model

Research on indoor radio propagation is an active field. A study of indoor radio propagation characteristics can be found in [55, 56]. A detailed description of earlier radio propagation models can be found in [57]. Based on the ray tracing technique, several statistical models have been analyzed recently [58, 59, 12]. When considering the *large-scale attenuation model*, most researchers

model the radio propagation path loss as a function of the *attenuation exponent* n (Please, see Equation 2.1), which is two for free space but greater than two for an indoor environment.

$$P(d)[dB] = P(d_0)[dB] - 10 \times n \times log_{10}(\frac{d}{d_0}) \tag{2.1}$$

where $P(d)$ is the power at distance d to the transmitter in meters; $P(d_0)$ is the power at a reference distance d_0, usually set to 1.0 meter. n is the attenuation exponent, which is often statistically determined to provide a best fit with measurement readings.

Based on the considered parameters in the radio propagation model, all radio propagation models can grossly be grouped into three categories: (1) Simple attenuation model; (2) Partition model; and (3) Site-specific model.

Simple attenuation model is in the form of Equation 2.1, and it is the base model for the other models. Hills, Schelegel, and Jenkins [12] used this model as part of an automated design tool to estimate the coverage areas for a set of APs. With point-by-point measurement, Hills *et al.* report that an attenuation exponent of 2.60 yields the best fit in the buildings on the Carnegie Mellon University campus. A difference of 3.0dB between the measurements and estimates is achieved in most cases.

Different from the simple attenuation model, the *partition model* reduces the pass loss effect from the attenuation exponent by additional consideration of the attenuation effects from the indoor partitions, like walls and floors. Many successful models belong to this group. A couple of famous examples include Phaiboon's statistical model [59], and *wall attenuation factor* model in RADAR [2]. Phaiboon's model considers multiple floor environment. The test results show the estimated signal strength from the partition model agrees better than that of simple attenuation

model [59]. In contrast, the RADAR system considers attenuation effects from walls along the direct path between the transmitter and the receiver on the same floor. RADAR's location search in the estimated signal strength map *SS-MAP*yields an average resolution of about 4.3m [2].

Site-specific model is similar as the partition model except that it relates to path loss with site-specific parameters (geometrics, materials, and thickness). Two representative models include Hassan-Ali and Pahlavan's probability model [58], and Lott and Forkel's multi-wall-and-floor model [60]. Hassan-Ali *et al.* compared the estimated signal strength with measurements using a probability model. The results of mean error of 2.77dB and standard deviation of 2.87dB are obtained [58]. Compared with the other models, the site-specific model does not depend on special assumption, so it works on most general building environment. However, it is complex and requires detailed site-specific parameters.

All these radio propagation models have the following shortcomings:

1. Tedious and extensive on-site reference measurements are required in order to determine the building-specific attenuation exponent and the attenuation coefficients of indoor partitions;
2. The measurements do not consider the dynamic behavior of the indoor radio propagation;
3. They only consider the path loss along the direct path between the transmitter and the receiver;
4. Detailed material characteristics and geometry properties are required if site-specific model is to be used.

Hence, in order to construct a signal strength map SS-MAP, most existing RF-based indoor localization systems are not able to cope with the dynamic property of the indoor radio propagation. They are unpractical for general large-scale applications. Therefore, to build a practical indoor

localization system, it is of particular interest to develop accurate and time-efficient methods that dynamically construct site-specific signal strength map.

2.2.2 Search module

As pointed out earlier, for localization of a wireless user, the signal strength map SS-MAP is searched based on the signal strength measurements at sniffers in the 'on-line' phase. To search the SS-MAP table, a general comparison metric is the least mean square error (LMSE).

$$D = min_{k=1}^{N}\{\frac{1}{n}(\sum_{i=1}^{n}(ss_{m,i} - ss_i)^2)^{\frac{1}{2}}\} \tag{2.2}$$

where D is the least mean square error, N is the total number of records in the signal strength map table, k denotes the k^{th} record in the *SS-MAP* table; n is the number of sniffers. $ss_{m,i}$ denotes measured signal strength at sniffer i of the mobile user, and ss_i is the signal strength record at a sniffer i in *SS-MAP* table. The *nearest neighbors in signal space* method by Bahl and Padmanabhan (2000) [61] is essentially this approach.

A problem with *LMSE* is that two or more very different locations could potentially have same signal strength, thus additional processing must be carried out in order to select a more accurate estimate. Therefore, more advanced localization methods are highly desirable.

Prasithsangaree, Krishnamurthy, and Chrysanthis (2002) [62] proposed a *closeness elimination scheme*. The main purpose is to find more than three locations from the *SS-MAP* table with signal strength close to the measurement. From these, the three closest positions are selected and their position average is used to denote the estimated location for the mobile user. Similarly, in [63], Pandey (2004) et al. used the second lowest MSE to assist the estimation. They found that if the

LMSE and the second lowest mean square error are physically adjacent, then the middle of their locations yields better estimates.

Hatami and Pahlavan (2004) [14] also proposed a modified *LMSE* algorithm, dubbed *prioritized maximum power*. This method sorts measured signal strength in descending order for all sniffers so that a contribution priority of each sniffer in the mapping procedure is obtained. According to the priority, the estimates are restricted to a set of reference points. Then *LMSE* or *closeness elimination scheme* is used to determine the final estimates.

Existing search algorithms works on accurate and dense grid signal strength map; and the search usually find an exact hit (*LMSE*) or a small set of positions close to each others (*closeness elimination scheme*). In reality, the signal strength map is imperfect. Measurements take time and the radio propagation is subject to change with environment, therefore, a search on the imprecise SS-MAP generally returns multiple hits or a set of scattered point clusters. Consequently, an efficient search algorithm that works with imperfect signal strength map is highly desirable.

2.2.3 Similar Indoor Localization Systems

The RADAR system [2, 61], the closest to ARIADNE, proposes an indoor radio propagation model for localization and tracking. Although the system requires extensive measurements and calibration, the achieved localization performance is not satisfactory. The RADAR radio propagation model does not fully capture the multipath phenomenon as it only considers radio propagation along the direct transmission path.

Similarly, Hatami *et al.* [14] used ray tracing software to generate a reference signal strength map *SS-MAP*. The proposed system uses five APs deployed in a building of 65×48 meter. To locate a mobile user, two different localization method (*LMSE* and *prioritized maximum power*)

are evaluated and compared. The results show that the *LMSE* method provides better estimation performance for users within the building. However, *prioritized maximum power* is less susceptible to reference grid resolution and can achieve better estimates when mobile user resides within the vicinity area outside of the building. The results from Hatami (Figure 3 in [14]) show a relation of 10 meters complementary cumulative positioning error with 54% probability for *prioritized maximum power* method. The research by Hatami [14] mainly focused on localization algorithms targeted at intruder detection. It uses ray tracing software to construct signal strength map *SS-MAP* without introducing the indoor radio propagation model.

Existing indoor localization systems failed to systematically consider the two modules, and thus when constructing signal strength map table, they either require tedious measurements or have to rely on special hardware. If the indoor radio propagation model is to be considered, they only consider the attenuation along the direct transmission path. Therefore, the research on the radio signal strength module is still a difficult topic. On the other hand, the search module in existing systems assumes that the precise and high resolution signal strength map table is available. And the search is based on the least mean square (LMS) algorithms.

Different from these systems, this dissertation introduces ARIADNE, a new indoor localization system. It holistically analyzes the indoor localization system by systematically considering two modules: map generation module and search module. In the map generation module, a new indoor radio propagation model is proposed; and the search module presents a clustering-based localization algorithm that works on imprecise radio propagation map tables. The ARIANDE system has been evaluated in two buildings under various sniffers deployment strategies.

Table 2.1: Number of sniffers used in previous research

Researchers and Systems	Number of Sniffers	Covered area ($m \times m$)
Bahl and Padmanabhan RADAR(basic) [2]	3	43.50×22.50
Bahl and Padmanabhan RADAR (enhanced) [61]	5	42.96×21.84
Hatami and Pahlavan Intruder detection [14]	5	65×48
Agiwal et al. LOCATOR [64]	3	45.11×32
Krishnan et al. LEASE [7] [13]	5	68.58×43.90
	6	76.20×53.34
Ladd et al. Localizer [65]	14	$\sim 65 \times 35$
Haeberlen et al. Practical localization [5]	33	12,000

2.2.4 Review of Sniffers Deployment Method

During the research, it is found that the sniffers deployment is critical for the performance of the indoor localization. Since the relation of the signal strength and the distance between the transmitter-receiver pair is not straightforward for the indoor environment, the sniffers deployment must first be designed to provide maximum signal coverage for the site; secondly, the deployment must also present best *discrimination* of signal strength for different locations inside the building.

In addition to the consideration of the position configuration, the available of larger number of sniffers may also improve the precision of the location estimation. Depending on the construction material and the partition, the effective 802.11 (b) radio range for the indoor transmission is roughly about $40 \sim 50$ meters. Table 2.1 surveys the number of sniffers used in the experimental analysis in related research, and it is shows that the test experiments generally use more sniffers to provide redundant signal coverage for the experimental site. A large number of deployed sniffers masks or smooths the requirement for the strategy of the sniffers deployment. However, a large number of sniffers cause more interference and impose extra cost on hardware and maintenance. Therefore, it is desirable to deploy just enough sniffers such that a good deployment strategy of minimal sniffers still provides optimal localization performance.

For all experiments, **none** of the existing research addresses the intrigue of the sniffers deployment. However, most researchers deploy the sniffer around the perimeter along the test building. And an interesting observable fact is that NONE of the above systems deploys the sniffers along a straight line. It appears intuitively that the straight-line deployment will not provide optimal localization results. Therefore it is of particular importance to derive a theoretical foundation of the sniffers deployment strategy such that optimal localization is achievable with known precision.

We review in the next section outdoor localization systems.

2.3 Literature Review for Outdoor Localization Systems

Research is currently quite active to provide location information for wireless ad hoc sensor networks. An ad hoc sensor network is a self-organized and rapidly deployable network. It usually consists of a large number of unattended sensor nodes that autonomously construct a network and communicate with each other over multi-hop paths. The unattended nature of the system necessitates a set of mechanisms to assure simple deployment while maintaining the self-organized nature. In sensor nodes, the *location information* of a node is often more important than its *identity*.

In the following sections, we first introduce, in Section 2.3.1, various localization algorithms, and in 2.3.2, we introduce the research on the realistic radio range irregularity model that greatly affects the localization performance.

2.3.1 Outdoor Localization Algorithms

To initialize the positioning process, a common approach is to have the interested node, who wants to find out it's position or to find an optimal routing path to the destination node, to first broadcast the hello packets to their neighbors. Through communication, neighboring nodes measure the

distance between themselves. Based on the distance measurements and the proposed localization algorithms, all nodes can estimate their relative positions with each other through distributed localization algorithms. In the presence of anchor nodes with known locations, absolute positions of all nodes can be derived accordingly.

Most of existing algorithms take advantage of mature micro-sensor technology. As introduced in Section 2.1 at the beginning of this chapter, time of arrival (TOA) measures the travelling time of radio signals, and time difference of arrival (TDOA) measures the time difference of the radio signals arrived at various antennas. The two methods are the most common methods for range estimation, and they have been applied in many infrastructure based systems (Oshman and Davidson, 1999 [66]; Priyantha, Chakraborty and Balakrishnan, 2000 [52]) and infrastructure free sensor networks (Niculescu and Nath, 2001 [67]; Savvides, Han and Strivastava, 2001 [68]; Savvides, Park, and Srivastava, 2002 [69]). Additionally, angle of arrival (AOA) has been proposed to estimate relative angles between neighbors (Niculescu and Nath, 2003 [11]). A good survey of these distance measurement systems can be found in Hightower and Boriello (2001) [70] and Tauber (2002) [71]. Still, another method of measuring the range is through received signal strength indicator (RSSI). RSSI does not require extra equipment and is widely supported by current transceivers. Consequently, RSSI is commonly adopted by many localization systems (Premaratne, Zhang and Doguel, 2004 [31]; Ji and Zha, 2004 [9]; Bahl and Padmanabhan, 2000 [2]; Sichitiu, Ramadurai and Peddabchagari, 2003 [32]).

Based on the reliance on the hardware support, localization algorithms can be classified into two main categories: range-based algorithms and range-free algorithms. Range-based algorithms rely more on hardware support by applying either one or a combination of TOA, TDOA, AOA, or RSSI technologies. A set of representative range-based algorithms include:

1. Convex positioning by Doherty *et al.* [72];

2. Ad-hoc positioning system by Niculescu *et al.* [67, 11, 36];

3. N-hop multilateration or AHLoS system by Savvides *et al.* [73, 30, 68, 69, 8];

4. Robust Positioning by Savarese [73, 30];

5. Probabilistic approach by Ramadurai *et al.* [38, 32];

6. Multidimensional Scaling by Ji *et al.* [9, 74].

On the contrary, range-free algorithms require less or no hardware support at all (Premaratne, Zhang and Doguel, 2004 [31]; He, Stankovic and Abdelzaher, 2003 [75]; Sichitiu, Ramadurai and Peddabchagari, 2003 [32]; Nagpal, Shrobe and Bachrach, 2003 [34]). Three exemplary algorithms include Nagpal *et al.*'s *Amorphous localization algorithm* [34, 35], Premaratne's *GGO* algorithm, and He *et al.*'s *APIT* algorithm.

The *Amorphous localization algorithm* by Nagpal *et al.* [34, 35] uses *Kleinrock's formula* [76] to estimate the average hop distance in a uniformly deployed network. However, *Kleirock's formula* can only be used in densely deployed networks. For more general networks, special transmitters may be used to assist the localization procedure. For example, Premaratne *et al.* in [31] assumes that anchors can transmit radio signals with multiple level transmission powers. And in [33], He *et al.* uses high-powered transmitters. These two transmission methods will let anchors cover much larger areas. Consequently, with these special transmitters, an anchor can send beacon packets directly to normal nodes that are far away from the anchor itself. When a normal node obtains enough information about the surrounding anchor nodes, Premaratne's method allows the normal node to estimate its position by grid overlaying all possible regions it may reside in. This method is called *geometric grid overlaying (GGO)* in [77]. In contrast, He's method isolates the area into

multiple triangular regions and use *approximate a-point-in-triangulation Test (APIT)* to obtain the position estimates.

This dissertation mainly focuses on the range-based algorithms. And in the following sections, we first introduce the radio propagation methods, and we describe a set of existing localization algorithms next.

Radio Propagation Mechanisms

Although some algorithms, like [72], do not explicitly describe the initialization communication method, most of them allow the anchor nodes to broadcast their position information first. When intermediate neighboring nodes receive the beacon message, they record the anchor's position information, and then rebroadcast packets to their neighbors with updated hop counts. The process continues until all anchor nodes' position information is delivered to every node in the network.

To partially alleviate the expensive flooding, several optimization techniques are applied. Some of them are as follows:

- *Flood limit* method provided in [78]: This method assumes the normal nodes can derive good position estimation from a limited number of anchors. Consequently, when a normal node has recorded enough anchors, it will stop forwarding further information;

- *Hop-count* method by [30], in this method, if a node receives multiple packets from the same anchor node, it maintains and rebroadcasts the one with the least hop counts while ignoring all the rest. This optimization will eventually lead to a shortest path to the anchors.

- *Time stamp* method by [67], in which a Time To Live (TTL) stamp is appended to each beacon packet. Out-dated packets are dropped silently.

To estimate the distance to anchor nodes, several different methods have been proposed. Based on the local neighborhood density, in [79] Kleinrock and Silvester provided a formula for the average distance between nodes. We present later this formula as *Kleinrock's formula*. Based on the information flooding, Niculescu [67] presented three methods: *DV-Hop propagation method*, *DV-Distance method* and *Euclidean method*. These methods are commonly used in many localization algorithms. For example, method *Hop-Terrain* by Savarese is similar to *DV-Hop* [67]; and method *Sum-dist* by Savvides is essentially the *DV-Distance* method [78].

- *DV-Hop* or *Hop-Terrain propagation method*: In this method, each node only communicates with its immediate neighbors as in distance vector routing. Starting from an anchor node, hop by hop information propagation allows a node to determine its distance, in hops, to that anchor.

  Note that during the message propagation, each node only maintains and rebroadcasts packets with the smallest number of hop counts. The minimum hop counts that nodes retain will eventually be the length of the shortest path to the anchor. The distance to the anchor nodes is determined as the product of the hop counts and the average hop distance.

  The average hop distance is determined by anchor nodes dynamically. When an anchor node receives a packet containing position information from other anchors, it determines the hop counts and the actual distance between them, and the average hop distance is obtained by averaging the distance with corresponding hop counts [67]. To avoid expensive flooding, a simple method of determining the average hop distance is to use the maximum radio range directly [30];

  The advantage of the *DV-Hop* method is that it allows fewer anchors for the initial estimation, and it work well in uniform networks.

- *DV-Distance* or *Sum-dist propagation method*: This method works similarly to *DV-Hop*, the difference is that each node now measures the pair-wise distance between neighboring nodes. The hop by hop information propagation transmits the cumulative distance instead of hop counts.

  Same as *DV-Hop* method, each node only selects and rebroadcasts packets with minimum distance to the anchors. Compared with *DV-Hop*, this method is sensitive to the measurement error because of the cumulative effect. However, this method is less coarse than the previous and may still perform well in randomly deployed networks.

- *Euclidean propagation method*: This method is based on the local geometry of nodes around the anchor. It assumes that the network is densely populated.

  An anchor node initiates the propagation process by flooding its own position information in a beacon packet. Upon receiving the message, intermediate nodes can estimate their distance to the anchor and rebroadcast the message with added distance information of their own.

  If node (A) receives messages from two neighbors (B and C) coming from the same anchor (D), and if the two neighbors (B and C) are also neighbors with each other (shown in Figure 2.2), then a quadrilateral (ABDC) may be formed with the knowledge of all sides and one diagonal. Note that node A can estimate its distance to B and C; and the diagonal is the distance from B to C, which can be determined since B and C are neighbors. The determination of the other diagonal (AD) is trivial, and the diagonal is the distance from node A to anchor D.

  The algorithm allows all nodes to estimate their distance to the anchor node, and it works from the inner most circle to the faraway outside like a circular-wave form centered at that anchor.

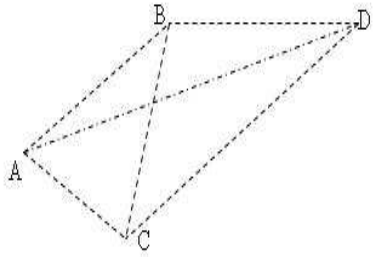

Figure 2.2: Euclidean propagation model

The advantage of this method is that it obtains precise distance estimation to the anchor node even under random deployment networks. However, it is complex and only works in high density networks. Furthermore, the error from pair-wise distance estimation is cumulative.

- *Kleinrock's formula*: Kleinrock and Silvester [79] showed that the average hop distance d_{hop} for a dense network depends only on local neighborhood density *N*:

$$d_{hop} = r\left(1 + e^{-N} - \int_{-1}^{1} e^{-\frac{N}{\pi}\left(\arccos t - t\sqrt{1-t^2}\right)}\, dt\right) \tag{2.3}$$

Nagpal *et al.* measured the average hop distance with variance *N* over several simulations from a random source [34]. Results validate Kleinrock's formula that it only slightly underestimates the measured average hop distance.

Nagpal further noticed that *N* of 15 is a critical minimal threshold for achieving low errors in distance estimation.

Range-based Localization Algorithms

After anchor nodes broadcast their location message, each normal node obtains anchors' position, and the estimated distance to each of them. Based on the distance information, different localization algorithms can be applied.

Convex Estimation *Convex* position algorithm models the known peer-to-peer communication in the network as a set of geometric constraints on the node positions [72]. Radio communication constraints are a set of geometry rules among the communicating nodes. The constraints can be radial and angular restrictions or a combination of them that are used to bound the position estimates. For example, if one node (*A*) can receive message from another node (*B*), then the distance between these two nodes should be less than node *B*'s radio transmission range. Under these constraints, a global convex optimization yields a feasible position estimation for all the normal nodes.

Based on connectivity and pair-wise angles between nodes, a linear program can be defined. A linear program (LP) is a problem of the form:

$$\textbf{Minimize } c^T x; \ \textbf{Subject to} : Ax \leq b \tag{2.4}$$

where x is the vector of variables to be solved for, A is a matrix of known coefficients, and c and b are vectors of known coefficients. The expression $c^T x$ is called the objective function, and the equations $Ax \leq b$ are called the constraints. The matrix A is generally not square (usually more columns than rows), and $Ax \leq b$ is therefore quite likely to be under-determined, leaving great latitude in the choice of x with which to minimize $c^T x$.

In two dimension system, the position is represented with pair *(*x,y), and a vector with all positions for the above equations can be formed as:

$$x = [x_1\ y_1\ x_2\ y_2\ ...\ x_m\ y_m\ x_{m+1}\ y_{m+1}\ ...\ x_n\ y_n]^T$$

where the first m entries represents anchor position and the remaining *n-m* are to be determined.

Geometrically, the connectivity of a network can be represented as a set of convex position constraints [72]. By utilizing the radio communication constraint models, it is possible to generate feasible positions for all the nodes in the network.

The Convex algorithm is a centralized algorithm. Through simulation, it is found that:

- For radically constrained connections, using a variable radius instead of a fixed radius improves overall estimation performance;
- For angle constrained connections, the angle estimation error affects the position prediction. If angle measurement can be combined with the distance information, the precision is improved.
- Results from radial and angular methods are not comparable.

Bounding-box or Min-max Method The bounding-box defines a possible area that a node may reside in. There are essentially two basic methods. The first method is a distance based approach proposed by Savvides [8], which uses the surrounding anchors' positions and the distance between them. The second method is a radio pattern based approach, which assumes regular circular radio pattern and constructs the bounding-box containing the overlapping or intersection area of neighboring nodes [72].

- **Distance based approach**: This is a general method which only assumes the knowledge of the distance from the estimating node to the anchors. Basically, we can construct a bounding-box for each surrounding anchor node. The possible area for the normal node is then determined as the intersection of those anchors' bounding-boxes.

  For example, suppose the distance from a normal node X to an anchor node A is d, and A's coordinates are (x_A,y_A), then the bounding-box for X is:

  $$[x_A - d, y_A - d] \times [y_A + d, y_A + d]$$

  If normal node X has multiple surrounding anchor nodes, then the bounding box for X is the intersection area of all anchor nodes' bounding boxes. The intersecting box is defined as:

  $$[max(x_i - d_i), max(y_i - d_i)] \times [min(x_i + d_i), min(y_i + d_i)]$$

  where (x_i, y_i) is the anchor *I*'s position, and d_i is the distance between the normal node X and the anchor *I*. In Figure 2.3 (a), node A and node E are anchors. Node B, C, and D are normal nodes. The bounding area for node D is:

  $$[x_E - d, y_E - d)] \times [x_A + (a + b + c), y_E + d]$$

  The final estimated position for the normal node is set at the center of the bounding box.

- **Radio pattern based approach**: Instead of the distance from the estimating normal node to the surrounding anchors, this approach requires circular radio propagation pattern. That is, when one normal node receives a beacon packet from a neighboring anchor, it assumes

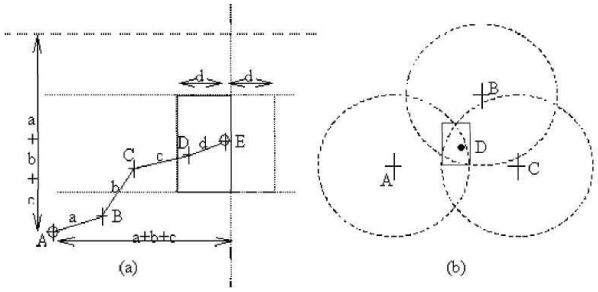

Figure 2.3: (a) Distance-based approach. (b) Radio pattern based approach.

it resides in the circular ring centered at that anchor. Accordingly, the bounding box for the normal node is the overlapping radio area from the neighboring (anchor) nodes. An example is shown in Figure 2.3 (b), which shows a rectangular box for node D (anchor nodes A, B, and C are D's neighbor).

Note that if the radio wave is circular, the second method is more precise.

Lateration Algorithm Lateration is a form of triangulation that uses least squares to estimate the position from a set of linearized equations in the form of $AX = b$.

Specially, from the radio propagation step, each normal node gets an estimated distance (d_i) to the anchor nodes with known position (x_i, y_i). If the normal node's position is *(x,y)*, then a set of equations can be expressed as:

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 = d_1^2 \\ \vdots \\ (x_n - x)^2 + (y_n - y)^2 = d_n^2 \end{cases} \tag{2.5}$$

Subtract the last equation from the first *n-1* equations, and reordering the equations, we get $AX = b$ with:

$$\mathbf{A} = \begin{bmatrix} 2(x_1 - x_n) & 2(y_1 - y_n) \\ \vdots & \vdots \\ 2(x_{n-1} - x_n) & 2(y_{n-1} - y_n) \end{bmatrix} \tag{2.6}$$

and

$$\mathbf{b} = \begin{bmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 + d_n^2 - d_1^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 + d_n^2 - d_{n-1}^2 \end{bmatrix} \tag{2.7}$$

then *X* can be estimated by least-squares with result: $X = (A^T A)^{-1} A^T b$.

The least-squares is efficient because it minimizes possible range estimation errors accumulated along the propagation path.

Probability Algorithm For an outdoor environment, received signal strength is a function of distance: the longer distance from the transmitter, the smaller the signal strength. It is noticed in [38] that the probability distribution of signal strength follows a normal distribution.

The probability algorithm starts with a table that records the signal strength (mean and standard deviation) with the changing distance. In addition, it assumes the network is fully connected, and every node in the network is present in the entire space with equal probability.

When a normal node receives a beacon packet directly from an anchor, it estimates itself to be located on a circular surface centered at that anchor(see Figure 2.4 (a) [2]). In Figure 2.4, the x axis and y axis denote the position in meters, and z axis represents the position probability. Each point on the circular surface represents a probability distribution at that location. A higher value means

[2] Adopted from [38]

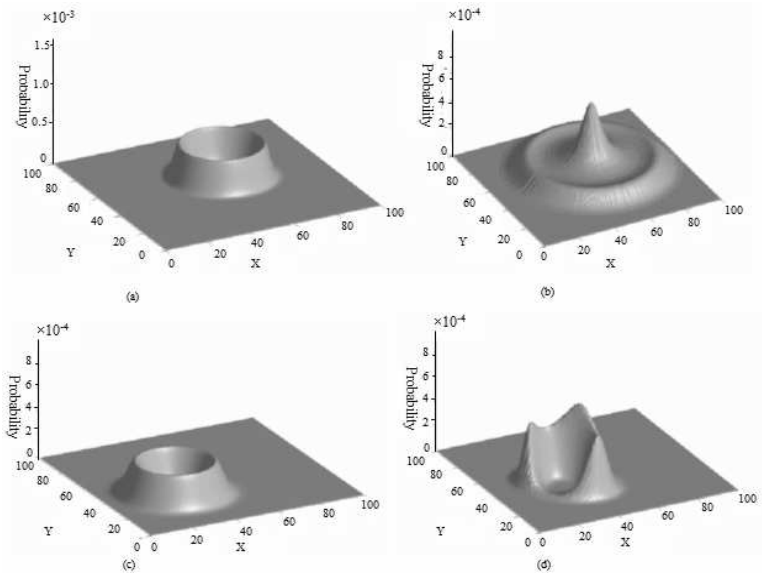

Figure 2.4: (a) Constraint with one-hop beacon message. (b) Constraint with two-hop beacon message. (c) Old position estimation of a normal node. (d) New estimation by intersecting the constraint with the old estimate.

higher probability of the node's position at that position. Upon estimation, the normal node will add its ID as well as its estimated mean and standard deviation to a beacon packet, and rebroadcast the packet to its neighbors.

If the beacon packet comes from other normal nodes, the location constraint for the normal node is not necessarily Gaussian. Actually, the packet contains a cascaded information of the estimated means and standard deviations of other normal nodes to the anchors. To estimate its position, the normal node has to process the cascaded distributions by *adding* all individual estimates (see Figure 2.4 (b)). The sum of these constraints is similar as the *convolution* of all the individual distributions.

After processing the beacon packet, the normal node updates its position by intersecting the new constraints with the current estimates. If the new estimates improve, it will wait for a specific

period of time and advertise the new position to all neighbors. Figure 2.4 (c) and (d) shows an example of the procedure. Figure 2.4 (c) gives the formal estimates of a normal node, and (d) shows the improved results to be advertised.

Multidimensional Scaling Method Multidimensional scaling (MDS) is an exploratory technique used to analyze the dissimilarity of data on a set of objects. Multidimensional scaling takes roots in two important traditions within psychology: psychophysics and psychometrics [80, 81], and it is called as "smallest space analysis" initially. Now MDS has already encompassed a collection of methods for general multivariate data analysis. Some excellent textbooks about MDS are [82, 83, 84].

Classical scaling technique is a metric multidimensional scaling technique [84], and it was originated in the 1930s by Young and Housholder [85]. Later, Schoenberg (1935) [86] and Young & Housholder (1938) provided the method for finding the original Euclidean coordinates from the Euclidean distances. In [9], the authors used classical multidimensional scaling algorithm in wireless ad hoc sensor networks for estimating the nodes' positions.

The MDS in [9] let anchors initiate the estimation procedure by broadcasting their position information to the whole network. Upon receiving a beacon packet, intermediate nodes will measure pair-wise distances to transmitting neighbors. If the intermediate node is an anchor, it will also estimate the average hop distance of the network. The average hop distance is determined by dividing the physical distance with the number of hop counts between the two anchors. If one routing path contains more than three anchors, and at least three anchors are not on a same line, classical multidimensional scaling can be used to estimate the coordinates of all normal nodes along the path.

Classical MDS works as follows [84]:

1) Compute the distance matrix $D = [d_{ij}]_{n \times n}$, where d_{ij} is the distance between node *i* and *j*, and *n* is the number of nodes in the path;

2) Compute a matrix *J* with $J = I - e * e^T / n$, where *I* is the identity matrix and $e = (1, 1, \ldots, 1)$;

3) Double center matrix *J* with $H = -\frac{1}{2} J D^2 J$;

4) Eigen decomposition matrix *H* and let $H = UVU^T$, then sort descending the Eigen matrix *V*, and change the matrix of *U* accordingly.

5) Choose an appropriate number of dimensions *p*. The coordinates of the estimated normal node in the *p* dimensional Euclidean space are given as the first *p* columns (two columns for 2-D case) in *U*.

The position information obtained through the above steps is the relative position with each other. To obtain the physical position of the nodes along the path, the estimated positions from MDS will, generally, be rotated and/or mirrored according to the reference anchor nodes along the path. Figure 2.5 shows an example.

Figure 2.5 gives a routing path containing six nodes in 2-Dimensional space, and the x-axis and y-axis denote the coordinates in x and y direction respectively. The '+' presents the true positions; 'o' denotes direct estimates through MDS; and '*' represents the rotated positions from estimates, which will further shift to the true position.

In summary of the related research for outdoor localization systems, most existing localization algorithms present good simulation results in their research, however, it is difficult to appreciate their performance because almost all algorithms require particular simulation environments (i.e. flat, open space), network conditions (i.e. anchor ratio, node density, static networks), and certain radio properties (i.e. bidirectional link, circular radio). Therefore, it is necessary to compare and

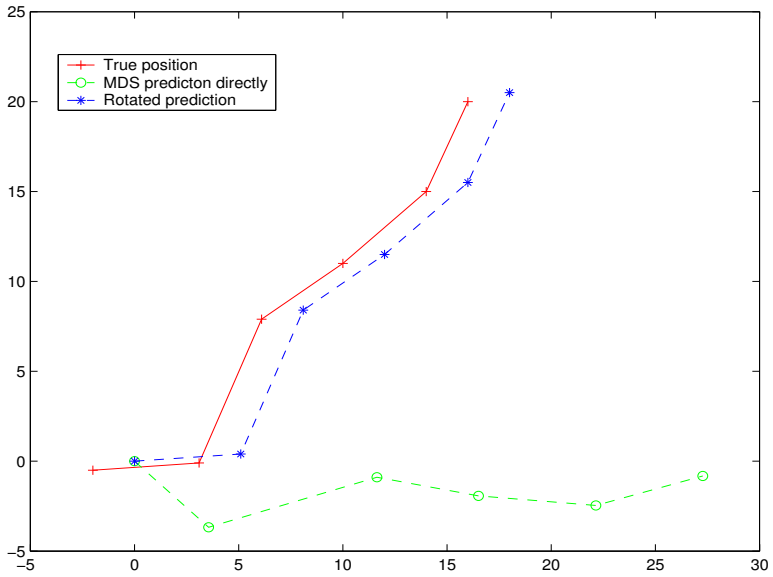

Figure 2.5: MDS localization example

identify the strength of each algorithm under the same network environment. Moreover, to our best knowledge, there exists no theoretical analysis on the estimation precision bounds for most localization algorithms. Thus it is very difficult to make critical decision if only the best estimation performance is available. Additionally, to identify or to develop an efficient localization algorithm that works on the most general environment with known precision is highly desirable.

2.3.2 Radio Range Irregularity

In recent years, localization research on wireless sensor networks has become very active. The research community usually exploits simple radio propagation models in simulations to investigate and evaluate a variety of localization algorithms and protocols. This simple model assumes spherical radio transmission pattern of fixed range at all wireless nodes [11, 16, 38]. However, results in the field widely differ from simulation results. The research community has become aware that

simplistic models may lead to flawed protocols. A key reason is that radio propagation is irregular [39, 33]. Pawlikovski *et al.* [40] surveyed 2200 published network simulation results and concluded that the simple model is adopted by most researchers and simulators. From experiments, Pawlikowski showed that the ideal radio model yields results radically different from reality. Kotz *et al.* [41] reviewed six of the most used simplistic assumptions and showed their individual impact on the obtained results. Thus, the credibility of applications/protocols based on simplistic models is dubious.

However, research on realistic models for the radio range irregularity is still in infancy, and most researchers mainly focus on these studies: a) To provide more evidence to further quantify this irregular phenomena [41, 87]; b) To clarify and demonstrate the weakness of the use of the ideal circular model in existing protocols and algorithms [40, 41]; and c) To develop other algorithms that are less dependent on the radio shape [77];

In [33], He *et al.*, for the first time, provided an irregular radio model: degree of irregularity (DOI). The model assumes an upper and lower bound on the radio propagation range. When a neighboring node is beyond the upper bound, then it is out of the communication range; and when it is within the lower bound, the nodes are guaranteed to be within communication range. If the distance between the neighboring nodes is between these two boundaries, the communication is dependent on the actual radio range in that direction. In simulation, the particular radio range in each direction is calculated based on a random number and a pre-assigned irregularity factor. While this is a good start for the research in this area, the DOI model does not take the environment into account and this model usually results in undetermined and abrupt changes of range values in all directions.

Zhou *et al.* [39] extended the DOI model by also considering the radio interference among devices. The new model is called radio irregularity model (RIM). RIM is based on experimental results made with a pair of MICA2 motes. RIM was developed to analyze the impact of radio range irregularity on MAC and routing protocols.

Apart from the complexity and the un-straightforward nature of the dependent path loss model and energy-fading pattern on the multipath environment, RIM is derived from measurements at a lower radio frequency below 916 MHz. The conclusions from those experiments may not be equally applicable for the IEEE 802.11b standard at 2.4 GHz.

Our preliminary measurements on various environments using IEEE 802.11 wireless Ethernet standard indicate that the radio transmission range at 2.4 GHz UHF (ultra high frequency) band may not smoothly vary in different directions depending on the presence of obstacles (ex. trees and stones) in vicinity. This find is significantly different from the conclusion by Zhou *et al.* [39].

Consequently, it is of particular importance to: 1) research a convenient radio range irregularity model for IEEE 802.11 that is easily adjustable according to actual operating environment; 2) identify the impact of the real radio propagation on the performance of existing localization systems; and 3) explore optimal radio propagation method that can improve the localization performance even under the irregular radio transmission.

CHAPTER 3

RESEARCH DETAILS FOR INDOOR LOCALIZATION SYSTEMS

ARIADNE consists of two modules as illustrated in Figure 3.1 - namely map generation and search - that are developed in Section 3.1 and Section 3.2, respectively. In Section 3.3, the localization on the indoor mobile user is analyzed. And Section 3.4 introduces the sniffers deployment method.

3.1 Map Generation Module

Map generation includes multiple steps: Subsection 3.1.1 develops the first step that consists of capturing the characteristics of the floor plan and producing a 3-D model necessary for ray tracing. Subsection 3.1.2 explains how ray tracing is used for the determination of the individual ray contribution to the signal strength on a grid of points. A propagation model is proposed in Subsection 3.1.3, and its parameters is solved in Subsection 3.1.4 and Subsection 3.1.4 using simulated annealing.

3.1.1 Floor plan interpretation

The main purpose of the floor plan interpretation is to integrate the geometry acquisition process as an automatic procedure. The major task of the interpretation process is to extract the structural parameters from construction CAD files or floor plan image files.

Structural information is extracted from the picture using basic image processing techniques, in which a picture is denoted as a matrix. Each element in the matrix has a value corresponding to the brightness of the pixel at the corresponding position, which is an integer between 0 and 255.

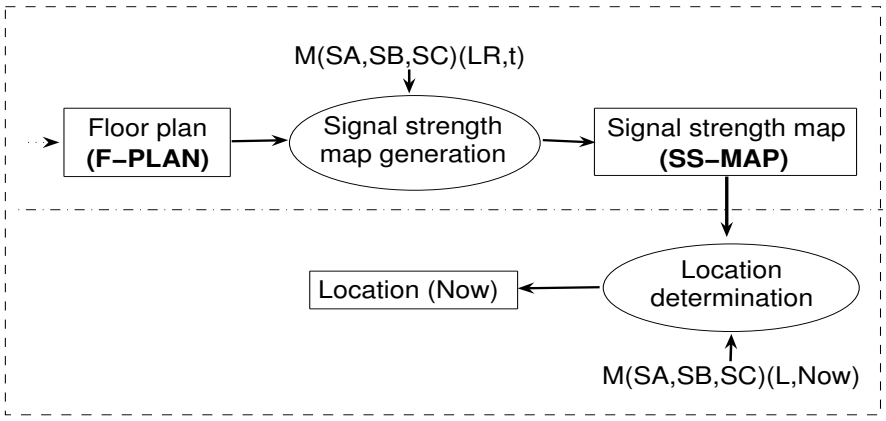

Figure 3.1: ARIADNE system

The 0 corresponds to black and 255 to white. If the pixel value of the lines in the picture is denoted by 0, then the grouping of a set of connected 0 value pixels, vertically or horizontally, yields a line.

The wall geometry information is constructed by extending the lines vertically in 2D image with base coordinates and the floor height. By stacking the wall information at each floor, overall structural representation of the building is obtained. Similar to most previous research [88], a wall/floor is modelled as a single plane in the middle. The offset between refracted and incident rays is ignored.

Figure 3.2 shows a site floor plan where ARIADNE was tested. The floor is about 150×120 *ft.*. Figures 3.3(a) and 3.3(b) display respectively the 2D topview and the 3D view with height information extracted from the original plan. Geometry information of the walls and floor/ceiling is stored in a database table for ray tracing.

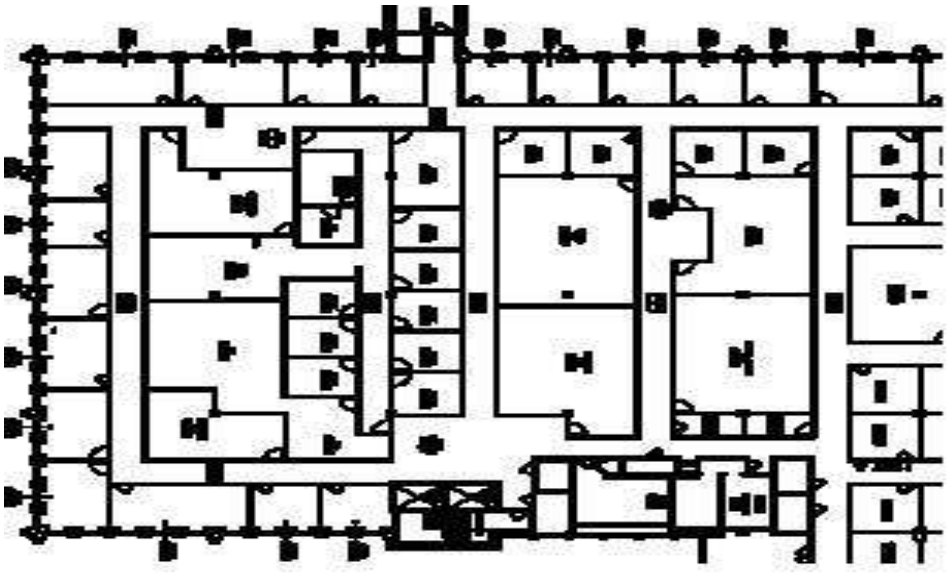

Figure 3.2: Original floor plan

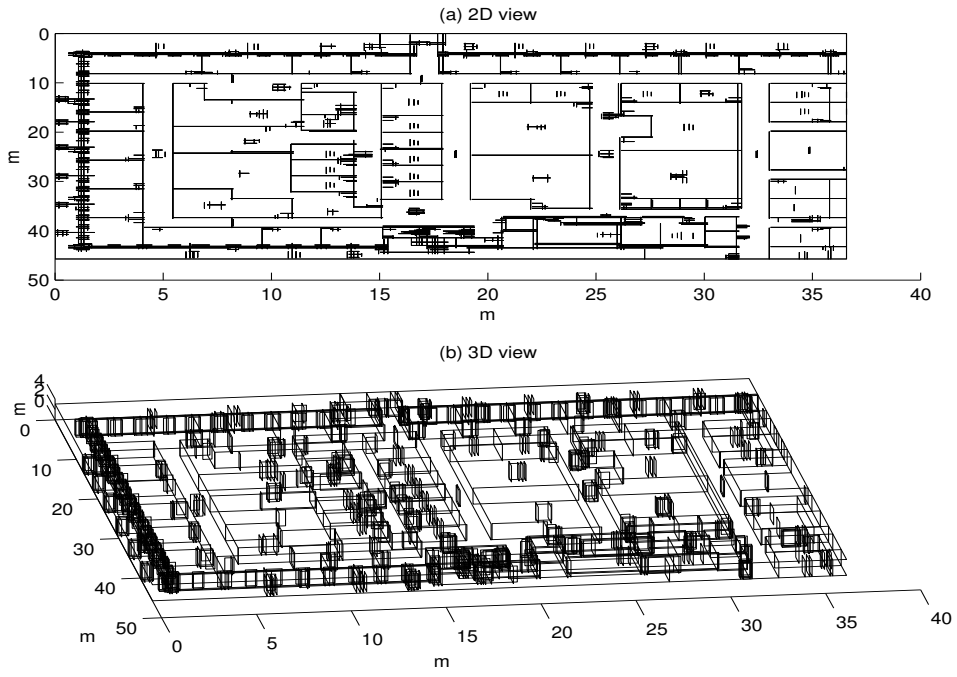

Figure 3.3: Floor plan interpretation output

3.1.2 Ray tracing

Ray tracing (RT) approximates the radio propagation with a finite number of isotropic rays emitted from a transmitting antenna [89]. For an omnidirectional antenna, each ray is assumed to transmit with the same amount of energy at the transmitter, and the energy of the rays will be attenuated at walls or floors due to reflections and transmissions. Ray tracing technique is widely used to simulate the radio propagation in indoor environment [90, 91, 88].

Ray imaging techniques are used to record each ray from the transmitter to the receiver. In the ray imaging technique, the transmitter is assumed to be reflected at each surface around it to produce image transmitters, the reflected rays to the receiver from the real transmitter are considered as direct paths from the mirror images of the true transmitter. Based on *geometrical optics* (GO), each ray from the transmitter to the receiver can be exactly determined. The detailed ray technique is omitted here for lack of space (a good reference can be found in [90, 92, 93]), but instead, several key points of ARIADNE are emphasized.

- Similar as the research by Hassan-Ali and Pahlavan in [58] and Bertoni et al. in [94], the diffraction and scattering effect are neglected in the proposed propagation model because of the minor contribution of the radio in this band;
- Only rays with power above a fixed threshold [95] are considered because highly attenuated rays do not reach the receiver in reality even though a transmission path exists in theory.
- Similar to [58], the multipath power at receiver is determined as the sum of all individual powers regardless of the phase of each path.

Figure 3.4 depicts a simple scenario where three rays are shown from the transmitter T to the receiver R. Each ray r_i is composed by multiple segments where distance of the j^{th} segment is d_{ij}.

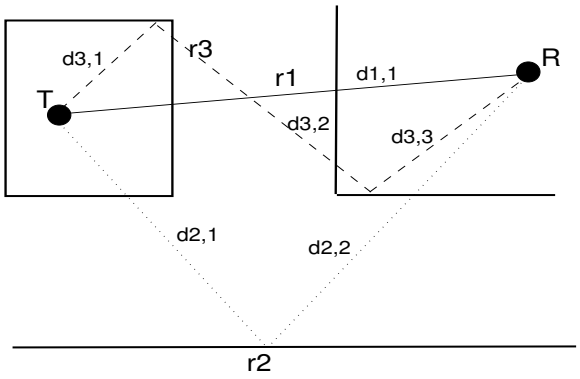

Figure 3.4: Radio propagation with ray tracing

Direct path (ray r_1) is denoted by a solid line. The other two paths (ray r_2 and r_3) are indirect and contain reflections. The faint dashed line (ray r_2) has one reflection and dotted line (ray r_3) has two reflections, respectively. The distances traversed by each ray is also depicted in the figure.

3.1.3 Radio propagation model

As explained in Section 3.1.2, the signal power at the receiver is the accumulated multipath power from all individual rays from the same transmitter. For each ray, the attenuation path loss includes three components:

1. The distance-dependent path loss, which is assumed as free space propagation loss;

2. The attenuation due to reflections, which is the product of the reflection coefficient and the total number of reflections from transmitter to the receiver;

3. The attenuation due to transmission, which is the product of the transmission coefficient and the total number of transmission walls.

Consequently, the model is defined as:

$$P = \sum_{i=1}^{N_{r,j}} (P_0 - 20log_{10}(d_i) - \gamma \cdot N_{i,ref} - \alpha \cdot N_{i,trans}) \tag{3.1}$$

where P is the power (in dB) at receiver, $N_{r,j}$ is the total number of rays received at the receiver j; P_0 is the power (in dB) at a distance of 1 meter; d_i, $N_{i,ref}$, and $N_{i,trans}$ represent the total transmission distance, the total number of reflections and the total number of (wall) transmissions of the i^{th} ray, respectively. γ is the reflection coefficient, and α is the transmission coefficient.

In Figure 3.4, the transmission distances for three rays (r_1, r_2, and r_3) are $d_{1,1}$, $d_{2,1} + d_{2,2}$, and $d_{3,1} + d_{3,2} + d_{3,3}$, respectively. Ray r_2 has one reflection, and ray r_3 has two reflections. All three rays have two wall transmissions.When starting from transmitter T, all three rays are assumed to hold the same amount of power. With different transmission conditions, the final signal power of each individual ray observed at the receiver R are different. And the overall signal power at the receiver R is the sum of the powers from all received rays.

The site specific parameters (N_{ray}, d_i, $N_{i,ref}$, and $N_{i,trans}$) in Equation 3.1 can be obtained directly from ray tracing as described in the Section 3.1.2. The other three parameters (P_0, γ, and α), in other similar research, are usually derived from tedious measurements. ARIADNE does not require extensive on site measurements. Instead, simulated annealing (SA) technique is used to determine optimal values for the three parameters of the proposed model. ONE reference measurement only is required.

3.1.4 Parameters Estimation

To estimate the radio propagation parameters (reference power of the ray P_0, reflection coefficient γ, and transmission coefficient α), some measurements at reference positions inside the

building are needed. If a maximum of n reference measurements are available, a linear system of $Ax = b$ (derived from equation 3.1) can be used to determine the three unknowns $x = [P_0\ \gamma\ \alpha]^T$.

To solve the linear equations, the method of least squares could be used. However, it is difficult. As stated earlier, only rays with power above certain threshold are considered in the radio propagation model. Or in other words, from ray tracing simulation, a maximum number of N rays may exist, theoretically, from the transmitter to the receiver. In reality, only n $(n < N)$ rays are actually received because of the different attenuation along each individual path. Since some rays are too weak to contribute the energy at receiver, they must be eliminated from the linear system. Such an elimination process is very difficult at this stage because of the lack of the energy information (again, the *Chicken and Egg Dilemma*). In this research, we use simulated annealing algorithm to search the optimal value of $x = [P_0,\ \gamma,\ \alpha]^T$.

Simulated Annealing Search Algorithm

Simulated Annealing (SA) [96, 97] is a method used to search for a minimum in a general system. It is based on the process of the way a metal cools down to the optimal state (the annealing process). SA's major advantage is an ability of a random search which not only accepts changes that decrease objective function, but also some changes that increase it. Thus, SA method can achieve global optimization without getting trapped at a local minima [98].

The original Metropolis scheme [96] indicates that an initial state of a thermodynamic system is chosen at energy E and a desired temperature T. Holding at that temperature T, the initial configuration is perturbed and the change in energy dE is computed. Applying *Monte Carlo sampling techniques*, the physical annealing process is modelled successfully by computer simulation

methods. A convenient formula can be borrowed from thermodynamics:

$$\widehat{P(E)} = exp(-\frac{E}{kT}) \tag{3.2}$$

which expresses the annealing probability $\widehat{P(E)}$ of a change on energy E at temperature T, where k is Boltzmann's constant.

Given initial values of $x = [P_0\ \gamma\ \alpha]^T$ at a temperature T, the power of each individual ray can be computed (Equation 3.1). (The initial values can be any positive numbers, however, better values will minimize the search time. Generally, better values can be derived from literature.) Neglecting those rays with power below the threshold, and summing the powers of all others, yield the multipath power at the receiver. The least minimum squared error (Equation 2.2) allows the comparison of the power estimates fitness with the measurements, and henceforth the adjustment of the parameters of x accordingly.

To adjust the parameters, a random movement is generated by adding a deviate from the Cauchy distribution to each parameter of $x = [P_0\ \gamma\ \alpha]^T$:

$$x_{i+1} = x_i + T \cdot tan(\widehat{P}),\ i = 1, 2, 3 \tag{3.3}$$

The cooling schedule for the temperature *T* can use a simple method similar to [97]:

$$T_{i+1} = a \cdot T_i,\ a \in (0, 1) \tag{3.4}$$

Consequently, the Simulated annealing search algorithm can be detailed below:

1) Define initial values for $x = [P_0\ \gamma\ \alpha]^T$.

2) Define the temperature, T_{max} for highest temperature and T_{min} for the cooling down value;

3) Calculate the annealing probability from Equation 3.2;

4) Update the displacement for the parameters using Equation 3.3;

5) Calculate the fitness between the estimates and the measurements using equation 2.2: if a better agreement is obtained, keep the displacement from the above step; else, keep the displacement with certain probability;

6) Update the temperature *T* by equation 3.4, and repeat steps 3, 4, and 5 until $T < T_{min}$ or specified minimum errors is achieved.

Simulated annealing method can effectively estimate parameter tride $x = [P_0\ \gamma\ \alpha]^T$ with only ONE reference measurement.

3.2 Search module: Clustering-based Search Algorithm

To locate a mobile user, the current user's signal strength measurement triplet is searched from the signal strength map *SS-MAP* for a match. Currently, most search algorithms are based on the *LMSE* and select a single location as the estimate. This method works if a detailed and precise *SS-MAP* for the building is available. As indicated in many papers, the signal strength is observed to be very dynamic at different measurement times, and to collect a fine-grid signal strength map is time-consuming for large scale building deployments. Consequently, the *LMSE* method will not generate optimal estimates in most circumstances. Therefore, it is difficult to make a decision if this method is to be used exclusively.

ARIADNE proposes a clustering-based search algorithm for the indoor localization of a mobile user based on the following findings:

- ARIADNE constructs fine-grid signal strength map *SS-MAP* based on the radio propagation model and the site-specific geometry of the building;
- The *SS-MAP* from the propagation model provides real-time estimates without further human intervention. However, it is only a close fit to the measurements, or in other words, it is imprecise and small estimation errors are expected for some locations;
- Consequently, *LMSE* may result multiple possible locations in the SS-MAP table, or the unique location corresponding to the LMSE is not necessarily the right position;
- If a set of positions (corresponding to low mean square error with respect to a predetermined threshold) is selected, the positions can be grouped into several clusters. The largest cluster will generally have higher probability to contain the true position for the mobile user.
- The location estimates with the clustering-based search method may provide larger errors for some positions, however, the overall estimation error gets lowered and the confidence is improved.

The clustering-based search algorithm is a two-phase search algorithm. The first phase is named as data collection and cluster preparation phase, and it is introduced in Subsection 3.2.1, where a set of candidate locations with lower mean square error within the threshold are selected and preprocessed with the purpose to neglect isolated locations from the set. The second phase is clustering phase, and it is presented in Subsection 3.2.2, where the remaining candidate locations are grouped into several clusters and the center of the largest cluster is chosen as the final estimate.

3.2.1 Data Collection and Cluster Preparation Phase

In this phase, the current signal strength measurement triplet $M(SA, SB, SC)(L, Now)$ of mobile M at some location L is compared with all records from the estimated *SS-MAP*. In stead of selecting only a single location for estimation, ARIADNE select a set of candidate locations according to a predetermined mean square error (MSE) threshold.

Because of the imprecise nature of the estimated *SS-MAP*, some of the selected candidate locations may be scattered around the floor plan. In order to prepare the candidate locations for clustering, the scattered or isolated (unlikely) location points must be detected and omitted from the set of candidate locations. The isolated position is characterized by a larger distance from its location to all other candidate locations. For example, if there are total N candidate locations in a selected location set, let x_i and x_j be two location clusters with m and n candidate locations, respectively, $m, n \in [1, N]$, and $m + n \leq N$; and let $d_{i,j}$ be the minimum pairwise distance from any member instances of these two clusters.

$$d_{i,j} = min(dist(x_{i,r}, x_{j,t})) \tag{3.5}$$

where r and t represent the position instance in cluster x_i and x_j, respectively; $1 \leq r \leq m$, and $1 \leq t \leq n$. If the candidate location cluster x_i has larger distance $d_{i,j}$ to every other clusters, it means that cluster x_i is an isolated cluster. If further this cluster contains much smaller populations, it may be omitted from the candidate location set.

Figure 3.5 shows an example of a set of positions in space. In the figure, position 8 is isolated from all others; positions 5 and 7 are close to each other and they may be treated as one group which is again separated from others. Figure 3.6 gives the distance information between (group) positions

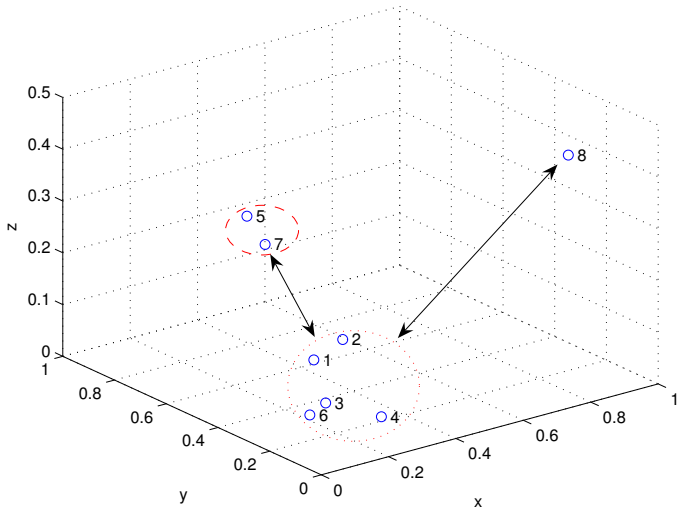

Figure 3.5: Position isolation example

for the data set in Figure 3.5. If positions of $\{1,2,3,4,6\}$ are grouped into one cluster, and positions $\{5,7\}$ form a second cluster, then the minimum distance between these two clusters is 0.3340. If positions of $\{1 \sim 7\}$ are to be grouped into a bigger cluster, and the position 8 is another group, then the minimum distance between them is 0.8311. For the data set in the example, positions of $\{5,7,$ and $8\}$ may be neglected during this preparation phase.

3.2.2 Clustering Phase

After the cluster preparation phase, most of the remaining positions have neighbors close to them. Consequently, the main purpose of the clustering phase is to determine the intrinsic grouping of the set for these positions, and to select the right cluster for the estimates.

To group the set of points in space, two common methods are available. The first one is an hierarchical clustering method, and the second one is K-clustering method.

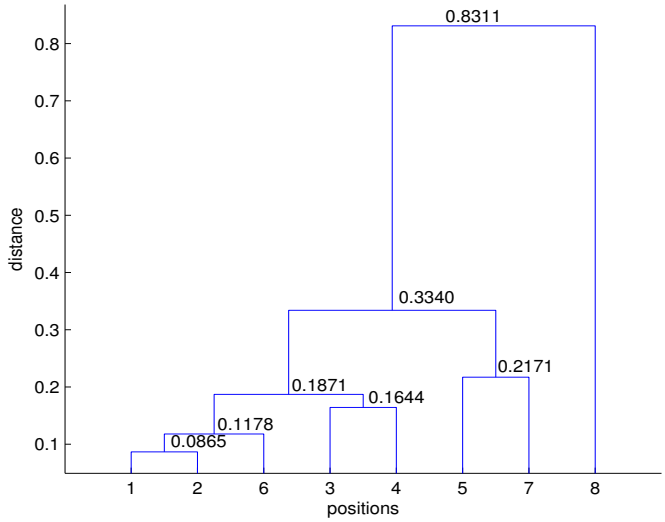

Figure 3.6: Distance between objects

- The hierarchical clustering method [99] produces a hierarchy tree structure of the original data set. The leaves are individual elements and internal nodes are sub-clusters. Each level of the tree represents a partition of the original data set of several sub-clusters. Figure 3.6 is an example of the hierarchical clustering method.

- K-clustering method searches the best k cluster centroids, and partition the data set by assigning each point to its nearest centroid. K-means clustering [100] is one of the most common K-clustering algorithm.

The clustering procedure is more observable if hierarchical structure of the original data set is obtained (Figure 3.6). If a minimum of three neighbors are selected ($\{1,2,6\}$), it translates to the *closeness elimination scheme* addressed by Prasithsangaree in [62]. And if only two neighbors are chosen ($\{1,2\}$), it is the *two closest neighboring scheme* by Pandey [63]. However it is difficult to determine the exact number of neighboring positions that should be selected in order to obtain

an optimal estimation. Hence, the idea of clustering is extended to incorporate larger number of estimates available by ray tracing. The K-means clustering algorithm is an unsupervised learning method. It starts with randomly selected cluster centers, and the final clustering performance is sensitive to them. Similarly, it is difficult to decide the optimal number of clusters for any given data set.

Unfortunately, there are no general theoretical solutions for these difficulties. This study heuristically explores these problems through simulations. It is found that the strict selection of fixed number of neighboring positions using hierarchical clustering algorithm does not yield satisfying results in most cases. On the contrary, after neglecting the isolated positions in the preparation phase, the k-means algorithm generally yields better estimates.

To determine the actual number of clusters, we select one that provides better separation of the original data as defined in Equation 3.6.

$$D_c = min(\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{j,i} - \overline{x}_{i,ctr})^2)^{1/2}) \tag{3.6}$$

where D_c is the total distance of all locations to the respective cluster centers, a small value represents better separation for the determined number of cluster. N is the predetermined number of clusters; M is the maximum number of position points in i^{th} cluster; $x_{j,i}$ denotes the coordinates of j^{th} position in i^{th} cluster; and $\overline{x}_{i,ctr}$ is the center coordinates of i^{th} cluster.

The sensitivity of the selection to initial cluster centers is minimized by running the clustering algorithm multiple times and an averaging result is used.

3.3 Mobility Analysis

To monitor a mobile user, a common method is to maintain and analyze a sliding window with a series of samples of signal strength from the mobile user within a time period. At each signal strength record, the stationary localization method is used individually. Then all location information is connected to give the user's position on a continuous basis.

ARIADNE exploits a similar idea as Abhijit, Ellis, and Fan [101] to track a mobile user in the building. The method is based on the fact that a mobile user does not move arbitrarily within the building. Instead, there is correlation between the current position with the previous location, for example, the distance between two continuous locations can not exceed certain limit. In this clustering-based search algorithm, the largest cluster is selected to denote the possible location of a stationary user. If the clustering history is used, the decision criteria may additionally be restricted with the history information. This may produce better estimates in reality. An example scenario is shown in Figure 3.7.

In Figure 3.7, the mobile user's previous location is denoted by P_1. Two candidate current positions are denoted by P_2 and P_3. If (stationary) clustering-based search algorithm is to be used, the location of P_3 should be selected because of the larger population in the clustering group. However, the distance between P_1 and P_3 is beyond the reasonable limit for the mobile user inside the building within the sampling time period. Consequently, the center at P_2 is chosen as the current location of the mobile user.

One shortcoming for this method is that it requires additional memory to maintain a history record. Moreover, the position estimation performance relies on the correctness of the history

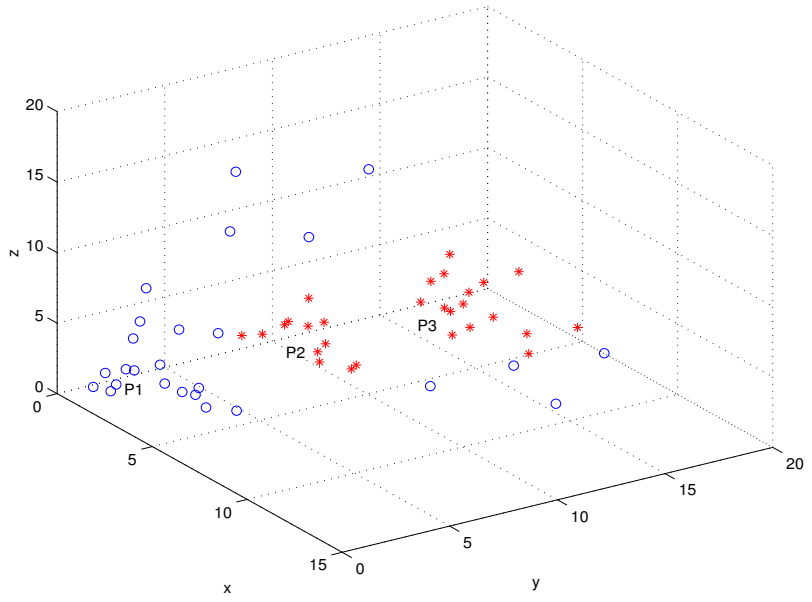

Figure 3.7: Decision making in tracking mobile user

record. If the initial estimation is not prefect, the subsequent position estimation tends to generate unacceptable results. Consequently, a mechanism to automatically reiterate the localization algorithm is required.

In this research, two principles are used to restart the localization algorithm:

1. If the history constraint results in a cluster with significant smaller populations according to a predetermined threshold;

2. If the history constraint results in a position on the other side of a partition;

3.4 Sniffers Deployment Analysis

This section addresses the impact of sniffers configuration on the performance of the indoor localization. The interest is in the following problems: i) how to optimally deploy the sniffers in

order to achieve the best precision in location estimation? and ii) for a given deployment, what is the highest achievable precision bound for the estimation?

This dissertation makes the following assumptions:

1. It does not consider the radio attenuation effect from partitions or obstacles like walls, and furniture;

2. It does not consider the effect of the radio coverage on the performance of the indoor localization, or it assumes that any mobile users at the interested locations are always effectively sensed by at least three sniffers;

3. The received signal strength SS for a given location at a sniffer is given:

$$SS \in [SS_{true} \cdot (1 - \delta), SS_{true} \cdot (1 + \delta)] \tag{3.7}$$

where SS_{true} is the true signal strength value (theoretical value) at that location, and δ is the maximum perturbation of the estimated signal strength from the indoor radio propagation model.

4. It assumes that the received signal strength SS_{true} uniquely maps to the range r_{true} from the transmitter at the position to the sniffer, i.e., $SS = f(r)$. Let the δ_r be the uncertainty of the radio range corresponding to the perturbation of the signal strength δ, then the range r is with the range:

$$r \in [r_{true} \cdot (1 - \delta_r), r_{true} \cdot (1 + \delta_r)] \tag{3.8}$$

If $r1, r2$,and $r3$ are radii from an interested position *P* to three sniffers, then the location of the position *P* is within the overlapping area of three rings (Please, see Figure 3.11).

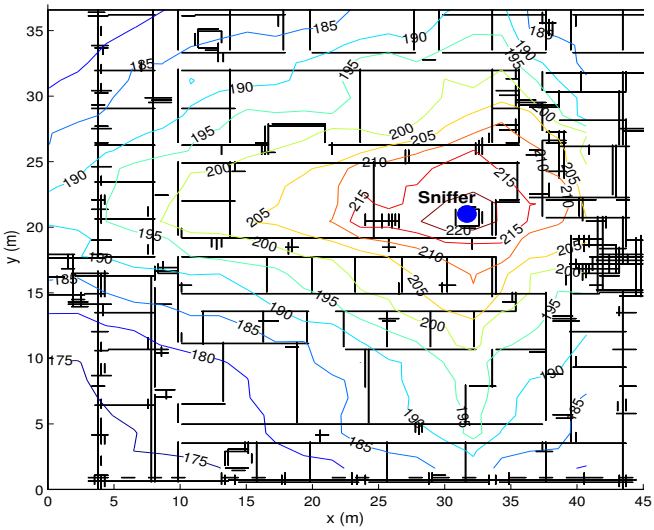

Figure 3.8: Signal strength vs. range for indoor radio propagation

Note that the radio range for the indoor radio propagation is not **circular** due to the attenuation from the partitions, and a typical indoor radio transmission is given in Figure 3.8. The figure shows the floor plan and the received signal strength at a sniffer at difference locations, the x- and y-axis denote the dimensions for the floor in two directions, and the *circular* lines are positions with particular signal strengths where the numbers on the lines are the signal strength readings. However, by simplify the site specific attenuation effects, it is potential to understand the essential relationship between the uncertainty of the localization and the sniffers deployment strategy.

Therefore, the problem of the sniffers deployment can further be expressed as the following two statements of problems:

1. **Problem a:** For a given interested position *P*, how can we optimally deploy the three sniffers with radii of $r1, r2, r3$, such that the coordinates of the given position *P* could be determined with minimal uncertainty?

2. **Problem b:** For a floor space with three deployed sniffers, if we do not consider the effect of the partitions and obstacles, what is the average uncertainty (in meters) of the position estimation for the floor?

We consider these two problems in Section 3.4.3, and we will further analyze the **problem b** in Section 5.2.1 at Chapter 5. For both problems, we first build a relative coordinate system. In this system, one of the sniffers fixes its position at origin, and by changing the relative positions of the other two sniffers, we analyze the impact of the sniffers deployment on the localization performance. We introduce this coordinate system in Section 3.4.1.

3.4.1 Definition and Coordinates

This section defines the terms that are necessary for the research.

[Definition 1] **Uncertainty Area**: For a position P covered by three or more sniffers $S_i,\ (i = 1, 2, \ldots, n)$, the **uncertainty area** for the position P is defined as the common overlapping area of multiple rings $RS_i,\ (i = 1, 2, \ldots, n)$ that centered at those sniffers. Suppose a ring surface RS_i is composed by a point set $\{q_{RS,i}\}$, then the **uncertainty area** of the position P is given:

$$S_{ua,P} = \{q_{RS,1}\} \cap \{q_{RS,2}\} \cap \ldots \cap \{q_{RS,n}\} \tag{3.9}$$

Figure 3.9 illustrates the scenario of the deployment for three sniffers of $O1, O2, O3$ and an interested position P. If there is no measurement error, the ideal radio range of the three sniffers are r_1, r_2, r_3, respectively; and consequently the position at the intersection of the three circles is the location P. Considering the measurement errors, and assume the range perturbations of the three sniffers are $\delta_i, (i = 1, 2, 3)$. Thus the possible sensing area at each sniffer would be a surface of a circular belt (or a ring surface between radii $r_i \cdot (1 \pm \delta_i),\ (i = 1, 2, 3)$), therefore, the probable

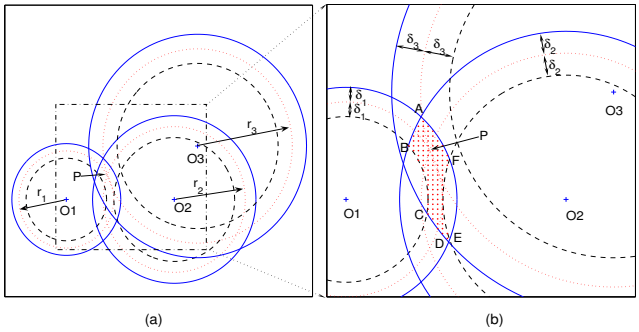

Figure 3.9: Uncertainty area of the indoor location estimation

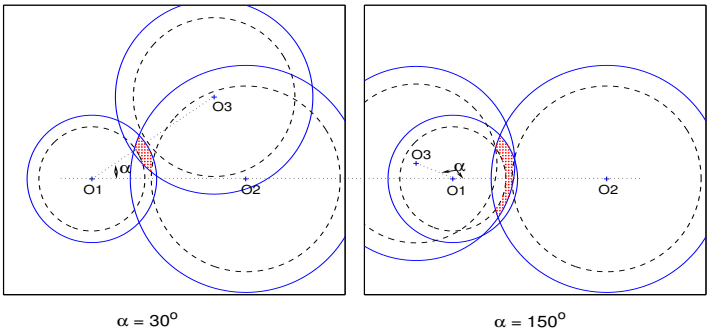

Figure 3.10: Uncertainty area of two different sniffers deployment

location for the position P would be an *area*, instead of a *point*, that are formed by the overlapping of the three rings.

Figure 3.9 (a) is the overall configuration of three sniffers. The interested area within the dotted rectangle in figure (a) is magnified and the details are presented in the figure (b), where the possible location of the position P is given by the surface $ABCDEFA$.

Figure 3.10 further shows the *uncertainty area*s for the same position P at two deployment strategies. It shows that different sniffers deployment methods do affect the *uncertainty area* of the

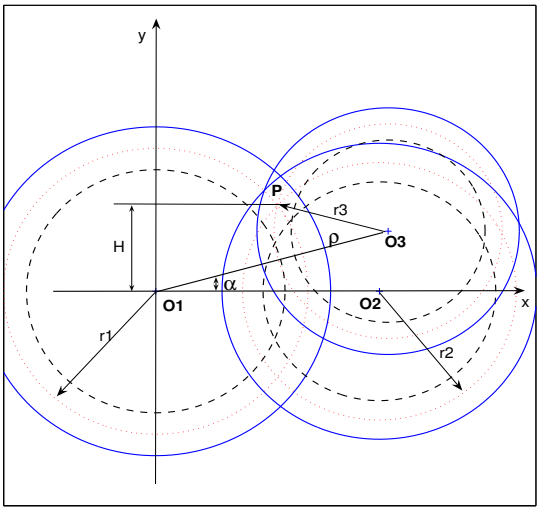

Figure 3.11: Sniffer configuration definition

location estimation. And it is obvious that smaller uncertainty area gives more confidence for the location estimation.

[Definition 2] **Deployment Coordinate System** for the space: assume three sniffers $O1, O2, O3$ are deployed on a floor, and the radio ranges from three sniffers to an interested position *P* are $r1, r2, r3$, respectively; and the maximum range perturbation of them is δ_r. As indicated in Figure 3.11, we let the sniffer *O1* fix its position at the origin, and we randomly select another sniffer *O2* to be on the right side of the x-axis. The y-axis is defined such that the third sniffer $O3$ is with positive coordinate in y direction for the coordination system.

As shown in Figure 3.11, we denote the angle formed by the line of $O3O1$ and the x-axis α and the distance from the sniffer $O3$ to the origin ρ. Therefore, the relative position between the

third sniffer $O3$ and the rest two sniffers can be uniquely determined by the combination of α and ρ. Additionally, the distance from the position P to the x-axis is denoted as H.

[Definition 3] **Average Uncertainty Distance** d_{AUD} is defined as the average distance of all points in the point set of the *uncertainty area*.

Suppose an uncertainty area S_{UA} contains total n points in the set (Please, see Definition 1 and Equation 3.9), and d_{ij} is the distance between two points i and j, ($i, j \in [1, n],\ i \neq j$), then:

$$d_{AUD} = average\{d_{ij}\} \tag{3.10}$$

[Definition 4] **Maximum Uncertainty Distance** d_{MUD} is the maximum distance of all points in the *uncertainty area*. Follow the above Definition 3, the *Maximum Uncertainty Distance* is given as:

$$d_{MUD} = \max\{d_{ij}\} \tag{3.11}$$

In the following sections, we first introduce the method to calculate the *Uncertainty Area* for the overlapping rings in Section 3.4.2; then we present the sniffers deployment strategy in section 3.4.3.

3.4.2 Method to Calculate the Overlapping Area for Multiple Circles/Rings

To mathematically formulate the overlapping area of multiple intersecting circles/rings as shown in Figure 3.11 is very complex because the common surface depends on the relative configurations of the rings. In this dissertation, in stead of presenting a theoretical formula for the problem, we introduce a straightforward method that can be used to numerically find the area of intersection of a number of circles/Rings on a plane.

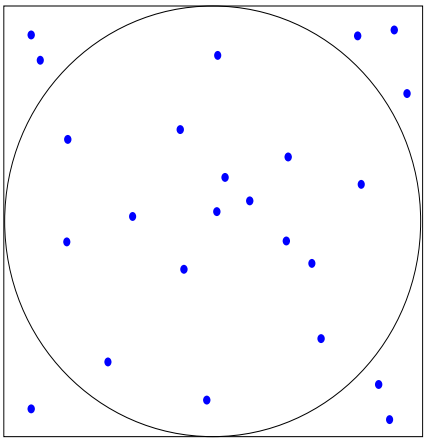

Figure 3.12: Area calculation method

Consider a single circle with the minimal bounding square as shown in Figure 3.12, suppose the radius of the ring is r, then the area of the circle is $\pi \cdot r^2$, and the area of the square is $4 \cdot r^2$. If we slice the space of the square into a grid of positions with ξ apart in both the x and y directions. When the grid resolution becomes smaller, i.e. $\xi \rightarrow 0$, and if one is going to count the total number of points on grid coordinates within the circle ($\Sigma(circle)$) and within the square ($\Sigma(square)$), the ratio of them would be approached closer to the ratio of the area of the circle and the bounding square. Or in other words:

$$\lim_{\xi->0} \frac{\Sigma(circle)}{\Sigma(square)} = \frac{\pi \cdot r^2}{4 \cdot r^2} = \frac{\pi}{4} \tag{3.12}$$

Similarly, to determine the intersection area of multiple circles on a plane, we developed the following procedures:

1. Select one circle (assume its radius is r), and slice the bounding square into a collection of grid positions at ξ apart from both x and y directions; and we assume the total number of points within the square is N;

2. Counting the total number of points n that are also within all other circles;

3. The intersection area of all circles is given by $\frac{4 \cdot r^2 \cdot n}{N}$.

The determination of intersection area of multiple rings/annuluses is similar as the above procedures except that the n represents the total number of points on the common surface for all rings (Please, see Figure 3.9).

3.4.3 Optimal Sniffers Deployment Method

This section analyzes the *uncertainty area* of interested locations under various sniffers deployment methods. The interest is to find reasonable solutions for the two problems defined at the beginning of the Section 3.4, and they are addressed in next two sections.

The Impact of Sniffers Deployment on An Interested position

This section tries to solve the first problem **Problem a**, i.e., for an interested position, how to deploy three sniffers such that the estimation uncertainty for this position is minimal. We proceed as follows: following the definition in Figure 3.11, we systematically change the relative position (by H) between the interested position *P* and two of the sniffers *O1,O2*, then at any instance of these position configuration, we update the third sniffer *O3* (by α), and calculate the *uncertainty area* for the position *P*. The sniffers configuration that generates the minimal *uncertainty area* is the optimal deployment method.

The experiment is carried as follows:

1. Randomly deploy three sniffers with range of $r1, r2, r3$, and assume a perturbation δ for all the range. The uncertainty area for each sniffer is a ring, which is defined as the part of the circle's space between radii of $r_i \cdot (1 + \delta)$ and $r_i \cdot (1 - \delta)$, $i = 1, 2, 3$;

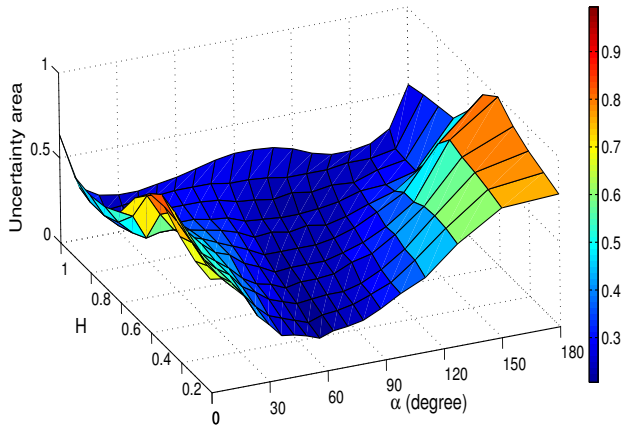

Figure 3.13: Uncertainty area at various configurations

2. The ring with the smallest radius is selected as the first sniffer $O1$, and the radius of the first sniffer $r1$ is normalized as unit length of 1, i.e., $r1 = 1$;

3. The second sniffer $O2$ moves along the right side of the x-axis, such that the distance H from the point *P* to the x-axis changes between $[0, 1]$;

4. At any configuration of sniffer $O1$ and $O2$ for a given distance H, the third sniffer $O3$ rotates its position around the interested point *P*. In the experiment, we let the angle α change between $[0, 180]$.

The normalized *uncertainty area* for the position *P* for all configurations are computed and the typical results are given in Figure 3.13. In the figure, the x-axis denotes the angle from 0^o to 180^o, the y-axis represents the distance of from the point *P* to x-axis, and the z-axis gives the normalized *uncertainty areas*.

Figure 3.14 further illustrates the changes of the *uncertainty area* (in shaded surface) for a set of similar configurations. In this example, the positions of sniffers $O1, O2$ and the position P are

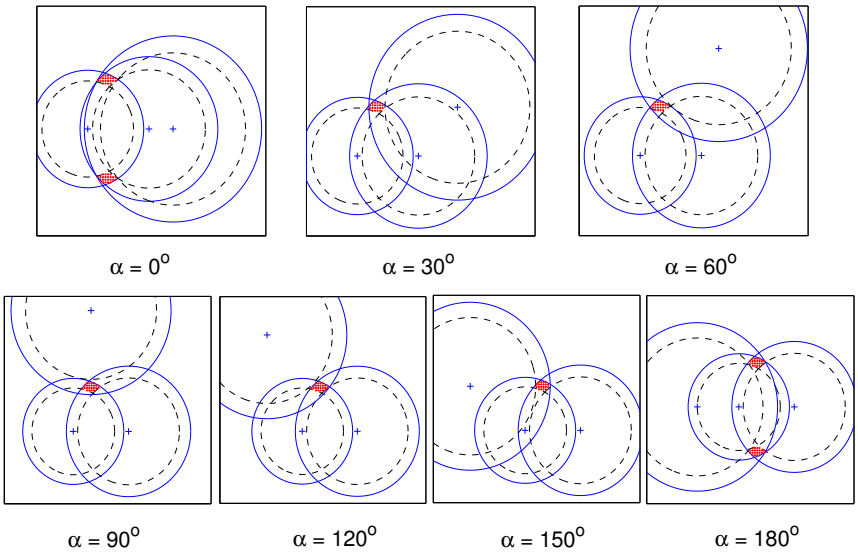

Figure 3.14: Uncertainty area of one configuration at various angles α

relatively fixed with each other, and the third sniffer $O3$ rotates around the position *P* from 0^o to 180^o. The range perturbation δ for the experiment is 10%.

The figures indicate three major results:

1. when the sniffer 2 $O2$ moves closer to the sniffer 1 $O1$ (H goes from 0 to 1), the *uncertainty area* for the position *P* is less affected by the deployment of the third sniffer. Or in other words, the *uncertainty area* remains relative stable when $H = 0.5 \sim 0.9$ and $\alpha = 30^o \sim 150^o$;

2. For all configurations of sniffer 1 $O1$ and 2 $O2$, the deployment of sniffer 3 $O3$ at the closer position to the x-axis gives the largest uncertainty area; or in other words, the linear (or close to linear) deployment of the three sniffers generates the worst localization performance.

3. When the distance between sniffer 1 (*O1*) and sniffer 2 (*O2*) are larger (i.e., $H- > 0$), the deployment of sniffer $O3$ at a position of $\alpha \simeq 60^o$ gives the minimal uncertainty area. However, when the first two sniffers (*O1* and *O2*) get closer (i.e., $H- > 1$), the angle of $\alpha \simeq 30, and 120^o \sim 150^o$ give the better estimation.

Therefore, in order to optimally deploy sniffers for the given positions, the sniffers should not be aligned, instead, they should be *separated* from each others in a triangular style. In reality, in order to minimize the cost of the system deployment (hardware and maintenance), the distance between deployed sniffers are kept larger so as to provide largest signal coverage. From the above simulation in this section, it appears that the deployment of the third sniffer at the angle $\alpha \simeq 60^o$ is of the most important. Considering similarity of the three sniffers and the positions in space, it is tempting to conclude that the deployment of three sniffers at *equilateral triangle* is the best. We will further address this problem in next section.

Sniffers Deployment Method on Given Floor

This section analyzes the sniffers deployment strategy for a given floor plan. We deploy three sniffer in various triangle styles from the right triangle, general acute triangles, and to the special equilateral triangle. At each single deployment, we calculate the overall *average/maximum uncertainty distance* for all points within and outside of the triangle formed by the three sniffers. For the analysis in this section, we assume the range perturbation is 10%.

Figure 3.15 shows a scenario when three sniffers are deployed in a square space of $40 \times 40\ m^2$. For both charts in the figure, the x-axis and y-axis represent the dimension of the space, and the z-axis in the left chart is the *average uncertainty distance* at each particular grid position on the space. The right chart is a contour representation of the left chart, and the number is the uncertainty

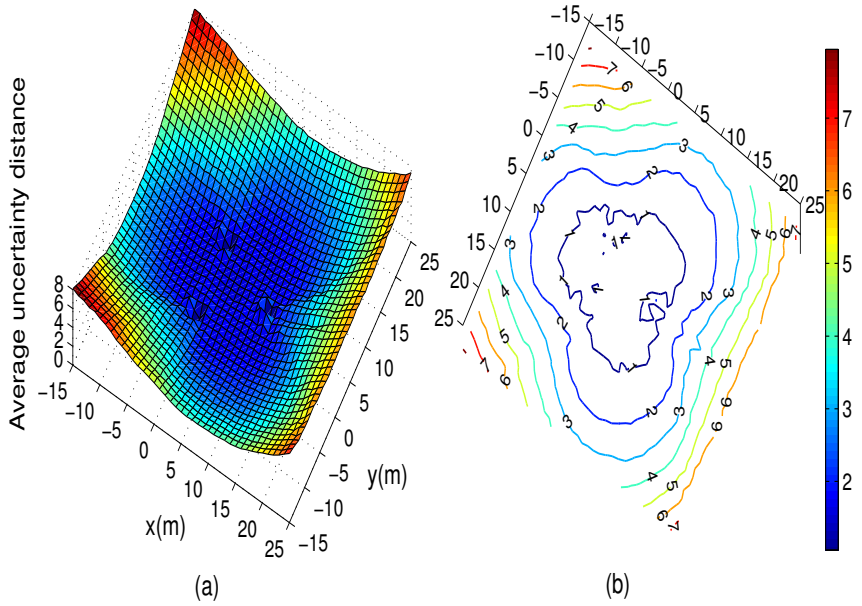

Figure 3.15: Average uncertainty distance at all grid positions in a square area

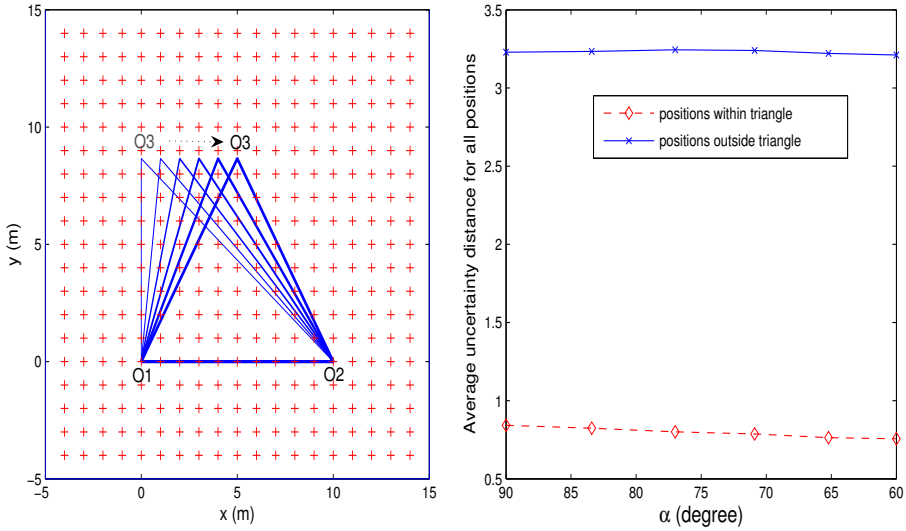

Figure 3.16: Average uncertainty distance of all grid positions within/outside the triangle

distance in meters. It can be seen that the positions within or close to the boundary of the triangle (formed by the three sniffers) obtains the smallest *average uncertainty distance*, i.e., within or around 1 meters in average; and the positions that are farther away from the triangle get larger *uncertainty*, i.e. $2 \sim 7$ meters depends on the distance to the triangle. This indicates that in order to achieve optimal localization performance, a set of sniffers should be deployed in a way to provide maximum coverage for the whole site. Or in other words, any particular interesting positions inside the building should reside within a triangular surface of three neighboring sniffers.

Figure 3.16 compares 6 different sniffers deployment methods where sniffers $O1, O2$ fixed their locations, and the third sniffer $O3$ changes its location and thus the triangles formed by the three sniffers changes from right triangle to equilateral triangle, i.e., $\alpha = 90^o \longrightarrow 60^o$ (Please, see the left chart). At each single deployment, we separate the grid positions for the space into two sets: the first set of the positions are within the triangle $\{P_{in,i}\}, i = 1, 2, \ldots, n$; and the second set is for all positions outside of the triangle $\{P_{out,j}\}, j = 1, 2, \ldots, N$. Then for each set of positions, we sum the *average uncertainty distance* at all positions and make the average distance for the set, i.e., D_{AUD,in_set} and D_{AUD,out_set}, respectively.

Suppose $d_{in,(i,j)}$ and $d_{out,(i,j)}$ $(i \neq j,\ i, j = 1, 2, \ldots, n\ or\ N)$ are the pairwise distance of two points in the two sets, i.e., $\{P_{in}\}$ and $\{P_{out}\}$, respectively, then the **overall** *average uncertainty distance* for the set is defined as:

$$D_{AUD,in\_set} = \frac{\sum(d_{in,(i,j)})}{n} \tag{3.13a}$$

$$D_{AUD,out\_set} = \frac{\sum(d_{out,(i,j)})}{N} \tag{3.13b}$$

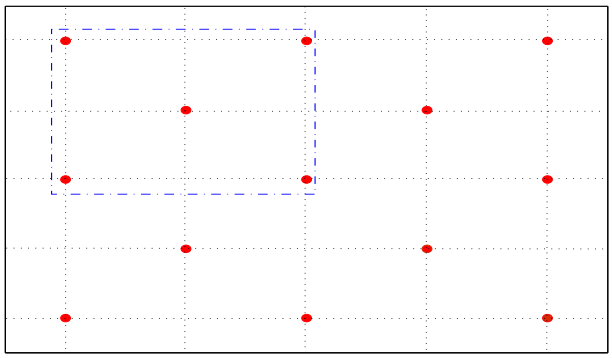

Figure 3.17: Ideal sniffers deployment method

The results are given in the right chart in Figure 3.16, where the x-axis is the degree α as defined in Section 3.4.1, and the y-axis is the mean value of the **overall** *average uncertainty distance* for the two point sets.

The Figure 3.15 and Figure 3.16 indicate two important results:

1. Different acute triangles of the sniffers deployment do NOT make much difference for the location estimation of the points within the triangle, although the deployment of the equilateral triangle ($\alpha = 60^o$) does provide the best localization performance;

2. For a given space, the uncertainty of the localization for positions within & outside of the triangle makes big difference, and the localization for positions that are outside of the triangle is not comparable with the positions within the triangle.

Therefore, in order to optimally estimate the positions within the interested building, it is necessary to include the important locations within the surface of a triangle formed by three sniffers. For a large square floor plan, it is necessary to deploy a set of sniffers (instead of three), and the deployment may take the *mesh* format as shown in Figure 3.17, where a set of sniffers are deployed at ***semi-grid*** positions in the space, and thus most positions are within a triangle of three nearest

sniffers. The mesh structure (formed by five sniffers) within the dash-dotted rectangular is the basic construction block, and only one such block may be required for most (smaller) floor plans in order to provide good location estimation for all positions within the building.

In Section 5.2.1 at Chapter 5, we will further address this problem and calculate the *average uncertainty distance* d_{AVD} and the *maximum uncertainty distance* d_{MUD} for all available deployments at the experimental building. We will show that the *average uncertainty distance* d_{AVD} gives the theoretical location estimation upper bound. Besides the sniffers deployment, indoor localization performance is also dependent on other site specific parameters, like the radio coverage, the partitions, and the indoor occupations.

CHAPTER 4

RESEARCH DETAILS FOR OUTDOOR LOCALIZATION SYSTEMS

This chapter introduces the methods to solve the problems addressed in Section 1.2 in Chapter 1. We consider the radio communication constraints [72] to infer raw estimates of the position of a node or the distance between nodes. And we propose two methods that are built upon two efficient algorithms: lateration algorithm [30, 67] and the multidimensional scaling algorithm (MDS) [84, 85]. We adapt these algorithms to work and yield accurate estimates even with imprecise and/or incomplete information.

In the following sections, we present the research settings and definitions in Section 4.1. Section 4.2 discusses the proposed localization approach. Section 4.4 presents the precision analysis, and Section 4.5 introduces a practical irregular radio propagation model.

4.1 Definitions and Assumptions

We consider a static ad hoc wireless network in a unit square area. We assume that all nodes transmit radio with the same *maximum* range R ($0 \leq R \leq 1$), and that all nodes can measure the distance to their directly connected neighbors. In addition, we assume the network contains a small set of anchor nodes with prior knowledge of their positions (the others are called normal nodes). If there are m anchors for an ad hoc network with total N nodes, the anchor density is defined as $\frac{m}{N}$.

4.1.1 Deployment Randomness

For a network,if all nodes strictly reside on crosspoints of a coordinate grid then the deployment randomness is null. Specifically, the deployment randomness is defined as the maximum

displacement x of each node that can swing from a crosspoint. Starting from a null deployment, each node is displaced from a crosspoint over a random distance in a random direction. The unit of the maximum displacement is the radio transmission range [102]. We associate a random number ρ_i ($\rho \in [0, 1]$) with each node n_i. For example, if the randomness is *x*,then the maximum displacement is $x \cdot R$,and the displacement of node n_i is $\rho_i \cdot x \cdot R$. Similarly, a random number θ_i is assigned for the direction of the displacement.

4.1.2 Estimation Error

For a set of nodes, we denote the estimation error as the average distance between true positions and the estimated positions. The estimation error is expressed as a fraction of the radio range. Let $X_i(i = 1, 2, ..., n)$ be the coordinates of *n* normal nodes where $X_i = (x_{i1}, x_{i2}, x_{i3})^T$. If $\widehat{X_i}$ are the estimated coordinates, then the estimation error is given by the Euclidean distance between X_i and $\widehat{X_i}$:

$$err = \frac{\sum_{i=1}^{n} \sqrt{(X_i - \widehat{X_i})^T (X_i - \widehat{X_i})}}{n \cdot R} \tag{4.1}$$

4.1.3 Fitness Function

Consider a path with *n* nodes that measure their pair-wise distances. The measured pair-wise distance of all nodes over this path is denoted by $\Delta = [\delta_{i,j}, i, j = 1, ..., n; i < j]$, where $\delta_{i,j} = (X_i - X_j)^T (X_i - X_j)$. Based on these measurements, a localization algorithm provides estimates of the positions of the nodes. If the pair-wise distance between the estimates is given by $\widehat{\Delta}$, with $\widehat{\Delta} = [\widehat{\delta_{i,j}}, i, j = 1, ..., n]; i < j)$ and $\widehat{\delta_{i,j}} = (\widehat{X_i} - \widehat{X_j})^T (\widehat{X_i} - \widehat{X_j})$.

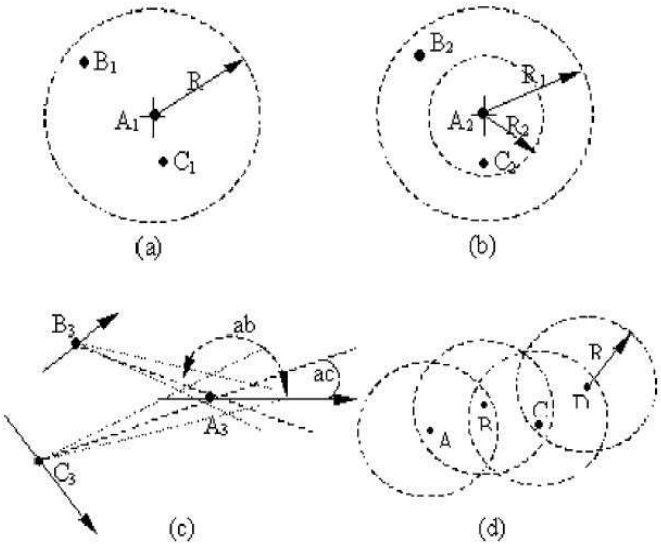

Figure 4.1: Radio communication scenarios

The fitness function is defined as the difference between $\widehat{\Delta}$ and Δ, and it is denoted by σ, where:

$$\sigma = \sum_{i<j} \frac{(\widehat{\delta_{ij}} - \delta_{ij})^2}{\delta_{ij}^2} \tag{4.2}$$

The goal of the localization algorithms is to minimize the fitness function, and to obtain minimum estimation error.

4.1.4 Radio Communication Constraints

Radio communication constraints are a set of geometry rules useful to bound position estimates. Generally, radio communication constraints are a combination of radial and angular constraints [72, 102].

Radial constraint means that if a node B can hear a node A, the distance between them is less than A's radio transmission range. In Figure 4.1(a), since nodes B_1 and C_1 hear node A_1, then the

distance of A_1B_1 and A_1C_1 is less than A_1's radio range R. In Figure 4.1 (b), assume that node A_2 can transmit with two different ranges: R_1 and R_2 ($R_1 > R_2$). If node B_2 hears node A_2 only when A_2 uses its highest power with range R_1, then B_2 is in the circular annulus between R_1 and R_2.

Angular constraint refers to the fact that when a node gets the best reception at a certain direction, it can estimate the relative angle to the source transmitter, which may be a cone bounded by a certain limit [72]. In Figure 4.1 (c), the location of node A_3 can be determined with relative angles of ac and ab. In addition, a small group of neighbors may form a set of triangles with local geometry constraints to further refine estimates [11]. In Figure 4.1(d), nodes *A, B, C* and *D* are neighbors along a path from *A* to *D*. The distance between *A* and *D* is less than *3R*. The distance between *A* and *C*, as well as the distance between *B* and *D* should be less than *2R*. In addition, *B* resides in the radio transmission intersection area of *A* and *C*, and node *C* is within the intersection area of *B* and *D*.

4.2 Augmented Multidimensional Scaling

The proposed localization method is a combination approach of the multidimensional scaling method and the lateration method. Basically, the algorithm works as follows:

1. Using the *DV-Hop* or *Euclidean* radio propagation method, anchors start the estimation process by broadcasting their position information to the whole network.

2. If a path $\mathcal{P}$ contains more than three anchors, the information collected through $\mathcal{P}$ is transmitted back to the original anchors for localization procedure.

3. Apply IT-MDS or SA-MDS algorithm for the nodes along the paths, and update the average hop size iteratively. (The average hop size is to be used by lateration algorithm).

4. Apply the lateration algorithm for the rest nodes (not on paths with three or more anchors).

5. Refine the position estimation using the constraints from neighbors

4.2.1 Classical and iterative MDS

Classical multidimensional scaling (MDS) is a metric MDS technique in 1930s by Young and Housholder [85]. A good reference for MDS is by Cox and Cox in [84].

To apply the MDS algorithm in the estimation of the nodes' position in a wireless network, accurate pair-wise distance among all nodes along the transmission path is necessary. If absolute position is required, each path must contain at least three anchor nodes in order to adjust the relative position accordingly. However, detailed pair-wise distance is not available in most cases. In reality, only the following information maybe achievable: 1) The anchors' positions; 2) 1-hop pair-wise distance; 3) Estimated average hop size; and 4) Hop counts between pairs.

Therefore, iterative approach is proposed to help the location estimation under general localization system with incomplete information. With limited pair-wise distance information, we modify the fitness function σ (equation 4.2) accordingly. Specifically, we assume the pair-wise distance matrix Δ and $\widehat{\Delta}$ include only the 1-hop pair-wise distance between nodes along the path, where Δ represents the measured 1-hop pair-wise distance and $\widehat{\Delta}$ denotes the estimates for each iteration. The reason is that the measured 1-hop pairwise distance is generally more precise than the estimated n-hop ($n > 1$) pair-wise distance.

Based on the neighborhood radio communication constraints and the accurate anchors' position information, the algorithm is able to adjust the pair-wise distance iteratively. At each iteration round, if the fitness function for the estimation is less than the previous step, then the present position

estimation is kept for the next loop until specific requirement is fulfilled. Therefore, the main procedure can be described as follows:

```
while (σ > ε) {
    Run MDS algorithm
    Communication constraints check
    Update the anchors' position
    Update the distance among anchors
    Compute fitness function
}
```

The *communication constraints check* is used to bound the position updates at each MDS procedure within the iteration loop. Based on the imprecise n, ($n > 1$) hop pair-wise distance from hop count, MDS algorithm is likely to yield unrealistic position estimates, therefore, it is necessary to bound or reposition those estimates using the neighborhood radio communications.

In [9], X. Ji *et al.* proposed a similar iterative method in their research. The detailed approach of the two iterative algorithms is different. Specially, the two methods differ as follows: 1) **Estimation coverage**: For a general network without special deployment, MDS based algorithm does not provide full coverage estimation for all nodes in the network. We further detail this issue in section 4.3 in this chapter; 2) **Constraints**: We specially consider the radio communication constraints during each iteration loop; 3) **Fitness function**: We use adapted fitness function without considering all pair-wise distance; 4) **Mechanism**: Our localization procedure does not require local maps to be constructed at all individual normal nodes; 5) **Performance**: X. Ji's iterative algorithm doesn't always provide optimal estimates.

The IT-MDS algorithm works well when the reference anchor nodes are not aligned. In the implementation, these methods may be used: i) if there are more than three anchors in the path, we can select the third reference anchor with the longest distance to the line formed by the starting and the ending nodes; iii) if one node resides on multiple routes and gets estimates more than once, the final position estimation for that node is given as the average of all predictions.

4.2.2 Simulated Annealing MDS Algorithm

Simulated Annealing (SA) [96] is an efficient search method (please, see Section 3.1.4 in Chapter 3 for details). To combine the SA with MDS algorithm for localization process in a wireless ad hoc network, we proceed the optimization procedure at each single node, i.e., for a given state with initial temperature *T*, all nodes in the path are considered one at a time.

Different from the search procedure in the indoor localization from the signal strength map table (SS-MAP), in the outdoor localization procedure, a random optimization movement is given by adding a deviate from the Cauchy distribution to each coordinates *(1,2, ..., p)* of node's previous position X_i in the format as:

$$x_{i+1,k} = x_{i,k} + T \cdot tan(P),\ k = 1, 2, ..., p \tag{4.3}$$

The overall procedure of the SA-MDS algorithm is similar to the search method in Section 3.1.4 in Chapter 3. But the fitness determination in step 5 is now adapted to use the classical MDS algorithm (equation 4.2) instead of the least mean square error (LMS). That is, if a lower fitness function is obtained, keep the displacement from the above step; else, keep the displacement with certain probability.

Accordingly, the main algorithm can be detailed below:

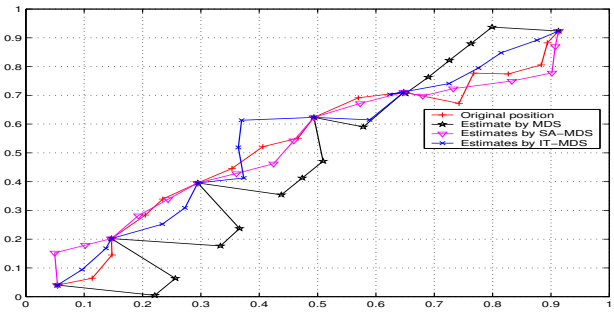

Figure 4.2: SA-MDS and IT-MDS algorithm example

while (not cooling down) {

```
  Random search a node's position
  Run MDS algorithm
  Communication constraints check
  Evaluate estimation errors
  if (errors < previous estimates)
     keep change
  else
     keep change with probability
  end
}
```

A typical SA-MDS example (with IT-MDS) is given in Figure 4.2. The figure shows a path in a network of unit square area. '$\bigtriangledown$' represents the estimated positions by SA-MDS method, and '$\times$' shows the estimates from IT-MDS. The true position is denoted by '+'. The figure shows that the SA-MDS may achieve similar or even better estimates than the IT-MDS method.

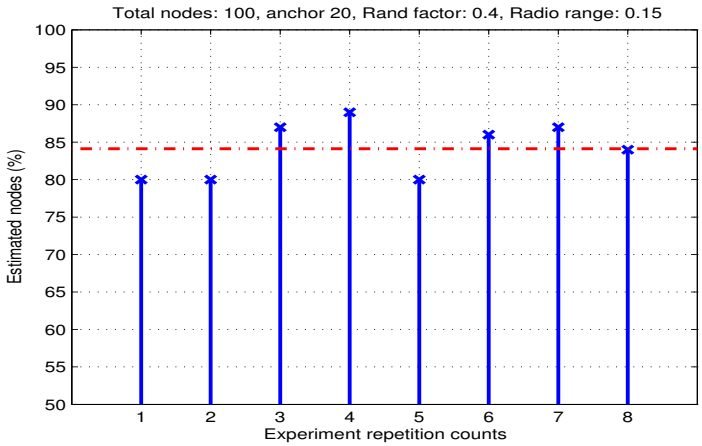

Figure 4.3: MDS estimation coverage analysis

4.3 MDS Processing Coverage Analysis

During the study, we found that MDS based algorithms are able to estimate only the positions for a subset of the nodes in network. The two major facts contribute this *partial* coverage property: 1) path requirements, where the MDS algorithm requires more than three anchors for a routing path, and 2) the *DV-Hop* propagation method only selects the shortest path when routing.

A typical example is given in Figure 4.3, where a set of repetition experiments is carried out for a network of 100 nodes (20 anchors). The x-axis denotes the experiment repetition, and the y-axis is the estimated nodes in percentage. The dotted horizontal line is the average estimation coverage. The results show the MDS algorithm only estimates the position for about 84% nodes in the network. Generally, if we enlarge the radio transmission range, then fewer nodes will get estimated.

This necessities two remedies: 1) Another algorithm to estimate positions for the rest nodes in the network. In this research, we choose lateration algorithm to estimate the positions for the rest

nodes, and 2) Anchor deployment strategy. Similar as the research from Savarese [30], Doherty [72], and Nagpal [34], we manually deploy four anchor nodes at the corner of the testing area.

4.4 Precision Analysis

In this section, we consider the position estimation performance for two basic algorithms: the least squares (LS) and the multidimensional scaling (MDS). We first derive the theoretical estimations for the two algorithms, then we consider a simple range measurement error model, and experimentally outlines the relationship between the localization precision with the range error.

4.4.1 Theoretical Estimates for Multidimensional Scaling Algorithm

Sibson [103] considered the effect of perturbing the distance matrix Δ to a matrix $\Delta(\epsilon)$, such that $\Delta(\epsilon) = \Delta + \epsilon F + O(\epsilon^2)$, where *F* is a symmetric zero-diagonal matrix. Sibson showed that the corresponding induced eigenvalue λ and the Eigen vector *V* will be:

$$\lambda_i(\epsilon) = \lambda_i + \epsilon\mu_i + O(\epsilon^2) \tag{4.4}$$

$$v_i(\epsilon) = v_i + \epsilon f_i + O(\epsilon^2) \tag{4.5}$$

where $\mu_i = -\frac{1}{2}v_i^T F v_i$, $f_i = \frac{1}{2}(B - \lambda_i I)^+ F v_i + \frac{1}{2}(\lambda_i n)^{-1}(e^T F v_i)e$.

The superscript symbol '+' denotes pseudo-inverse, that is for a symmetric matrix *M*, with spectral decomposition $\sum \lambda_k v_k v_k^T$, then M^+ is the matrix $\sum \lambda_k^{-1} v_k v_k^T : \lambda_k \neq 0$.

Consider a simple perturbation model *F* where there is a constant off the diagonal and zero on it: $F = \epsilon(ee^T - I)$. From [103], we can get $\lambda_i(\epsilon) = \lambda_i + \frac{1}{2}\epsilon$.

Substitute *F* to equation 4.5, it is trivial to derive the Eigen vector $v_i(\epsilon) = v_i$ [1]. Consequently, the perturbation on the coordinate *X* will be:

$$x_i(\epsilon) - x_i = (\lambda_i + \frac{1}{2}\epsilon)^{\frac{1}{2}} v_i - \lambda_i^{\frac{1}{2}} v_i \tag{4.6}$$

According to the binomial expansion, we can simplify the coordination perturbation as:

$$\begin{aligned} x_i(\epsilon) - x_i &= (\lambda_i + \frac{1}{2}\epsilon)^{\frac{1}{2}} v_i - \lambda_i^{\frac{1}{2}} v_i \\ &= \lambda^{\frac{1}{2}} v_i + \frac{1}{2}\lambda^{-\frac{1}{2}}(\frac{1}{2}\epsilon)v_i + \cdots + (\frac{1}{2}\epsilon)^{\frac{1}{2}} v_i - \lambda_i^{\frac{1}{2}} v_i \\ &= \frac{1}{2}\lambda^{-\frac{1}{2}}(\frac{1}{2}\epsilon)v_i + \cdots + (\frac{1}{2}\epsilon)^{\frac{1}{2}} v_i \\ &\simeq (\frac{1}{2}\epsilon)^{\frac{1}{2}} v_i + O(\epsilon) \end{aligned} \tag{4.7}$$

Formula 4.7 gives the position displacement of a node under constant distance perturbation. The result shows that the constant distance perturbation will not directly alter the directions (Eigen vectors) of the deployment configuration, however, it does modify the amount of the stretch for those directions (Eigen values). In other words, the difference of the position coordinates is determined only by the variation of the Eigen values.

4.4.2 Theoretical Estimates for Least Square Algorithm

This section derives the formula for the least square method that is addressed at Section 2.3.1 in Chapter 2. We assume that three anchors $A_i,\ i = 1, 2, 3$ are available around the interested position P, and their coordinates are $(x_i, y_i),\ i = 1, 2, 3$, and the distances from the anchors to the position P are $d_i\ , i = 1, 2, 3$, respectively. Consequently, according to the Equations (2.5), (2.6),

[1] Proof: $f_i = \frac{1}{2}(B - \lambda_i I)^+ F v_i + \frac{1}{2}(\lambda_i n)^{-1}(e^T F v_i)e = \frac{1}{2}\epsilon(B - \lambda_i I)^+ v_i + \frac{1}{2}(\lambda_i n)^{-1} e^T(-\epsilon v_i) = 0.$

and (2.7), the linearized equation of $AX = b$ is with A and b as follows:

$$\mathbf{A} = \begin{bmatrix} (x_1 - x_2) & (y_1 - y_2) \\ (x_1 - x_3) & (y_1 - y_3) \end{bmatrix} \tag{4.8}$$

and

$$\mathbf{b} = \frac{1}{2} \cdot \begin{bmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 + d_n^2 - d_1^2 \\ x_1^2 - x_3^2 + y_3^2 - y_3^2 + d_1^2 - d_3^2 \end{bmatrix} \tag{4.9}$$

Thus, the determinant and the inverse of the matrix A is:

$$det(A) = x_1 \cdot (y_2 - y_3) - x_2 \cdot (y_1 - y_3) + x_3 \cdot (y_1 - y_2) \tag{4.10}$$

$$A^{-1} = \frac{1}{det(A)} \begin{bmatrix} (y_1 - y_3) & (y_2 - y_1) \\ (x_3 - x_1) & (x_1 - x_2) \end{bmatrix} \tag{4.11}$$

Consequently, the coordinates of the position P are:

$$\begin{aligned} X &= \frac{1}{2 \cdot det(A)} \begin{bmatrix} (y_1 - y_3) & (y_2 - y_1) \\ (x_3 - x_1) & (x_1 - x_2) \end{bmatrix} \cdot b \\ &= \frac{1}{2 \cdot det(A)} \begin{bmatrix} y_1 \cdot \Delta_1 + y_2 \cdot \Delta_2 + y_3 \cdot \Delta_3 \\ -(x_1 \cdot \Delta_1 + x_2 \cdot \Delta_2 + x_3 \cdot \Delta_3) \end{bmatrix} \end{aligned} \tag{4.12}$$

where $\Delta_i,\ i = 1, 2, 3$ are:

$$\Delta_1 = x_3^2 - x_2^2 + y_3^2 - y_2^2 - (d_3^2 - d_2^2) \tag{4.13a}$$

$$\Delta_2 = x_1^2 - x_3^2 + y_1^2 - y_3^2 - (d_1^2 - d_3^2) \tag{4.13b}$$

$$\Delta_3 = x_2^2 - x_1^2 + y_2^2 - y_1^2 - (d_2^2 - d_1^2) \tag{4.13c}$$

Now suppose there is a perturbation for the distance of $d_i,\ (i = 1, 2, 3)$ between the anchors to the interested position, such that the measured distance is:

$$\widehat{d_i} = d_i + \delta_i,\ i = 1, 2, 3 \tag{4.14}$$

Then a location estimation error exists between the true position X and the estimated position $\widehat{X}$, and the difference between is:

$$\widehat{X} - X = \frac{1}{2 \cdot det(A)} \begin{bmatrix} -(y_1 \cdot \delta_{23} + y_2 \cdot \delta_{13} + y_3 \cdot \delta_{12}) \\ (x_1 \cdot \delta_{23} + x_2 \cdot \delta_{13} + x_3 \cdot \delta_{12}) \end{bmatrix} \tag{4.15}$$

where $\delta_{i,j},\ i, j = 1, 2, 3,\ i \neq j$ are:

$$\delta_{23} = 2d_3\delta_3 + \delta_3^2 - 2d_2\delta_2 - \delta_2^2 \tag{4.16a}$$

$$\delta_{13} = 2d_1\delta_1 + \delta_1^2 - 2d_3\delta_3 - \delta_3^2 \tag{4.16b}$$

$$\delta_{12} = 2d_2\delta_2 + \delta_2^2 - 2d_1\delta_1 - \delta_1^2 \tag{4.16c}$$

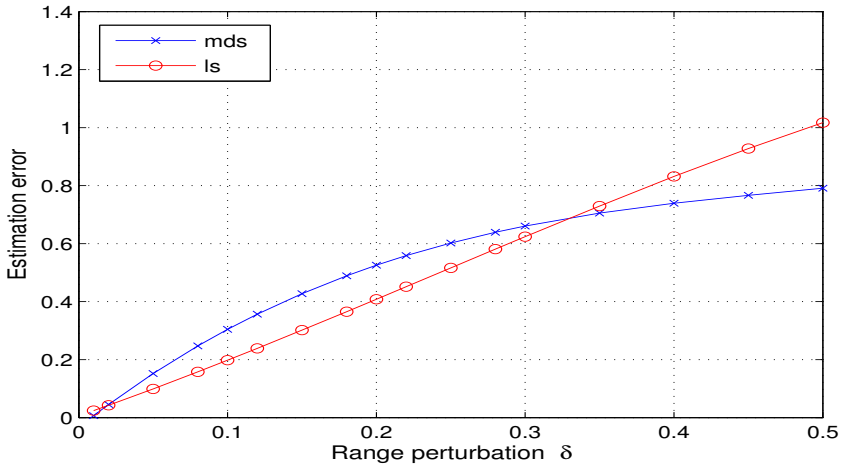

Figure 4.4: Theoretical estimation errors under various perturbations

4.4.3 Algorithm Precision at Various Perturbations

With the analysis from Section 4.4.1 and section 4.4.2, we calculate the theoretical estimation errors of two algorithms at various range perturbations. The experiment is carried as follows:

1. Randomly generate a configuration of three anchors and one normal position in an unit square area;

2. Assume the perturbation of the range r is δ, thus the actual range measurement is given as $r = r \cdot (1 \pm rand \cdot \delta)$;

3. Apply the MDS and LS algorithm to estimate the location of the normal node.

A comparison of the localization estimation is given in Figure 4.4, where the x-axis is the range perturbation δ, and the y-axis is the relative estimation error in terms of radio range R. In the figure, the line with circle denotes the estimation of least squares (LS), and the line with '$\times$' represents the estimation of MDS. The radio range R in this experiment is set to 0.4, and the results show in the

figure is the estimation with 95% confidence for over 5,000 simulation runs. The figure indicates that the LS algorithm achieves better performance when the error of the range measurement is small; however, when the error of the range measurement becomes larger, the MDS algorithm is more precise.

In Chapter 6, we will reiterate this problem together with actual simulations. By outlining the average estimation error at all anchors' configuration, we tentatively build the relation between the location estimation error and the perturbation for both algorithms, i.e., MDS and Least Square.

4.5 Radio Range Irregularity Research

The research on Radio Range Irregularity (RRI) is based on experiments. The experiments were carried out on various environments using IEEE 802.11 wireless Ethernet standard, and it was found that the radio transmission range at 2.4 GHz UHF (ultra high frequency) band may not smoothly vary in different directions depending on the presence of obstacles (ex. trees and stones) in vicinity. This observation is significantly different from the conclusion by Zhou *et al.* [39]. Based on this result, this dissertation develops a realistic RRI model that takes into account, for the first time, actual effects from various operating environments and heterogeneous properties of the wireless devices (transmission with different powers).

Based on the new RRI model, this dissertation analyzes the performance of a set of representative localization algorithms that were initially based on the ideal radio model. These algorithms include two multidimensional scaling algorithms [77] , bounding box algorithm [8], and lateration algorithm [30]. Through simulation, it is found that the realistic RRI model greatly degrades the performance of all algorithms. Specially, the effects can be identified in the following ways:

1. Radio propagation method: the maximum-distance *greedy* forwarding approach is not suitable anymore [38];

2. Radio coverage pattern: the radio transmission coverage is not circular, consequently the neighborhood density is different from (or less than) the ideal counterpart [38];

3. Pair-wise hop distance estimation: under realistic environment, there exists no reliable average hop distance between communicating nodes if no other mechanisms is available [77, 7];

4. Refinement procedure: it is not appropriate to apply the same the refinement procedure with information from bidirectional connected neighbors [8, 76].

To remedy the adverse effect of the radio irregular propagation on the localization performance, this dissertation proposes and evaluates an **optimized constrained-greedy forwarding** radio propagation method. Test results indicate that better location estimation is achievable.

The rest of the introduction is organized as follows: In Section 4.5.1, we present the analysis for the radio irregularity and the measurement results. Section 4.5.2 introduces the radio irregularity model, and Section 4.5.3 provides an optimized radio propagation method for the localization algorithms.

4.5.1 Radio Range Analysis and Measurement

The irregularity of the radio propagation is mainly caused by the multipath environment, and the power heterogeneity of the transmission devices. To conduct experiments, two laptops with IEEE 802.11b wireless interface were connected in ad hoc mode to measure the signal strength. One of the laptops is a Compaq Pavilion v2000 with Intel Pro 2100b card, and the other is an HP Pavilion ze4900 with D-link DWL-G630 C card. We consider three settings in this dissertation:

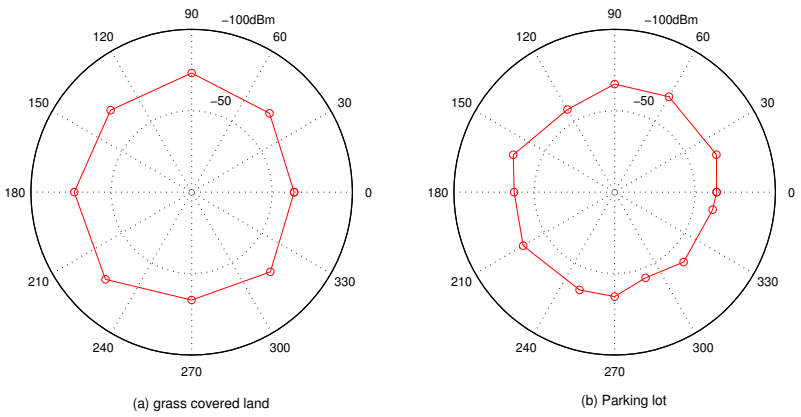

Figure 4.5: Signal strength in clear space

1. Forest land with scattered trees of various diameter on a clear ground;

2. Grass-covered land with grass height around 20 cm;

3. Flat and clear parking lot.

The measured signal strength in grass covered land and the parking lot are plotted in Figure 4.5. The left chart (figure (a)) is the measurement from grass covered land, where the receiver is 30 meters away from the transmitter; and the right chart (figure (b)) denotes the measurement from a parking lot, and the distance between the transmitter and the receiver is around 50 meters. In both charts, the center denotes the position of the transmitter, and the measured signal strength in *-dBm* is denoted by the circular ring around it. It can be seen that the measured signal strength at different directions at fixed distance to the transmitter is approximately circular.

Similarly, the signal strength measurements in two different forest-lands are plotted in Figure 4.6. Figure 4.6(a) depicts the measurements in a wood with the tree density[2] of roughly 20, the

[2]Tree density is determined by the number of trees N in an area of 100 sq. meters.

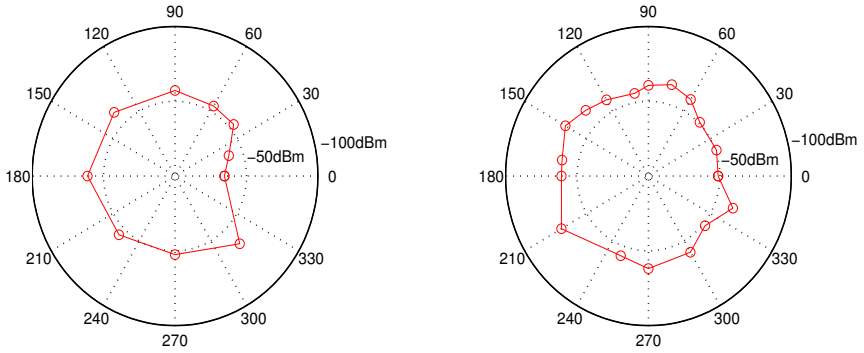

Figure 4.6: Signal strength in forest land

distance between transmitter and the receiver is about 20 meters. Figure 4.6(b) plots the results of a wood of 35 tree density, and the distance is 35m. It is clear that the existence of trees heavily affects the received signal strength.

To estimate the effect of a tree on the received signal strength, a simple experiment (as depicted on the left chart of Figure 4.7) was carried out. The tree is approximately 60 cm in diameter, and is roughly 25 meters away from the transmitter. The measurement positions are about 1.0 meter beyond the tree, and three positions around the ray sector (denoted by dotted lines on left chart) are considered. The measurement values are given on the right. The x-axis denotes the three measurement positions, and the y-axis is the signal strength in -dBm. The results indicate that the tree strongly attenuates the signal power level in the vicinity of the *hidden* part behind the tree, and the signal strength does not smoothly vary.

A similar experiment that considers a rock (instead of a tree) was performed. The rock is about 10 meter away from the transmitter, and the receiver is 1.0 meter after the stone (left chart of Figure 4.8). The measurement results are given in the right chart of Figure 4.8. The x-axis denotes the measurement positions, and they are roughly 3 degree apart; and the y-axis is the signal strength in

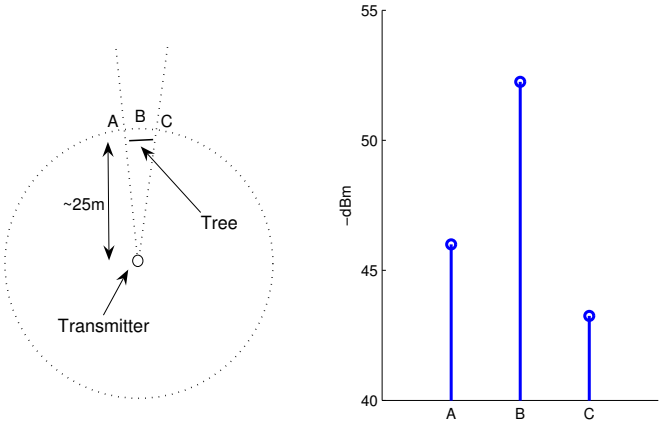

Figure 4.7: Signal strength near a tree

-*dBm*. This chart clearly indicates that the stone greatly affects the received signal strength level, and the signal strength does not change continuously.

Finally, we measure the signal strength across a dense bush, where the transmitter and the receiver is about 20 meters apart, and the bush is roughly 1 meters before the receiver. The signal power is given in Figure 4.9. The x-axis is the measurement positions in degrees to the first measurement location; and y-axis is the signal strength in *dBm*. It can be seen that the bush affects the signal power level irregularly.

In conclusion, the radio propagation at 2.4GHz has the following properties:

1. The radio transmission range is environment dependent. It is roughly circular on a clear space with hard and flat ground (parking lot).

2. The maximum radio range is determined by the transmission power at the transmitter;

3. In field deployment, the radio transmission is not regular; and the signal strength variation at different directions is not smooth around obstacles.

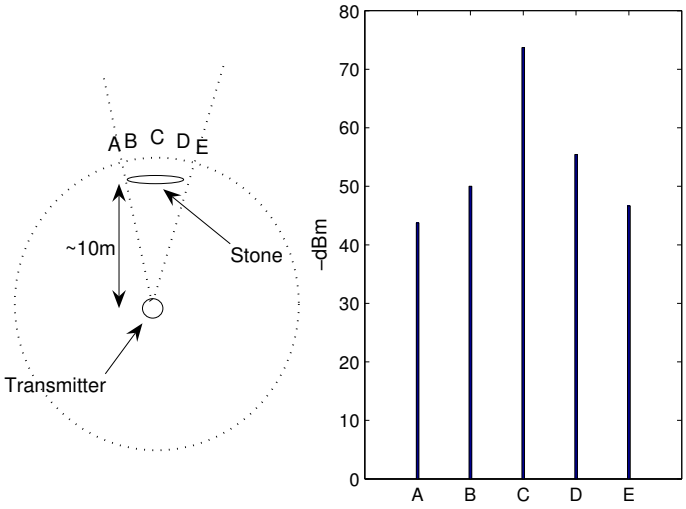

Figure 4.8: Signal strength near a stone

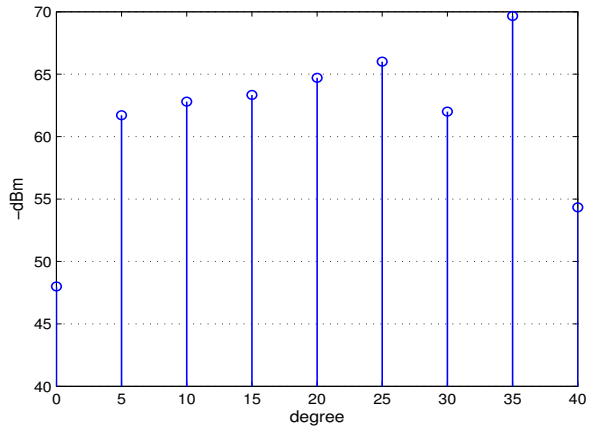

Figure 4.9: Signal strength cross bush

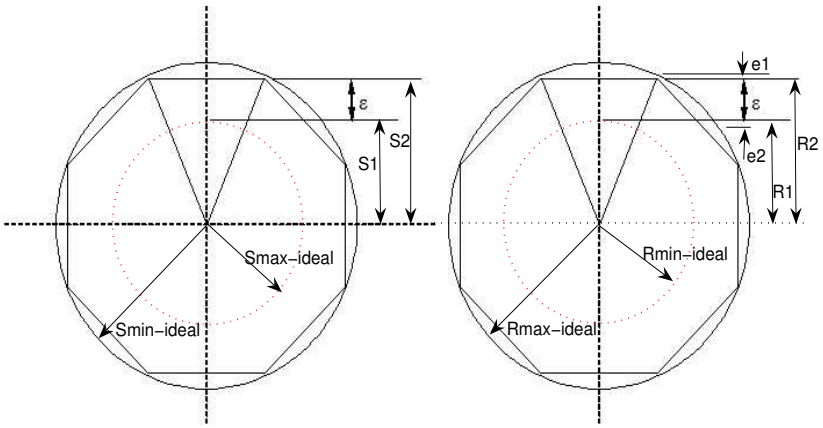

Figure 4.10: Radio irregularity model

4. Depending on the obstacles in each direction around the transmitter, the radio coverage area may be approximated as a collection of propagation sectors starting from that transmitter. Each sector experiences a specific attenuation in that direction.

4.5.2 Radio Range Irregular Model

Based on the findings from experiment, we propose to use a set of triangular sectors to emulate the radio propagation in various directions. Within ONE sector, we assume the ideal radio transmission: the signal strength level is modeled as a set of arcs with decreasing power towards the periphery.

To model the radio propagation with diverse obstacles, the ideal circular radio pattern is sliced into multiple sectors with different signal strength levels. An example is given in Figure 4.10(a), where the radio power at a given distance is assumed to be within the range between $S_{max-ideal}$ and $S_{min-ideal}$. To determine the actual radio signal strength, the ideal circular radio shape is

modeled as a collection of sectors where each one corresponds to a particular attenuation in that direction. The number of sectors is environment dependent (an octagon is used in this example). If the minimum and maximum signal strength values from the polygon are $S1$ and $S2$, approximation of the signal strength values corresponding to $S_{max-ideal}$ and $S_{min-ideal}$ respectively, then at each sector, the signal strength level could be any value between $S1$ and $S2$ as shown in Figure 4.10.

For the localization algorithms, the relationship between the radio range and the received signal strength is exploited. From the signal strength model in Figure 4.10(a), we know that the radio in one sector has only one signal strength level and the radio in this sector transmits across the same media. If we assume that the radio transmission range is only related with the signal strength for a given environment, the radio propagation range on each arc (within a given sector) should be equal. Thus, the radio coverage is similar as the layout of the signal strength distribution.

Consequently, the radio range irregularity model is depicted in Figure 4.10(b): the ideal max and minimum radio range are $R_{max-ideal}$ and $R_{min-ideal}$, and they correspond to the signal strength level of $S_{min-ideal}$ and $S_{max-ideal}$ respectively. Using the same sector style, the approximate radio range is between $R1$ and $R2$.

Therefore, to analyze wireless protocols and algorithms, the radio irregularity model can be constructed as follows:

1. Determine the operating environment, and configure an appropriate polygon that emulates the radio transmission. Generally, for a flat and clear space with no or few obstacles, approximated circular transmission is expected, thus close to circular range model is used. On the contrary, for a field with trees and rocks, a set of sectors according to the obstacle in each particular direction will provide similar simulation settings;

2. Determine the maximum radio transmission range which is related with the transmission power;

3. Determine the maximum range attenuation ε (Please, see Figure 4.10) according to the possible obstacles around the transmission node;

Let us denote the ideal radio transmission range $R_{max-ideal}$, the number of edge of the polygon N, and the maximum range attenuation ε. The actual transmission range R at each sector is given by:

$$R = R_{max-ideal} \cdot sin(\pi(0.5 - \frac{1}{N})) + e1 - rand \cdot \varepsilon - e2 \tag{4.17}$$

where $e1$ and $e2$ are two random numbers with values between the ideal circular transmission range and the approximated polygon at maximum and minimum borders respectively, i.e., $e1 = rand \cdot (R_{max-ideal} - R2)$ and $e2 = rand \cdot (R_{min-ideal} - R1)$.

We analyze, in the next section, the effect of the radio range irregularity model on localization algorithms.

4.5.3 Effect of RRI and Optimization

Most localization algorithms aim to determine the physical location information of nodes in the wireless network, where a limited number of nodes (called anchors) know their absolute positions. The determination of location information for nodes other than anchors requires the distance between nodes, especially the distance from normal nodes to anchors. To determine the distance information, four representative radio propagation methods have been developed (please, see Section 2.3.1 in Chapter 2): 1) *DV-Hop* method [11]; 2) *DV-Distance* method [11, 8]; 3) *Euclidean*

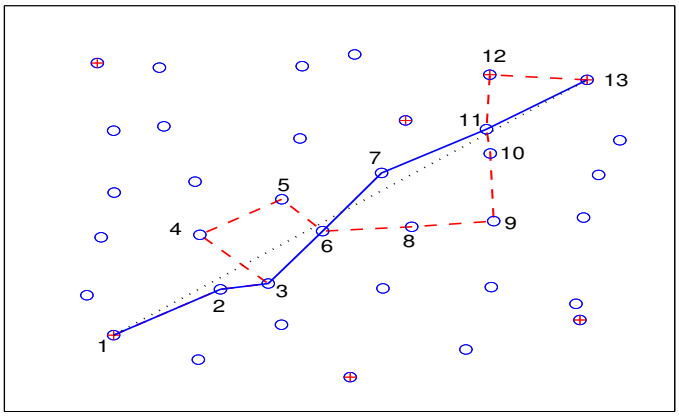

Figure 4.11: Routing path under regular and irregular radio transmission

method [11]; and 4) *Kleinrock's* formula [79]. In the next two sections, we briefly analyze the above radio propagation methods in the irregular radio networks.

DV-Hop method for Irregular radio networks

DV-Hop method allows each node, during message propagation, to only maintain and rebroadcast packets with the smallest number of hop counts to original sources. Therefore the final routing path is the one with minimal transmission hops (maximum distance greedy forwarding). For the ideal circular radio propagation, this method usually generates regular routing path among all nodes. However, under realistic operating environment, the radio transmission is irregular at all nodes, the routing paths between communicating nodes are also irregular. Therefore, a network-wide average hop-size is not reliable as the ideal model counterpart. Consequently, the distance estimation between normal nodes and anchors contains much larger errors, and thus the localization performance is greatly affected. We explain these using a real routing path in Figure 4.11.

Figure 4.11 shows a network of 36 nodes randomly deployed in a unit area. There is a communication starting from node 1 to node 13, and the figure gives two routing path: the solid line show the ideal path under regular radio transmission, and another path in dotted line under irregular radio situation. Additionally, the figure shows a direct line (dotted line) between node 1 and node 13, which is the ideal one-hop direct communication.

Two very interesting routing positions are at node 3 and node 6. For the *DV*-based radio propagation method, each node only maintains and rebroadcasts packets with the *smallest* number of hop counts. This way, the final routing path is the one with minimal transmission hops, and thus it is *maximum-distance greedy forwarding*. As a result, the routing path is fairly close to the direct dotted line between communicating nodes. Therefore, in the ideal situation, the total hop counts between node 1 and node 13 is six.

With irregular radio transmission shape, the *DV*-based method may still work, but a different propagation path is selected this time. Because node 6 is not the directed neighbor of node 3, in order to find a path from node 3 to the destination node 13, the information packets from node 3 have to route back to node 4 (which is not the neighbor of node 2). Similarly, at node 6, the next routing intermediate node is node 8. Consequently a routing path is constructed irregularly, and total hop counts between node 1 and node 13 is now eleven, almost as double as the regular counterpart.

Comparing the two routing paths between two communication nodes, the average hop size estimation is very different. As it can be seen from Figure 4.11, the value of the estimated average hop distance for the irregular radio transmission network is much smaller. Although relative good distance estimation may still be obtained for particular routing path, due to the irregularity nature of various propagation paths, a network-wide average hop size value is not available. Consequently, the

pairwise distance estimation at irregular radio transmission networks contains larger errors, which will finally affect the location estimation performance of all localization algorithms.

Other methods

Similar to the *DV-Hop* radio propagation method, the distance estimation performance for the other three methods also degrades.

DV-Distance method is sensitive to the measurement error, and it works just like *DV-Hop*: it generates irregular routing path, and consequently it gets unreliable distance estimates.

Euclidean method assumes dense network environment, which will be affected by the irregular radio transmission. Additionally, the communication link is not bi-directional, thus the local geometry information may even not be obtainable. This means that the *Euclidean* method may not work correctly for some nodes.

We mainly consider the *DV-Hop* method in this research. Note however, the proposed methodology would equally work with the other three methods. And in the next section, we discuss the corresponding remedy methods on these negative impacts.

4.5.4 Constrained-Greedy Forwarding Radio Propagation Method

In order to obtain reliable distance estimates, it is critical to select a regular routing path for the communicating nodes. Accordingly, the key point turns out to identify best candidate nodes that are used to forward packets to the destination. We develop a *constrained-greedy forwarding* approach in this research. The idea is to constrain intermediate nodes within an optimal range, i.e. only the nodes within a predetermined range are selected to forward packets. This way, the possible routing deviation or irregularities are prevented.

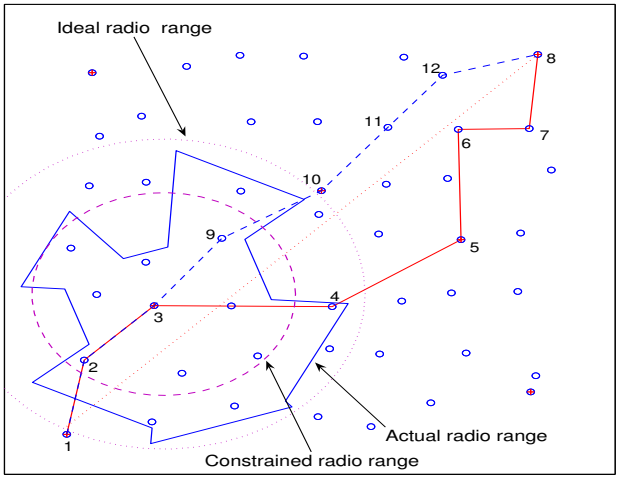

Figure 4.12: Routing optimization under irregular radio transmission

An example is given in Figure 4.12. The figure depicts a network with 49 nodes in a unit area. Two different routing paths are given between node 1 and node 8. The first path *1-2-3-4-5-6-7-8* denotes the routing by the DV-Hop method, and the second path *1-2-3-9-10-11-12-8* represents the routing by the constrained-greedy forwarding method.

A critical routing node is node 3, where two very different forwarding nodes (4, 9) are to be selected. Both candidates are within the radio range of node 3, the difference is that node 9 is close to the center within the predetermined constrained range. It is interesting to see that both paths contain seven hops, and it is clear that the second path is more regular than the first one.

In this dissertation, the constrained radio range is determined as the average value of all radio ranges in different directions. If the operating environment is a clear space, and the radio is close to circular, then the constrained radio range is the actual transmission circle. On the contrary, if the environment is complex, the signal strength along some directions may be heavily attenuated, and thus the transmission range in those sectors is affected. By average the radio range at all directions, a small constrained radio range is obtained. We let those nodes that are within this constrained radio

range to be the candidates for the information forwarding. Note however, the selected candidates should also within the actual radio range, which is environment dependent.

In reality, when a node wants to transmit a packet, it measures and incorporates its transmission power in the packet header. Upon receiving the packet, the neighboring nodes measure the associated signal strength, and compare it with the original transmission power from the header. If the received signal strength at a node is above a threshold, then the intermediate node assumes that it is within the constrained radio range, thus it forwards the packet out if the hop count of the packet is also minimal.

CHAPTER 5

APPLIED RESULTS FOR INDOOR LOCALIZATION SYSTEMS

Extensive experiments have been carried out at two very different buildings to evaluate the ARIANDE system. We present the first experiment results in Section 5.1, and then Section 5.2 introduces the comprehensive experimental results in another building at Auburn University.

The experiment for the first building mainly focuses on the overall performance of the ARIADNE system, and the experiment in the second building is mainly interested in the sniffer's position configuration. We start from the experimental setup for both experiments

5.1 Experiment I

5.1.1 Experiment Setup

Figure 5.1 shows the floor plan of the building used for this study. Three sniffers A, B, and C are deployed inside the building with sniffers A and C deployed close to west and east boundaries respectively. Sniffer B is slightly south of the center of the building. Each sniffer was implemented on an IBM T30 ThinkPad running RedHat 9 operating system. Through the AP, sniffers connected with the global monitor, which is used to process the signal strength data. The global monitor also stores the signal strength tables and estimates the users' current location based on the signal strength readings by sniffers A, B, and C.

In order to validate the proposed indoor radio propagation model, the signal strengths were collected at 30 different locations. A Toshiba laptop with Linksys WAP 11 wireless card was used for data collection. At each location, about 100 sample packets were emitted at an interval of 0.5 seconds and measured at sniffers A, B, and C. The positions of data collection are marked in

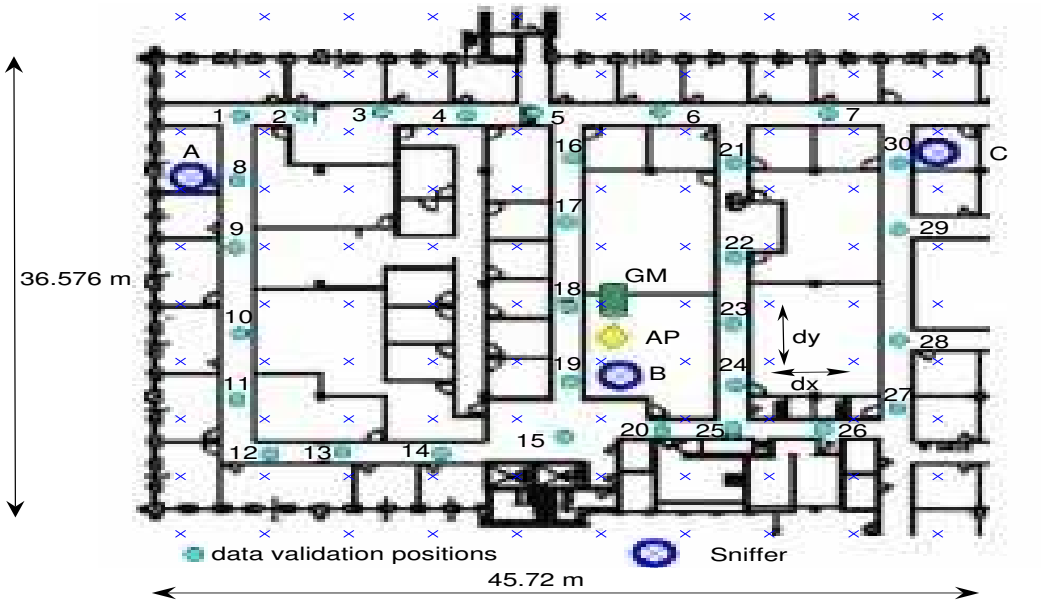

Figure 5.1: Floor plan with sniffers and data validation positions ('×': reference positions in *SS-MAP* table)

Figure 5.1 with faint dots as marked from 1 to 30. The '×' denotes grid positions in estimated *SS-MAP* table.

The same series of measurements at the 30 locations were repeated on 6 different days noted in the following as Day_i.ARIADNE radio propagation model (see Section 3.1.3) is evaluated using the 30 signal strength measurement triplets collected in the building for all 6 days. Note that ONE of the 30 signal strength triplets is randomly selected to estimate site specific parameters in the ARIADNE radio propagation model and then to compute an estimated *SS-MAP* table for a specific set of grid positions. Given the estimated *SS-MAP* and the measured signal strength triplet $M(SA, SB, SC)(L, Now)$ for a mobile at some location L, different search localization algorithms were evaluated.

Table 5.1: Measured signal strength indicator variability over days

	Day1-2	Day1-3	Day1-4	Day1-5	Day1-6
Max difference	5.48	6.77	6.30	7.18	7.66
Average difference	2.21	2.79	2.46	3.02	2.81
MSE	0.48	0.58	0.55	0.65	0.63

5.1.2 Measurements

The signal strength was collected on six different days at the data collection positions indicated in Figure 5.1. Taking Day_1 as reference, Table 5.1 reports the variability of the signal strength from day to day: each column Day_{1-i} displays the variability between Day_1 and Day_i. The variability is captured using the maximum and average difference, and the mean square error MSE respectively defined as:

$$\begin{cases} \text{max diff} & = max_{i=1}^{n}\{abs(SS_{Day_j,i} - SS_{Day_k,i})\} \\ \text{average} & = mean_{i=1}^{n}\{abs(SS_{Day_j,i} - SS_{Day_k,i})\} \\ MSE & = \frac{1}{n}(\sum_{i=1}^{n}(SS_{Day_j,i} - SS_{Day_k,i})^2)^{\frac{1}{2}} \end{cases} \tag{5.1}$$

where $SS_{Day_j,i}$ and $SS_{Day_k,i}$ represent the signal strength measurements at location i on j^{th} day Day_j and k^{th} day Day_k respectively; n is the number of SS triplets ($n = 30$).

Table 5.1 illustrates the dynamic nature of the indoor signal strength over time. Such a variability shows that any search localization method on a static signal strength map *SS-MAP* will in general perform poorly.

5.1.3 Radio Propagation Model Validation

The parameters $[P_0, \gamma, \alpha]$ of ARIADNE radio propagation model are estimated for a given day Day_i using ONE randomly selected signal strength measurement triplet among the 30 triplets from day Day_i. A signal strength measurement triplet is *randomly selected* because a potential

user of ARIADNE may take the reference measurement anywhere in the building. ARIADNE radio propagation model is then evaluated using the 30 signal strength measurement triplets from the same day Day_i. The mean square error MSE between estimates and measurements is used to evaluate the model fitness.

The influence of the maximum number of reflections and the maximum number of traversed walls on the accuracy of ARIADNE radio propagation model was investigated:

1. As the *number of reflections* increases, the attenuation of the signal increases. After some number of reflections, the contribution of power becomes negligible. Based on the simulations, taking into account more than 3 reflections induces heavy computations without any improvement of the precision or accuracy. This conclusion concurs with other researchers (Valenzuela, Fortune, and Ling[104]), therefore, the maximum number of reflections is set to 3;

2. A ray path will traverse at most a limited number of walls. As the ray traverses walls on a direct path, it weakens in power. We call the *maximum number of transmission walls* MW the number of walls traversed before the ray "dies". From simulations in the building considered here, no improvement in precision or accuracy is achieved for MW over 20. Results are reported here for MW taking values 15 and 20 for comparison.

5.1.4 Simulation Results

Extensive simulations were carried out, and the average results of all simulation runs are compared against the 30 signal strength measurement triplets. For each test run, ONE signal strength measurement triplet is randomly selected as a reference among the 30 measurement triplets. Good agreement is obtained between the estimated signal strength map and the measured one.

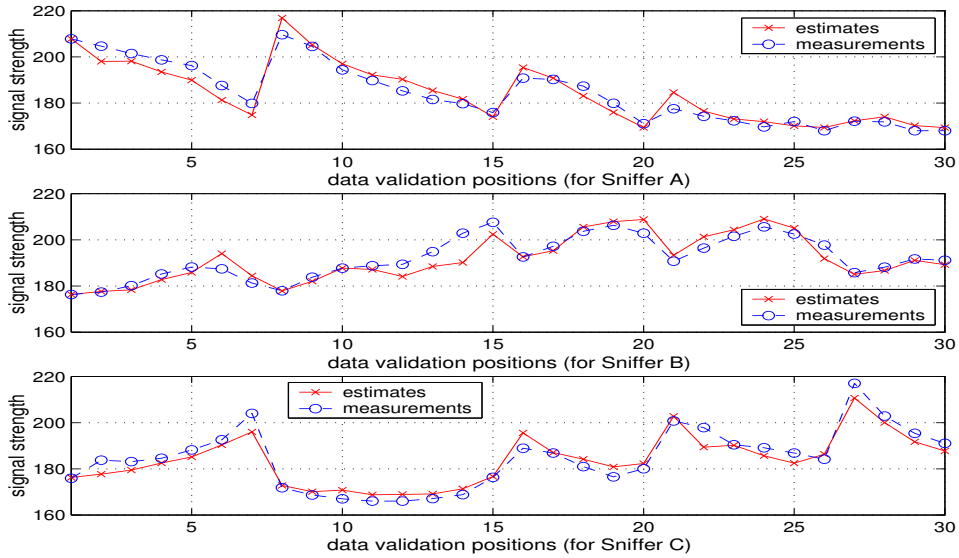

Figure 5.2: Estimation and comparison with measurements at data validation positions

Typical comparison results are shown in Figure 5.2 that consists of three plots respectively for sniffers A, B, and C. For each plot, the x-axis represents the 30 positions from the data collection and the y-axis denotes the signal strength measured as received signal strength indicator (RSSI). The maximum number of reflections is 2 and the maximum number of transmission walls transmission is 20. The points with symbol '$\ominus$' are the signal strength measurements, and the points with symbol '$\times$' are the estimates.

Table 5.2 reports for Day_1 the difference between estimates and measurements and illustrates the impact on ARIADNE accuracy of the maximum number of reflections (2 or 3) and the maximum number of transmission walls (15 or 20). Each simulation run uses as reference *one* signal strength measurement triplet *randomly* selected within the 30 data collection positions from Day_1. Each number reported in the table is the result averaged over more than 20 simulation runs.

Table 5.2: Radio propagation model verification, maximum RSSI=255

Estimation vs. Measurement		Sniffer A		Sniffer B		Sniffer C	
		Wall:15	Wall:20	Wall:15	Wall:20	Wall:15	Wall:20
2 reflections	Max difference	8.2913	8.3820	15.3067	15.3385	9.4413	7.9786
	Average difference	3.3916	3.3487	5.3976	6.3310	3.9688	3.1910
	MSE	0.7472	0.7472	1.2309	1.3504	0.8525	0.6976
3 reflections	Max difference	8.4514	8.8607	15.4559	13.9691	8.7444	8.2063
	Average difference	3.5720	3.5842	6.0425	6.1516	3.5482	3.1853
	MSE	0.7723	0.7711	1.3343	1.2967	0.7710	0.6886

Table 5.2 shows that results are quite similar for 2 and 3 reflections. This shows that higher order reflection rays marginally affect power estimation accuracy. This conclusion agrees with Valenzuela *et al.*'s results [104]. Consequently in the following test runs, rays are restricted to at most of two reflections. Table 5.2 also illustrates that the maximum number of transmission walls of 15 or 20 yield close results.

However, Table 5.2 suggests an interesting observation: estimates for sniffer B are not as good as estimates for sniffers A and C. The difference is systematic and significant. It was observed that the wireless card at sniffer B always provides lower readings than those of sniffer A and C. Despite this, the search localization on the estimated signal strength map *SS-MAP* still works well because the received signal strength from a mobile user at sniffer B is always comparable with B's former readings. So no calibration was done to produce measurement data used in this work.

Finally, Table 5.2 indicates that ARIADNE radio propagation model yields good estimates (the *maximum* signal strength difference is within $3\% \sim 5\%$ of the maximum RSSI, see Figure 5.2): It is worth noting that the variability over days (see Table 5.1) is quite close to the difference between estimates and measurements.

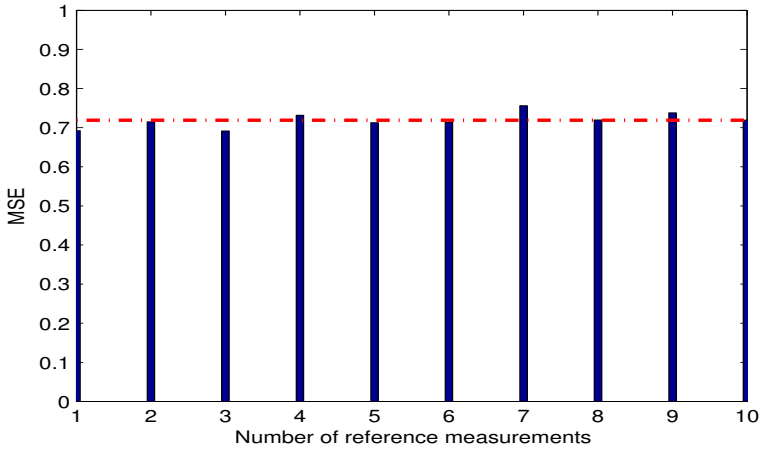

Figure 5.3: MSE vs. number of reference measurements

5.1.5 Number of Necessary Reference Measurements And Location Dependency

In Section 3.1.4, the simulated annealing searching algorithm is used to estimate the parameter tride of $[P_0, \gamma, \alpha]$ using one reference signal strength measurement triplet. This section addresses the question whether multiple reference measurement triplets would yield estimates that are closer to measurements. The answer is surprising: one reference measurement triplet will yield estimates as good as estimates from 2, 3, or 10 reference measurement triplets.

Figure 5.3 confirms the findings. The x-axis denotes the number of reference signal strength measurement triplets used for estimating the parameters of the ARIADNE radio propagation model, and the y-axis represents the mean square error MSE of signal strength between measurements and estimates. For each run of x references, x references are *randomly* selected to be used to estimate the radio propagation model and construct the signal strength map. Each point on Figure 5.3 is an average over 20 runs.

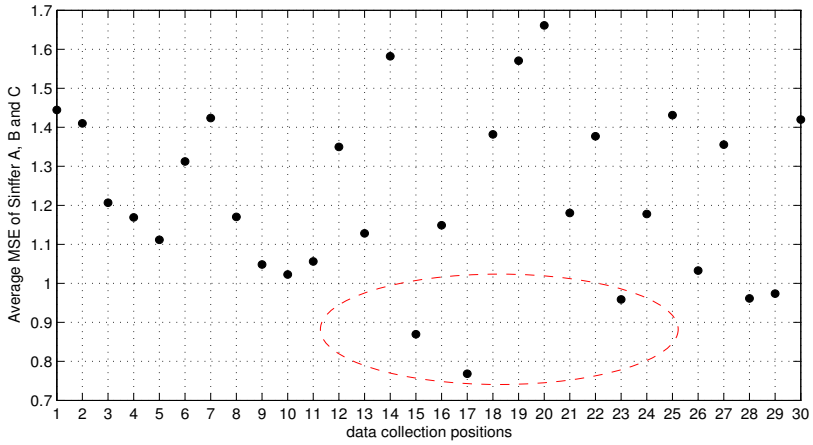

Figure 5.4: Reference Measurement Selection

To evaluate the impact of the location of the reference measurement on the performance of the signal strength estimates, the signal strength map is constructed using each individual 30 collected signal strength measurements. Figure 5.4 provides the average minimum squared error (MSE) for sniffers A, B, and C when using as reference measurement one of the 30 collected signal strength measurements. The x-axis is the location number where the signal strength measurement was made. The y-axis is the average MSE over all sniffers. Results point out that lowest MSE is obtained with signal strength measurements 15, 17, and 23. These measurements were made, as shown in Figure 5.1, close to the center of the building. This appears to suggest that the reference measurement should be made at the center of gravity of the sniffers. We'll reiterate this problem in the next experiment where various sniffers deployment configuration are considered.

Note however, for a building with non-uniform walls (i.e. different construction materials and thickness), or with spatially different (human) occupation, it is not appropriate to only select a single measurement at a reference position. In this situation, it is necessary to first identify all representative regions for the building such that each region has homogeneous construction material

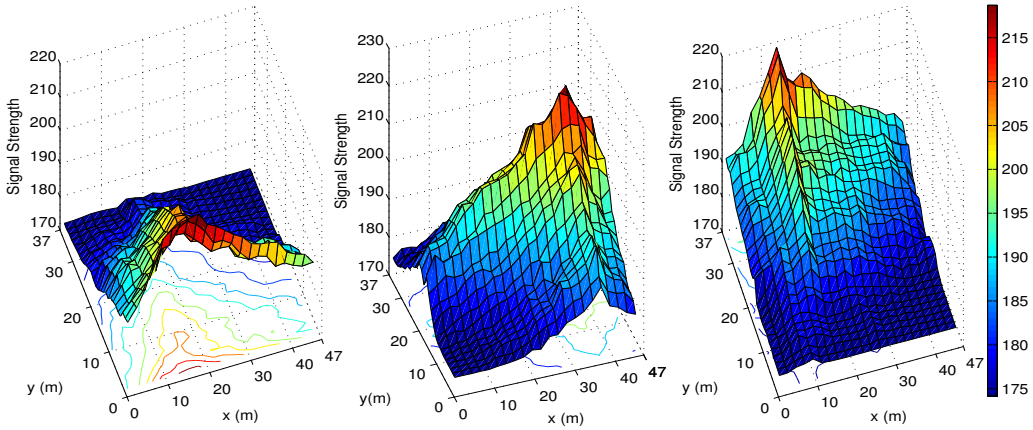

Figure 5.5: Signal strength shaded surface at sniffers A, B, and C

and in uniform space. Thus a regional SS-MAP can be constructed in advance, and the overall building based SS-MAP can then be stitched.

It is worth noting that ARIADNE radio propagation model is quite accurate even though the sniffers are not deployed in an optimal fashion. The next section addresses the poor deployment of sniffers through a coverage analysis.

5.1.6 Coverage Analysis

The study in this first experiment does not have access to the sniffers and thus cannot modify sniffers' deployment. The work presented in this dissertation just exploits a data set collected in the work [63]. A gross coverage analysis shows that the sniffers are not optimally deployed.

Figure 5.5 provides a coverage map for each sniffer (A, B, and C). In each 3D plot, the xy plan is the building floor and the z-axis is the estimated signal strength (RSSI). As the reader can observe, a considerable portion of the area is dark for each sniffer. These dark areas are quite flat as

Table 5.3: Localization performance of six experimental measurements

	Clustering		LMSE		2-N		3-N	
	err	std	err	std	err	std	err	std
Day1	2.8372	2.4304	2.7442	2.0349	3.7355	2.9256	3.5412	3.0458
Day2	2.5330	2.2388	3.5297	2.3543	4.5651	3.5070	4.1878	3.2926
Day3	2.7076	2.1568	3.7510	2.6856	4.1948	2.7037	4.0549	2.6667
Day4	2.9063	2.4727	2.9170	2.5019	2.7875	2.5861	2.6399	2.6080
Day5	3.0004	2.5388	3.6431	2.1429	4.3931	2.5808	3.9705	2.5022
Day6	3.1074	1.7975	3.0704	1.7990	3.5920	2.1638	3.5151	2.1886
Avg	2.8487	2.2725	3.2759	2.2531	3.8780	2.7445	3.6516	2.7173

there is not much difference in the signal strength for locations in this area. Figure 5.5 raises a hope that better results can be achieved with a better deployment of sniffers.

Alternatively, more sniffers could be deployed to provide better coverage for the building. This way, only the three best signal strength measurements are selected for future applications, i.e. if the received signal at one monitor is too weak, the sniffer will not report the signal strength measurement to the global monitor (GM).

5.1.7 Localization Performance

ARIADNE radio propagation model constructs an imprecise signal strength map *SS-MAP* on a grid of locations. To locate a mobile M "sniffed" as $M(SA, SB, SC)(L)$ at a location L, the signal strength map *SS-MAP* must be *searched* for a best match. This section evaluates the Least Minimum Squared Error, the multiple nearest neighbors, and the proposed clustering-based search techniques. The impact of the grid resolution is also evaluated.

As explained in Section 2.2.2 and Section 3.2, *LMSE* picks only the position with LMSE to the "sniffed" signal strength of the mobile user. The scheme of multiple nearest neighbors (*nearest neighbors in signal space, closeness elimination scheme*), as the name suggests, selects multiple closest neighbors and computes the average of these neighbors' positions for the estimates. This

work evaluates both 2 and 3 nearest neighbors. The *Clustering-based* search method works similarly to multiple nearest neighbors technique, however, it is more flexible in that it does not restrict or fix the number of neighbors. Instead, the clustering-based search algorithm selects a set of candidate positions with signal strength close to the mobile user, and group these positions in space into multiple clusters. Then the algorithm picks the center of the largest cluster as the estimate.

A signal strength map *SS-MAP* is built over a grid of known locations (reference points) based on the proposed radio propagation model. The horizontal and vertical distances between reference points are 0.75 meter and 1.5 meter, respectively (see Figure 5.1). Six different *SS-MAP*s for the 6 days Day_i were constructed to evaluate the localization performance of the three strategies.

With the "sniffed" signal strength triplet as an input, the signal strength map *SS-MAP* is searched using *LMSE*, *multiple nearest neighbors*, or *clustering-based* techniques. Table 5.3 summarizes the error distance in meters between the real location and the estimated location. The error and standard deviation are reported for each search method for the six days. The last row provides an overall average of the six days. The *clustering-based* localization algorithm in general outperforms all other techniques. For a grid positions of 0.75×1.5 $meter$ apart, and for the floor plan in Figure 5.1, clustering-based method gives the position estimation with average error of 2.8487 meters. The estimation with *clustering-based* is respectively 14.99%, 36.13% and 28.18% closer than with other techniques.

Note that the localization performance reported in the table 5.3 is based on a dynamically estimated SS-MAP over a grid of positions with resolution of 0.75×1.5; and the system has only three sniffers that are not optimally deployed inside the building (Please, see Section 5.1.6). Thus we anticipate that better localization performance can be obtained if the system is with optimal configurations and finer reference grid resolution (as will be discussed in the next section).

Table 5.4: Grid resolution on the performance of Localization

	0.75×1.5 m		1.5×1.5 m		3.0×3.0 m	
	err	std	err	std	err	std
Avg	2.8487	2.2725	3.2861	2.0494	3.4481	3.5999

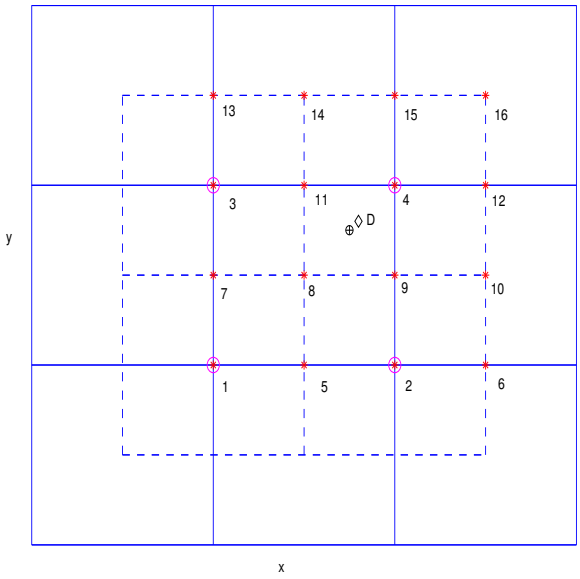

Figure 5.6: Grid resolution on the performance of clustering localization

5.1.8 Impact on Grid Resolution between Reference Points

This section addresses the question whether a finer resolution grid signal strength map would yield better accuracy for the *cluster-based* search technique. A simple example is provided to explain why finer grid resolution yields better results and simulations confirm this in Table 5.4. Table 5.4 provides the error and standard deviation for 3 different grid resolutions (in meter): 0.75×1.5, 1.5×1.5, and 3.0×3.0.

Lower estimation error at finer grid resolution is due to more candidate points in the vicinity of the true position. Figure 5.6 illustrates the impact of grid resolution. In Figure 5.6, the true position of the mobile user is denoted by $\diamond$ at point D. If coarse grid resolution is to be used (reference

Table 5.5: Mobile vs. static localization

	Mobile	Static
Path 1-7	2.0798	2.8558
Path 8-11	1.7476	2.6384
Path 16-19,15	2.0742	2.3065
Path 21-25	3.3953	4.9681
Path 27-30	1.3847	1.7620
Avg	2.1363	2.9061

positions at cross points of solid lines), a set of four points ($1 \sim 4$ with sign $\otimes$ in the figure) may be selected as the final cluster group of estimates. The center of this cluster is given at position 8. Alternatively, if finer grid resolution is to be used (reference positions at cross points of both solid and dashed lines), then all points in the figure ($1 \sim 16$ with sign * in the figure) may be included in the final cluster. And the center will be in position $\oplus$, which is much closer to the true location of the mobile user. Hence, fine grid resolution yields better localization performance.

5.1.9 Mobile User Localization

If a user is mobile, better accuracy can be achieved due to geometric constraints and physical limits (maximum speed, movement patterns along corridors, and users do not step on walls unless drunk!). As in RADAR, mobility helps improve accuracy. Assume a mobile user is moving along a corridor in Figure 5.1. The distance limit of the mobile user between two continuous locations within a sampling period is no more than 5.0 meters. Experiments with mobile users were conducted in this work. Table 5.5 provides the localization performance for stationary and mobile users.

The path information in Table 5.5 is corresponding to the data collection positions along the corridor in Figure 5.1. For example, *Path 1-7* denotes the scenario of a mobile user moving from

position 1 to position 7 along the corridor. The numeric value in the table is the average estimation error in meters for all data collection positions along the path. The bottom row is the overall localization performance for both cases.

5.2 Experiment II

The second experiment was carried out in a basement at Auburn University from April to July, 2005. The floor plan is given in Figure 5.7. The positions from 1 to 22 are the measurement locations for the indoor radio propagation validation, and the three-bit number associated with each room denotes the office number. In this floor, room 101, 107, 109, 112 and 111 are classrooms, room 110 is the computer lab, and the rest rooms are offices. Typical offices usually contain computers, (metal) bookshelves, and (metal) cabinets. The computer lab in 110 is occupied with three rows of computers, additionally, there are metal bookshelves and cabinets around all four walls. Besides the normal office/classrooms, there are a set of rectangular construction columns, as well as a set of storage rooms for the air conditioners.

Based on the observation from the first experiment, this second experiment will mainly focus on the sniffer position configuration and the floor structure modeling. Therefore, three, four or five sniffers will be deployed (Please, see Figure 5.8) in the building to study the deployment on the performance of the indoor wireless localization. For floor structure modeling, we will try various methods to model the construction columns and the office furniture (i.e. bookshelves and cabinets).

Different from the previous experiment, each sniffer in this experiment was implemented on a HP Pavilion v2000 with Orinoco Golden card running Linux Fedro II operating system. Through the AP, sniffers connected with the global monitor by the internal BroadCom wireless interface.

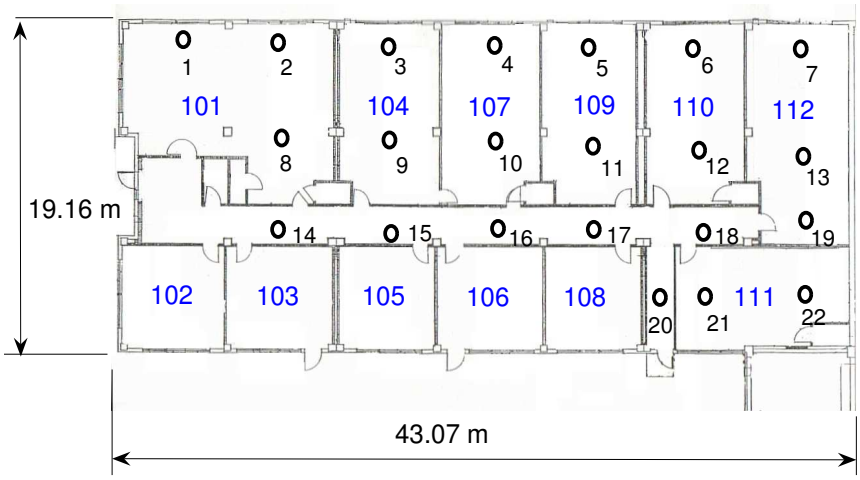

Figure 5.7: Floor plan with data validation positions

5.2.1 Sniffers Configuration

In order to analyze the sniffer's position on the performance of the indoor localization, we vary the **number** and the **position** of sniffers. We use three to five sniffers in this experiment, and a comprehensive illustration of the deployment details can be seen in Figure 5.8.

As it can be seen from Figure 5.8, we mainly takes five different deployment configurations in the experiment:

1. Linear: configuration (a) and (e);
2. Acute triangle: configuration (c) and (d);
3. Obtuse triangle: configuration (b);
4. Redundant zigzag: configuration (f), (g) and (h);

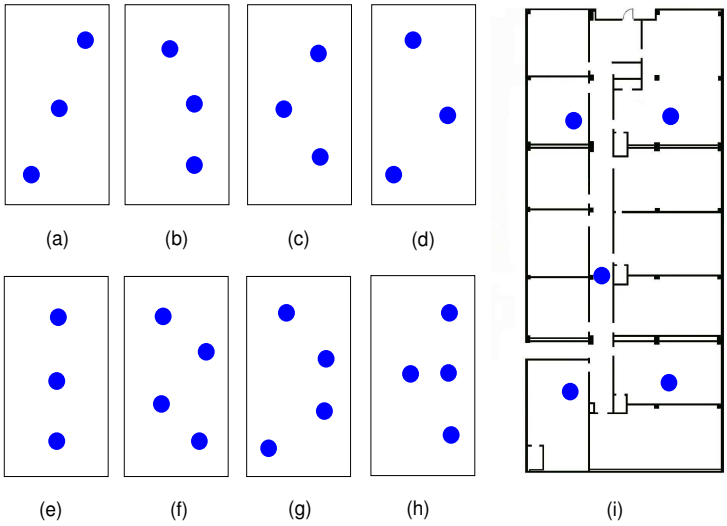

Figure 5.8: Various sniffer deployment

5. Semi-grid: configuration (i).

In the following sections, we will first evaluate adaptability of the indoor radio propagation model by comparing the estimated signal strength with the actual measurements at all data evaluation positions in the building. Then we construct the signal strength map, and analyze the localization performance under various deployment configurations.

5.2.2 Signal Strength vs. Distance

In order to estimate the relation between the distance and the received signal strength indicator (RSSI) along a straight line inside the building, we carried out two experiments: the first one measured the RSSI in the corridor, where there is no walls between the transmitter and the receiver (line-of-sight); and the second one measured the RSSI within rooms from 101 to 112. The results are given in Figure 5.9. In both figures, the x-axis is the distance in feet, and the y-axis denotes the

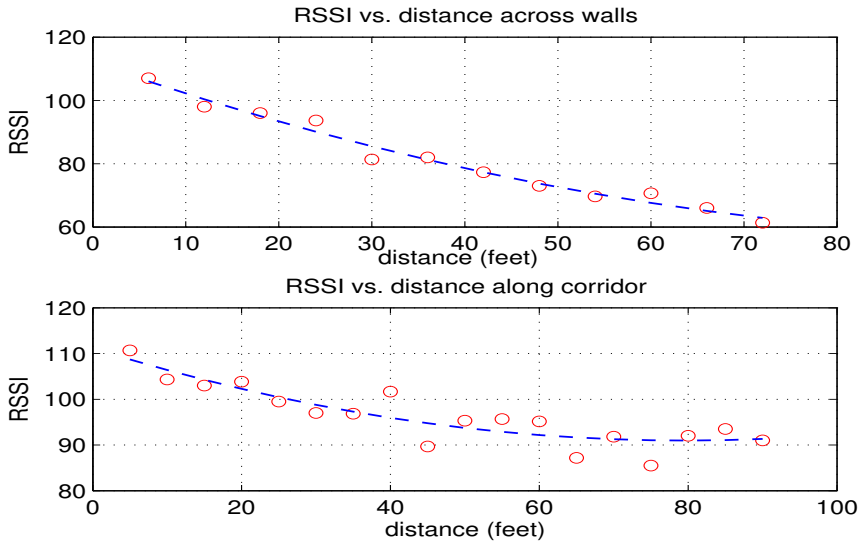

Figure 5.9: Indoor RSSI vs. distance

RSSI. The circle dots are measurement results, and the dashed lines represent the fitted trend-lines of the measurements.

It is obvious that there is no simple relationship between the RSSI and the distance. Comparing the measurements with the two configurations (with or without walls), the signal strength attenuation is higher when there are walls between the transmitter and the receiver. In both cases, the signal strength decaying slops (attenuation speed) are not linear. The experiments show that the simple radio propagation attenuation model in Section 5.1.2 may not be suitable for complex building environments.

5.2.3 Radio Propagation Model Validation

When three sniffers are deployed in the building as shown in figures (a) to (e) in 5.8, the ARIADNE system randomly selects ONE measured signal strength triplet from the 22 data validation

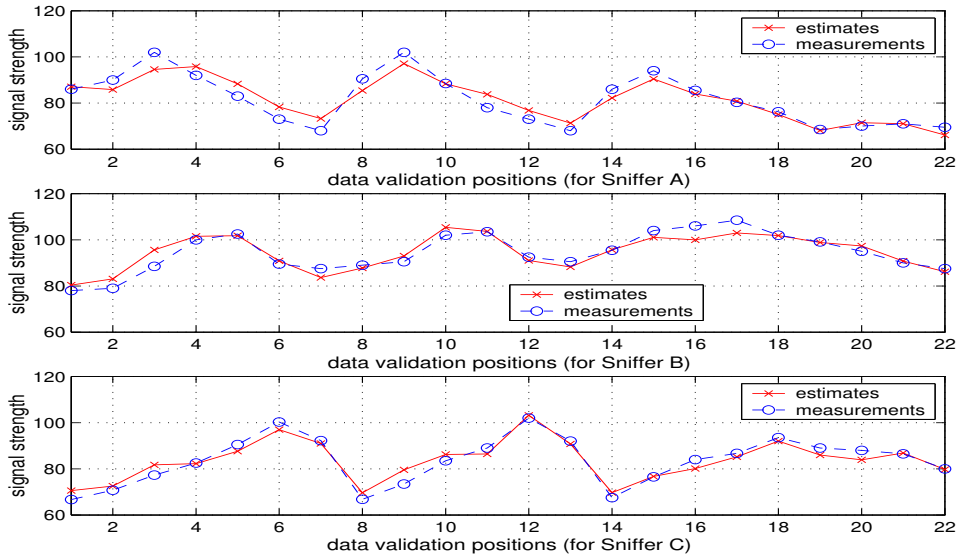

Figure 5.10: Indoor radio propagation model validation

positions. Then it regenerates the signal strength for all 22 positions. We validate the proposed radio propagation model by comparing the estimates with the measurements. Similar as the previous validation in another building in Section 5.1.2, very good agreement is achieved for all configurations. A typical comparison is given in Figure 5.10, and the detailed comparisons for the five configurations are given in Table 2. In Figure 5.10, the x-axis represents the 22 positions from the data collection and the y-axis denotes the signal strength measured as received signal strength indicator (RSSI). Points with '$\ominus$' are the signal strength measurements, and the points with symbol '$\times$' are the estimates.

Table 5.6 provides the detailed comparison for all configurations. The results shown here are the average difference between estimates and the measurements, and each estimation process uses only ONE reference measurement that is randomly selected from the 22 positions. Or in other words, the results are the average values of 22 comparisons. For each sniffer, the comparison gives

Table 5.6: Indoor radio propagation model validation for all sniffer configurations

	Sniffer A			Sniffer B			Sniffer C		
	err	std	max dif	err	std	max dif	err	std	max dif
config (a)	0.9598	0.0930	8.37%	0.7639	0.1040	6.91%	0.9975	0.1205	9.16%
config (b)	0.7179	0.0734	5.65%	1.0559	0.2381	9.36%	0.8811	0.2728	6.35%
config (c)	0.9144	0.0795	8.06%	0.7575	0.1079	6.64%	0.8193	0.1382	5.64%
config (d)	0.7373	0.0862	5.87%	1.0771	0.1359	9.40%	0.9405	0.1755	8.35%
config (e)	0.8073	0.0934	7.48%	0.9152	0.1412	8.71%	0.8341	0.2582	6.99%

mean square error (**MSE**), the standard deviation (**std**), and the maximum relative difference (**Max diff**) between the estimates and the measurements.

The maximum relative difference is calculated as the maximum difference between estimates and the measurements divided by the maximum measured signal strength values. When expressed by formula, it is in the format:

$$maximum\ relative\ difference = \frac{max(abs(SS_{est} - SS_{mea}))}{max(SS_{mea})} \tag{5.2}$$

From Figure 5.10 and Table 5.6,it can be seen that the introduced indoor radio propagation model effectively estimates radio signal strengths: the mean square error (MSE) of between measurements and estimates is less then 1.0, and the relative maximum difference is within 10% for all configurations. The results present in this paper are very similar to those of another building that were reported in Section 5.1.2. Thus we conclude that the radio propagation model proposed in the ARIADNE system is valid. Comparing the signal strength estimation performance at all configurations with three sniffers, it can be seen that there is no preference for the signal strength estimation on a particular configuration. Or in other words, given any sniffers configuration, the

Table 5.7: Localization performance for all sniffer configurations (Experiment II)

	Clustering		LMSE		2-N		3-N	
	err	std	err	std	err	std	err	std
configuration (a)	3.7575	2.1063	5.0022	2.6465	4.8774	2.7854	4.1998	2.7869
configuration (b)	4.1758	3.1056	4.4543	2.9503	4.4298	2.9801	4.1149	3.2755
configuration (c)	2.6986	2.0957	3.5987	2.5229	3.4119	2.5004	3.2570	2.0454
configuration (d)	2.8961	1.6098	4.4191	3.1258	4.5212	3.0957	3.3934	1.9603
configuration (e)	4.2239	2.9542	4.4217	3.4447	4.2486	3.3568	4.0814	2.3930
configuration (f)	2.3781	1.7420	4.1639	2.1107	4.1131	2.1008	3.0554	1.6391
configuration (g)	2.2646	1.4928	3.3410	1.8016	3.0167	1.8379	2.8114	1.8875
configuration (h)	2.4235	2.2844	2.9646	1.8999	2.9159	1.8679	3.4603	2.0390
configuration (i)	1.9176	1.8013	3.3090	2.4472	3.0140	2.3437	3.0516	1.9145

signal strength estimation using the proposed radio propagation model, for the same building and same sniffers, achieves similar precision.

5.2.4 Impact of Different Sniffers Position Configuration

Based on the radio propagation model and the reference, we built a signal strength map (SS-MAP) over a gird of positions with the resolution of 0.55m in both x and y directions. Similar to Section 5.1.7, we simulate the localization process with four algorithms, and the results are given as configuration (a) $\sim$ (e) in Table 5.7.

With the estimated signal strength map table, the clustering-based localization algorithm yields an estimation error of $2.5 \sim 4.3m$ for all configurations for the complex basement environment. Of the four algorithms, the clustering method proposed in this paper produces the best results for all configurations.

Comparing the localization performance for all configurations, the configurations (c) and (d) yield a much better location estimation performance than all others. The results indicate that the

Table 5.8: Average/Maxumum uncertainty distance for the building (Experiment II)

	d_{AUD}	d_{MUD}	standard deviation of d_{AUD}	*location estimates*
configuration (a)	4.8907	12.5976	0.4850	*3.7575*
configuration (b)	3.0809	8.1816	0.2146	*4.1758*
configuration (c)	2.0694	5.7051	0.3972	*2.6986*
configuration (d)	1.7815	4.5425	0.3408	*2.8961*
configuration (e)	6.3066	13.3697	0.4102	*4.2239*
configuration (f)	1.4752	4.4445	0.1952	2.3781
configuration (g)	1.4445	4.2666	0.2625	2.2646
configuration (h)	1.5556	4.4927	0.4064	2.4235
configuration (i)	1.1319	3.1726	0.1661	1.9176
sniffers configuration for **experiment I**	2.2774	6.1752	0.3888	*2.8487*

triangular configuration maximizes the discrimination of the signal strength triplet. This outcome corresponds to the conclusions in Section 3.4 at Chapter 3

In order to understand the intricacy of the sniffers configuration on the indoor localization performance, we compute the *average uncertainty distance* and the *maximum uncertainty distance* for all configurations from (a) to (e). Then we average the distances at all grid points for the whole building, and the results are given in Table 5.8. Figure 5.11 additionally shows the *average uncertainty distance* in z-axis at all grid points for the configuration (c) and (d). In the figures, the x- and y-axis give the floor plan in two dimensions. Both figures show that the area outside of the triangle of the three sniffers gives larger *average uncertainty distance*.

Table 5.8 also gives the *average/max uncertainty distance* for all other configurations form (f) to (i). The forth column is the standard deviation of the *average uncertainty distance*, and the fifth column is the location estimates for the corresponding sniffers configurations (adopted from Table 5.7). Comparing the values in column 2 and column 5, it is found that the errors of most actual estimates are larger than the *average uncertainty distance*. Or in other words, the

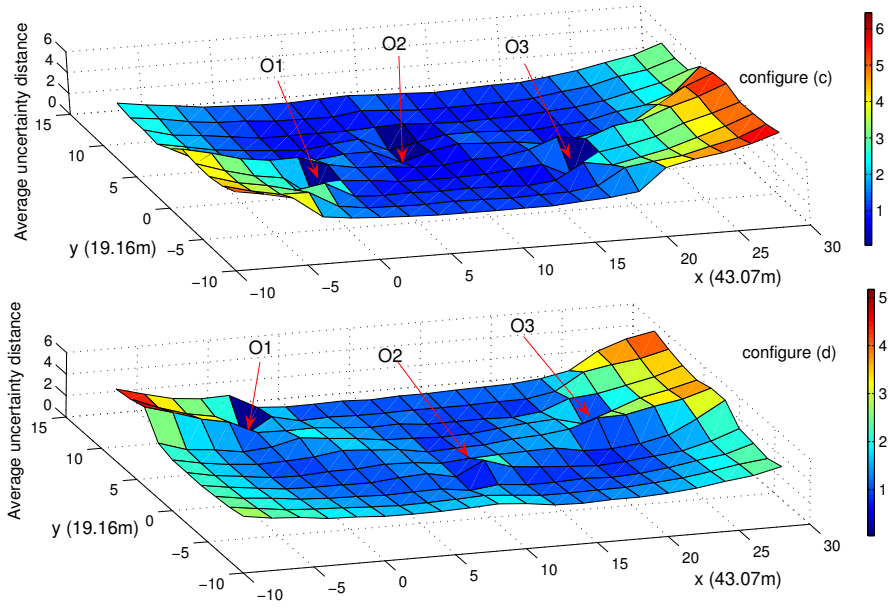

Figure 5.11: Average uncertainty distance of the estimation for configuration (c) & (d)

average uncertainty distance of the deployment gives the optimal achievable bounds for the indoor localization performance.

However, there are two exceptions in the configuration (a) and (e). The reason may lie in the fact that the linear sniffers configuration usually gives two candidate positions at both sides along the line formed by the sniffers. For an idealized environment without partitions, the two positions are symmetric along the central line, thus the *average uncertainty distance* equals the half of the distance between the two possible positions. In reality, due to the multipath effect of the indoor radio propagation, there may exist no such symmetric estimation at certain positions. Therefore, the theoretical *average uncertainty distance* over-estimates the errors for the aligned deployment scheme.

The results in Table 5.8 also indicate that the configuration (d) is the most optimal deployment strategy and the *average uncertainty distance* is within 2.0 meters. The configuration (c) ranks the

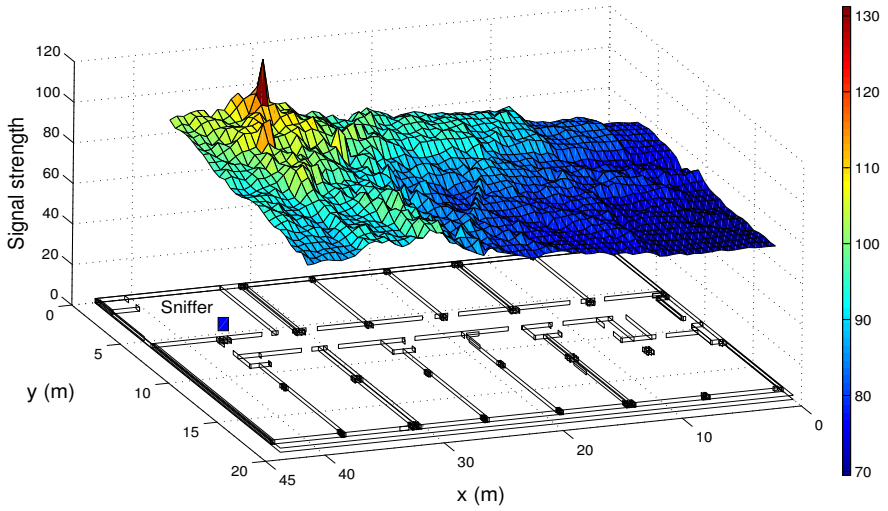

Figure 5.12: Signal Strength of the Sniffer at the lower corner (configuration (d))

next and the *average uncertainty distance* is around 2.0 meters. However, comparing configurations (c) and (d), the localization performance of (d) is a little worse than that of the configuration (c). This shows that the sniffers deployment should also consider many other site specific parameters in the building. An illustration is given in Figure 5.12, where it gives the signal strength map for one of the sniffers together with the corresponding floor plan. In the figure, the x-axis and y-axis denote the floor plan in two dimensions, and z-axis represents the received signal strength at this sniffer. As shown in the figure, the received signal strength for positions around the opposite corner (lower right side) is almost flat, or in other words, they are indistinguishable. Therefore, the deployment of this sniffer does not provide optimal signal strength coverage.

Therefore, to achieve optimal localization performance, the sniffers should not be aligned, instead, the deployment should be in acute triangle style in order to produce lower *average uncertainty*

distance for the interested locations. And the configuration must also provide full signal coverage for the estimated locations.

5.2.5 Number of Available Sniffers

In order to analyze the amount of available sniffers and their deployment strategy on the performance of the indoor localization, we deployed four and five sniffers inside the building (Figure 5.7) as shown in Figure 5.8 from (f) to (i). We repeated the whole localization process and finally we obtained the average localization error within 2.50m (see Table 5.7) for all three configurations. The improvement is systematic as compared with that of three sniffers in configuration (a) to (e). This conclusion agrees with the enhancement study from the RADAR system [61].

Similar as the configurations of three sniffers, we also calculate the *average uncertainty distance* for configurations from (f) to (i), and the results are given in Table 5.8. It can be seen that the deployment of more sniffers gives much smaller *average uncertainty distance*. Comparing the location estimation errors of configurations from (f) to (h) with four sniffers, the deployment of (h) is worse than the other two deployments in (f) and (g). It appears that the middle sniffer on the right side (of the three sniffers) improves the discrimination of the signal strength, however, it is within the coverage of the triangle formed by the other three sniffers, and therefore its contribution is limited.

The configuration (i) is in *semi-grid* style although we could not deploy the sniffers ideally due to administration restrictions. Initial experimental results in Table 5.8 indicate that this deployment strategy does provide the best *average uncertainty distance* and localization performance (within 2.0 meters).

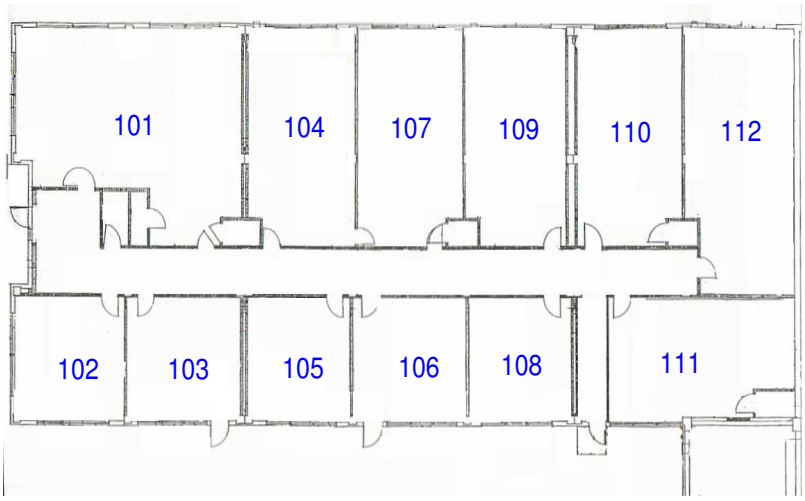

Figure 5.13: Floor plan without construction columns

Therefore, if the location-based applications require higher precision indoor localization estimation, the deployment of extra redundant sniffers could be a reasonable approach. And if the sniffers are deployed in the *semi-grid* style, the location estimation will achieve optimal performance.

5.2.6 Column/Post Modelling

We analyze the construction columns on effect of the indoor localization system in this section. For the same basement show in Figure 5.7, a floor without columns is given in Figure 5.13.

We compare the signal strength estimation performance and the localization performance based on the sniffer configuration (c) in Figure 5.8. The results are given in Table 5.9 and Table 5.10. From the two tables, it shows that the construction columns do affect the performance of both the indoor radio signal strength estimation and the localization, and the consideration of the columns in the floor plan model will help the indoor localization system.

Table 5.9: Signal strength estimation under various floor configurations using sniffer configuration (c)

	Sniffer A			Sniffer B			Sniffer C		
	err	std	max dif	err	std	max dif	err	std	max dif
Columns	0.9144	0.0795	8.06%	0.7575	0.1079	6.64%	0.8193	0.1382	5.64%
No Columns	1.1020	0.1103	8.76%	0.8758	0.0831	6.70%	1.0490	0.1765	7.56%
Furniture	0.8940	0.1094	7.46%	0.8528	0.1942	6.70%	0.8189	0.1896	6.37%

Table 5.10: Localization performance under various floor configurations using sniffer configuration (c)

	Clustering		LMSE		2-N		3-N	
	err	std	err	std	err	std	err	std
with Columns	2.6986	2.0957	3.5987	2.5229	3.4119	2.5004	3.2570	2.0454
without Columns	2.9628	2.4059	3.6045	2.0442	3.5934	2.1388	4.0710	2.7069
with Furniture	2.5706	2.0910	4.2124	28853	4.2120	2.8343	2.9353	2.0208

5.2.7 Furniture Modelling

As introduced in Section 5.2, some of the rooms ($\#102 \sim \#106$, and #108) in this basement are offices, and the room #110 is computer lab. All offices contain some furniture (bookshelves and cabinets); And the computer lab is extremely different from general offices and classrooms in that it contains three rows of computers; additionally, there are bookshelves and cabinets around the room along all four walls (except the door). The furniture is mainly made of metal, and the average height of them is around 1.70 meters. They should affect or even block the radio propagation.

To consider the furniture's effect on the indoor localization, we roughly model the furniture as additional walls along with the floor plan. We repeated the localization process for the sniffer configuration (c) in Figure 5.8, and the results are also included in Table 5.9 and Table 5.10.

The results show that the consideration of the furniture slightly improves the overall performance of the indoor localization system. However, the results also indicate that to simply model

Table 5.11: Measured RSSI at different humidity environment

	Sniffer A	Sniffer B	Sniffer C
Max difference	18.0	17.0	9.0
Average difference	7.8091	4.6134	3.1636
MSE	1.8674	1.2782	0.8467

the furniture as additional walls does not optimally capture the radio attenuation effect of the indoor radio propagation.

Moreover, in this basement, there are three storage rooms and two closets that are used mainly for air conditioners and sundry goods. We could not access these rooms, therefore the internal structure is unclear. We believe these storage rooms also affects the performance of the indoor localization system.

5.2.8 Cell-based Localization

Cell-based localization means that we localize a mobile based on the room resolution instead of absolute geometry coordinates. In this way, the only requirement is to correctly position a mobile within a right room without actual consideration of the detail location.

In this experiment, we treat each room as a cell, and the long and the short corridors (between room #108 and #111) also denote two cells. For the tested positions in this experiment, over 85% probability is obtained to correctly estimate a mobile within the right room. The only 2 or 3 missed positions are around the computer lab (#110), and they are usually found in the nearby cells. This indicates that cell-based localization could be reliably used for some localization based applications; more over, the correct modelling of the furniture is a necessary in the future research.

5.2.9 Humidity Effect

IEEE 802.11b standard uses radio frequency in the 2.4GHz band. Signal of this frequency is subjected the influence of the environmental humidity. We conducted an experiment in such an environment with humidity of over 95%, and we compared the signal strength measurement with that of normal dry environment (humidity around 30%). The comparison for sniffer configuration (d) in Figure 5.8 is given in Table 5.11.

Compared with the signal strength measurement variance in the first experiment (Please, see Table 5.1), the humidity does greatly affect the signal strength measurement. This again indicates that dynamic signal strength map SS-MAP construction is necessary for the real-time indoor localization.

5.2.10 Dynamic SS-MAP Update

A 'stationary emitter' as described in [105] judiciously positioned in the building can be used to be periodically "sniffed" to provide the reference signal strength measurement to dynamically generate a real time signal strength map *SS-MAP*. Such reference device would capture dynamic changes in the environment. The question is whether the map can be computed fast enough to take into account swift changes in the building.

The signal strength map *SS-MAP* table consists a grid of known locations in which the values of corresponding signal strength are stored. Generally, a higher resolution table is required in order to obtain better location estimation. Higher grid resolution induces more computation for the construction of the ray tracing from each point to a set of receivers (APs or sniffers). For example, the approximate time taken to run the ray tracing for 30 positions and 3 sniffers is about 2 hours on a machine of x86 family processor at 1.4GHz with 256 MB physical memory. However, a building

floor plan rarely changes. Therefore, ray tracing can be processed once and ray information can be stored in advance. The stored ray information can be fed to ARIADNE to generate a dynamic signal strength map. Given the ray information, the construction of a *SS-MAP* table with 300 points and 3 sniffers takes less than one minute. So, a dynamic realtime signal strength map is possible as long as structure conditions in a building remain stable.

5.3 Discussion

To improve the localization performance, the following three problems must be solved.

1. Accurate signal strength readings from all sniffers: In the first experiment, the readings from sniffer B contain system error (see Section 5.1.3). The performance of the proposed localization scheme should improve for the floor plan in the experiment;

2. Optimal sniffer deployment: In Section 5.1.6, it is found that deployment of the three sniffers (for the first experiment) is not optimal. Specially, coverage-insensitive areas do exist at corners in the studied floor of the building. To improve this, sniffers A and C may be placed a little closer to the center in y direction (see figure 5.5).

   From the second experiment, it shows that the triangular sniffer deployment (configuration (c) and (d) in Figure 5.8) gives much better localization performance. Additionally, more sniffers will generally help the performance of the indoor localization system as given in Table 5.7. And if the sniffers are deployed in the *semi-grid* style, optimal localization performance can be achieved.

3. Furniture modelling: As discussed in Section **??**, the simple modelling method for the furniture can improve the localization performance, however, to model the furniture as additional

walls does not fully capture the attenuation properties of the furniture. Additionally, the closets and AC storage rooms are not modelled in the second experiment.

4. Grid resolution: The grid resolution of the reference positions in the SS-MAP table affects the localization performance, and as indicated in Section 5.1.7, finer resolution grid will improve the localization performance. However, finer resolution SS-MAP generally require more computation for both table construction and the real-time localization.

Thus, the precision of the indoor localization really depends on many parameters (building structures, sniffer configuration, available sniffers, furniture, grid resolution, and many others), and it is inappropriate to evaluate the performance of the indoor localization systems only based on the best reported localization precision.

CHAPTER 6

APPLIED RESULTS FOR OUTDOOR LOCALIZATION SYSTEMS

This chapter dedicates to present the simulation results for the outdoor localization algorithms. Section 6.1 will first introduce the simulation environment. Section 6.2 presents the simulate results for the proposed algorithms, in Section 6.3, the dissertation compares the performance of various localization algorithms, and Section 6.4 introduces the results for the radio range irregularity.

6.1 Environment and Settings

The simulation is built in Matlab (version 6.1). We measure the estimation errors (see section 4.1.2) under various network settings. To simulate the ad hoc nature of the network, we run each test case multiple times with the same network settings, and take the average results for this dissertation.

For the simulated annealing MDS (SA-MDS) algorithm, we select the initial temperature T_{max} value as the half radio range (i.e. $T_{max} = \frac{R}{2}$). And for all simulation runs, we use a network of 100 randomly deployed nodes in a unit square area. In addition, the network can update the following parameters: 1) Anchor node density, 2) Radio transmission range, 3) Deployment randomness, and 4) Range measurement errors.

In the following sections, we first present the simulation results for MDS, IT-MDS and SA-MDS algorithms under various network conditions. Then, we outline the reported localization performance from other research groups. In the end, we compare, under the same network settings, the performance of proposed IT-MDS and SA-MDS algorithms with other representative algorithms.

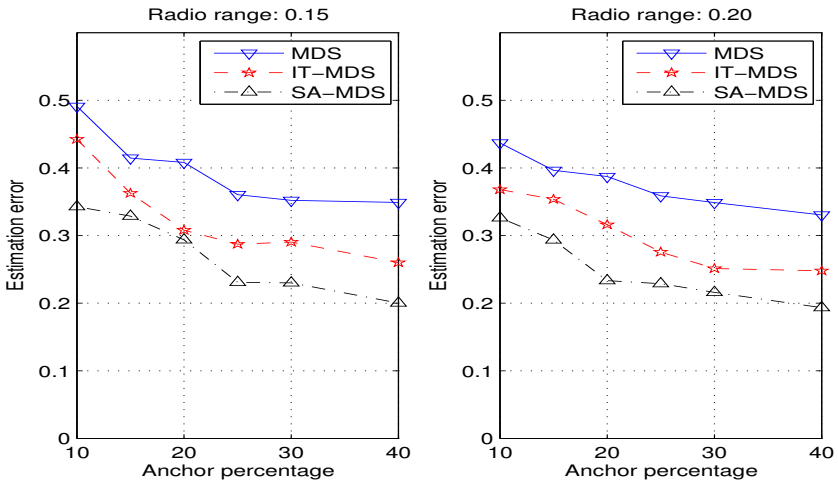

Figure 6.1: Precision with the anchor ratio

6.2 Algorithm Simulation Results

To understand the dependency of the estimation precision on the anchor population, we run the IT-MDS algorithm and the SA-MDS algorithm with different number of anchors in a set of networks with same network parameters. Figure 6.1 shows the relation between the precision and the anchor population. The deployment randomness is 0.4. The x-axis of the figure represents anchor percentage, and the y-axis denotes the estimation errors in term of radio transmission range.

The left chart of Figure 6.1 shows the results of radio range of 0.15, and the right chart shows that of radio range of 0.2. Both results indicate a trend that more anchors will lead a better performance. Figure 6.1 also shows the relation between the estimation error and the radio transmission range. The result is that both algorithms give better estimates at larger radio transmission range.

Figure 6.2 gives the experimental results of the deployment randomness (x-axis) on the performance of the positioning precision (y-axis). It shows MDS based algorithms are insensitive to the deployment randomness.

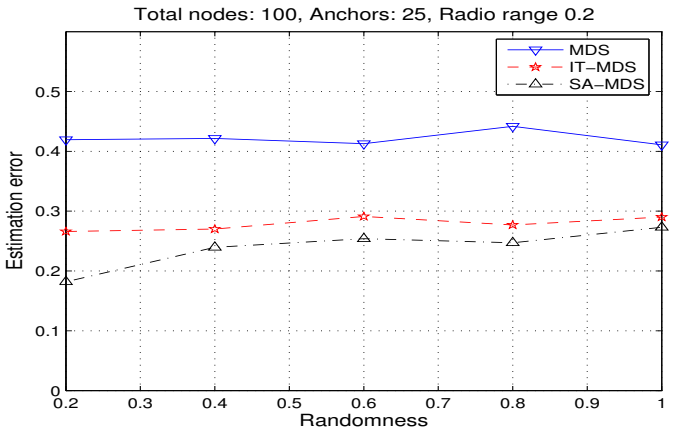

Figure 6.2: Precision with randomness

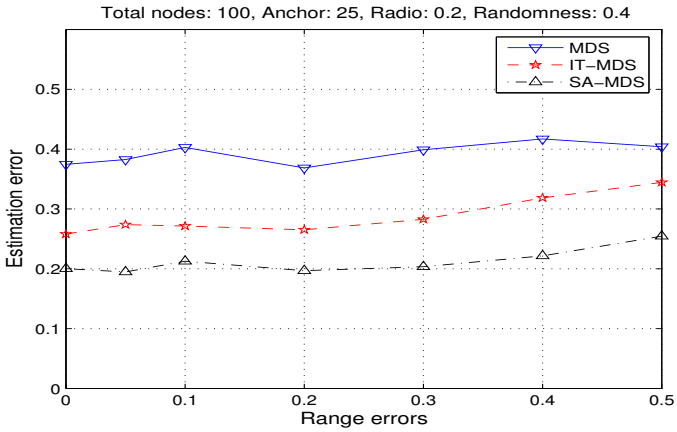

Figure 6.3: Precision with the 1-Hop range error

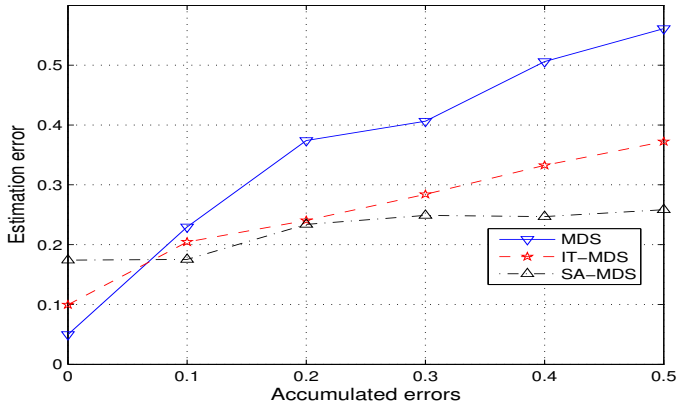

Figure 6.4: Precision with the accumulated Euclidean error

To determine the effect of the 1-hop range measurement error on the performance of the estimation precision, we simulate the 1-hop range error for both algorithms. The range error is determined relatively to the radio range; and to simulate the real situation, we define a maximum variance for all range measurements. The actual range error is determined dynamically during the experiment by the production of the maximum variance and a random number between -1 and 1. The results are given in Figure 6.3. It shows that the estimation performance does not significantly decrease with the increasing 1-hop range errors. The main reason may because of the incorporation of the communication constraints during the iteration. Also we relocate all anchors' positions and adjust the corresponding distance among them after each estimation iteration. These procedures can potentially minimize the range errors.

To measure the accumulated errors for the *Euclidean* propagation method, we conducted similar experiments as for *DV-Hop*. The results is given in Figure 6.4. The figure shows the increasing estimation errors when the accumulated errors increase. However, with *Euclidean* method, both proposed algorithms give better estimation performance if small error is presented.

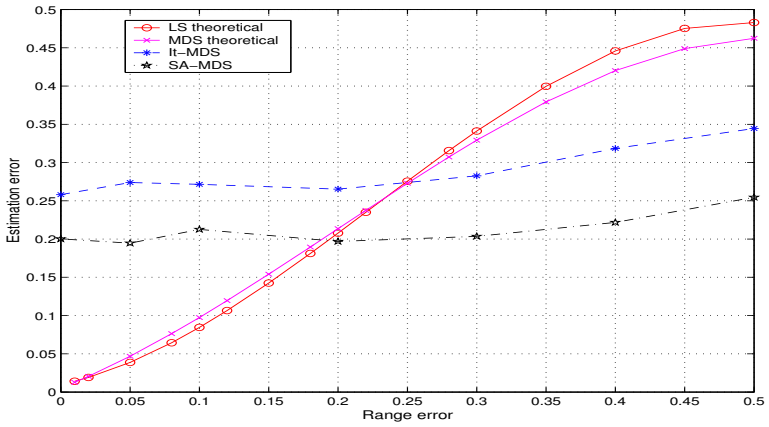

Figure 6.5: Precision analysis

Figure 6.5 shows the theoretical estimation precision for both the MDS and the Least Squares under constant perturbation (x-axis) (Please, see Section 4.4 in Chapter 4). It also gives the estimation results for IT-MDS and SA-MDS algorithms. The network for the proposed algorithms is a network with 100 nodes, anchor ratio is 20%, the radio range is 0.2, and the randomness is 0.4. The results show that the proposed algorithms are fairly robust on the 1-hop pair-wise distance errors because of the radio communication constraints that have been adopted in the methods. Comparing the theoretical results for LS and MDS, it shows that the estimation errors for both algorithms are pretty close, and at larger range measurement error, the location estimation of the LS is a little worse than that of the MDS algorithm.

In Figure 6.5, it also indicates that when the range error is less than 0.25, the theoretical estimation error is less than that of the proposed algorithms. The reason may lie in the fact that the proposed algorithms use very limited range information between pairwise nodes during the location estimation.

Table 6.1: Reported estimation error in literature

Algorithms	**Condition & settings**	**Estimation error**
Probability algorithm by Ramadurai and Sichitiu [38]	transmission range: 20m	47%
Ad hoc positioning system by Niculescu and Nath [67, 36]	isotropic topology, 10% anchor isotropic topology, 20% anchor anisotropic topology, 10% anchors anisotropic topology, 20% anchors	35% 25% 100% 90%
Robust algorithm by Savarese [30]	connectivity >7; anchor density >5%	33%
Approximate point-in-triangulation test (APIT) algorithm by He [33]	higher power transmitters at anchors	45∼50%
Amorphous algorithm by Nagpal [34]	local neighborhood density about 20	20∼37%

6.3 Algorithm Comparison

6.3.1 Reported performance

In order to evaluate the localization performance of the proposed algorithms, in this section, we first illustrate the reported localization performance from other research groups.

The results are given in Table 6.1. It can be seen that the location estimation performance is closely related to special network settings. Consequently, it is very difficult to compare them only with these values. From the table, it can be seen that the best reported estimation error is within 20%∼50% for networks with normal (isotropic) topology.

6.3.2 Localization Performance under Same Network Environment

In this section, we first compare our IT-MDS and SA-MDS with the iterative MDS algorithm by X. Ji in [9]. Then we compare the localization performance of the proposed algorithms with three other representative algorithms under same networks of different deployment randomness.

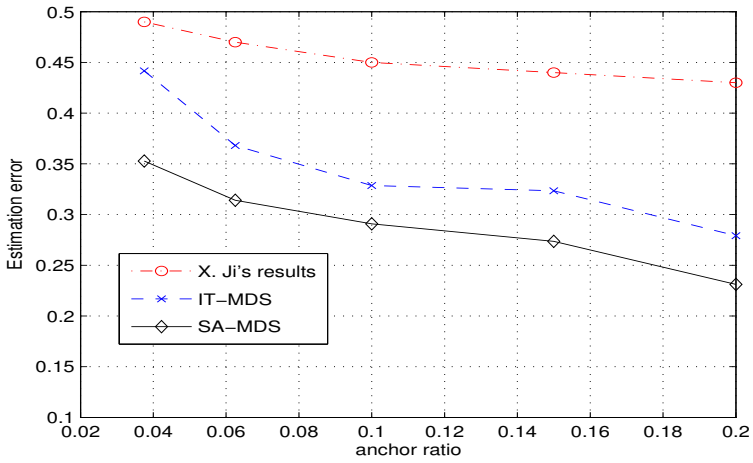

Figure 6.6: Comparison with X. Ji's algorithm

Figure 6.6 gives the comparison results of our algorithms with the iterative algorithm by X. Ji. The test environment includes 400 nodes in a unit square area, the distance measurement error is 0.5 radio range. The x-axis denotes the anchor ratio, and y-axis is the estimation error. X. Ji's estimation results is adopted directly from [9]. The simulation results show that the localization algorithms proposed in this dissertation give better estimation performance.

Figure 6.7 compares various localization algorithms under network environments with different randomness (see section 2.3.1). The x-axis denotes the network randomness, and y-axis is the estimation error. The selected Grid overlaying method is the APIT method by He *et al* [33], and the bounding box method is by Savvides *et al* [8].

The test network contains 100 nodes, and anchor ratio is 20%, radio transmission is 0.2. In this experiment, we use *DV-Hop* radio propagation method to estimate the distance between communicating nodes. (The setting in this experiment is different from the original testing environment of the APIT algorithm (by He [33]) where higher power transmitter is used at anchors, so the estimation performance for this APIT algorithm is a little worse.)

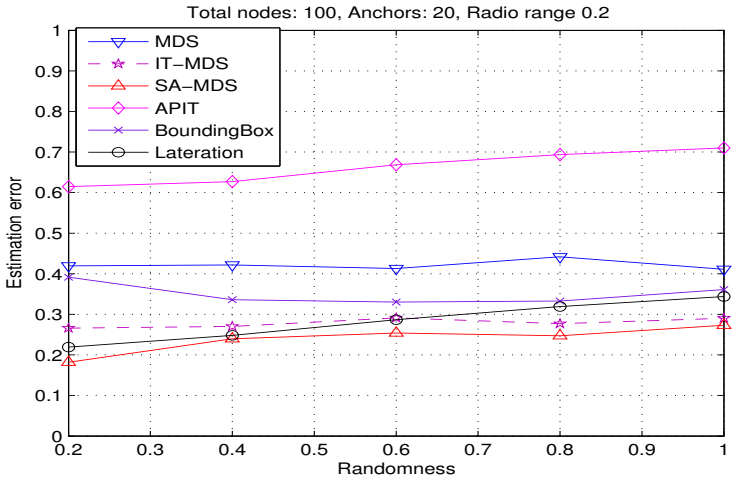

Figure 6.7: Algorithms comparison

From Figure 6.7, we can find that the SA-MDS algorithm performs better than all other algorithms, and the IT-MDS also achieves better estimation performance at larger deployment randomness.

For the test run (20% anchors), we find that SA-MDS and IT-MDS algorithms present 18%∼25% and 25%∼30% estimation error, respectively, independent of the deployment randomness. Compare with the reported localization performance in literature (see Table 6.1), the proposed localization algorithms in this dissertation do yield better estimation performance.

6.4 Simulation Results for Radio Range Irregularity

This section evaluates four localization algorithms under the irregular radio transmission model: 1) Iteration MDS (IT-MDS) and simulated annealing MDS (SA-MDS) (please, see Section 4.2 in Chapter 4); 2) bounding box algorithm [8]; and 3) lateration algorithm [30]. A detailed survey on these algorithms can be found in [76].

We will first introduce the network settings and localization metrics in Section 6.4.1, then we evaluate in Section 6.4.2 these localization algorithms; and in Section 6.4.3, we assess the same algorithms using the optimized propagation method.

6.4.1 Network setting and localization metrics

1) Network environment: We randomly deploy 100 nodes in a unit square area. We define the node deployment randomness as the maximum displacement, in terms of radio transmission range, of each node that can swing from the cross-points of a coordinate grid (also see Section 4.1 in Chapter 4).

2) Estimation error: The estimation error is expressed as a fraction of the maximum radio range. Let $X_i(i = 1, 2, ..., n)$ be the coordinates of n normal nodes where $X_i = (x_{i1}, x_{i2}, x_{i3})^T$. If $\widehat{X_i}$ are the estimated coordinates, then the estimation error is given by the Euclidean distance between X_i and $\widehat{X_i}$:

$$err = \frac{\sum_{i=1}^{n} \sqrt{(X_i - \widehat{X_i})^T (X_i - \widehat{X_i})}}{n \cdot R} \tag{6.1}$$

In the charts at this section, we use the y-axis to denote the estimation error; and use x-axis to represent network environment conditions, such as anchor ratio or deployment randomness.

6.4.2 Performance under Irregular Radio Networks

In this section, we will evaluate the radio irregularity on the performance of a set of representative localization algorithms. Specifically, we analyze these parameters: i) number of sectors or the number of edge of the polygon N; ii) maximum range attenuation ε; and iii) maximum ideal radio range R.

Table 6.2: Estimation errors under various irregular attenuation

	IT-MDS	**SA-MDS**	**Bounding-box**	**Lateration**
$\varepsilon = 0$	0.2408	0.2330	0.2737	0.2371
$\varepsilon = 0.2$	0.2437	0.2465	0.2848	0.2770
$\varepsilon = 0.5$	0.3367	0.3091	0.3574	0.3322

Overall localization performance under irregular radio transmission

The overall effects of the radio irregularity on the localization estimation performance are given in Table 6.2. The testing network includes 20% anchors, the maximum radio transmission rage is 0.2 unit, and node deployment randomness is 0.4. The table shows the location estimation under three different maximum attenuation. The second row ($\varepsilon = 0$) represents the ideal case where the radio is perfectly circular, and the third and the forth rows give the results with maximum attenuation of 0.2 and 0.5 respectively. We use a regular polygon with 16 edges (16 sectors) to simulate the operating environment.

It is not surprise that higher attenuation generates larger performance degradation. And the results agree with the conclusion by He *et al.* in [33].

In Figure 6.8, we show the estimation results of these algorithms at various anchor ratios. The x-axis denotes anchor ratio, and y-axis denotes the estimation error. And the four charts denote, respectively, the simulation results for all four algorithms. The network randomness is 0.4, the edge of the polygon is 16, and the maximum radio transmission range is 0.2. The results indicate that the increase of the anchor ratio could remedy the negative effect of the irregular radio transmission.

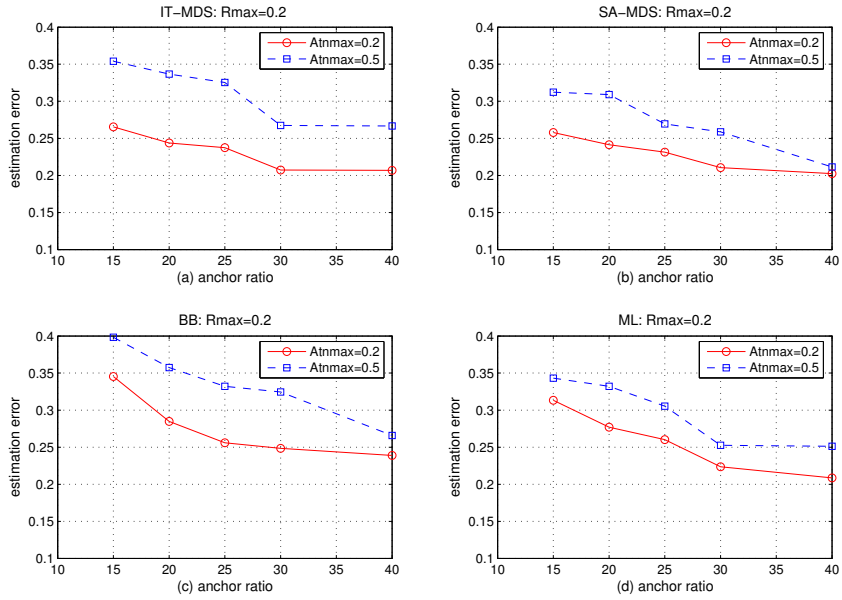

Figure 6.8: Maximum attenuation on the performance of localization

Number of Sectors on the Localization Performance

To under the environment determined radio sectors (expressed by the polygon edge number) on the performance of the localization estimation, we change the sector number of the radio transmission, and we include the results in Figure 6.9. Considering the space limitation, these four charts are on different network conditions. Figure (a) denotes the simulation results of IT-MDS algorithm with maximum radio transmission range of 0.2, and the maximum attenuation is 0.2. Figure (c) represents the estimation of bounding-box algorithm at the network of maximum attenuation of 0.5, and maximum radio range of 0.3. Figure (b) and Figure (d) are on the same network environment with maximum radio of 0.2, and maximum attenuation of 0.5 respectively.

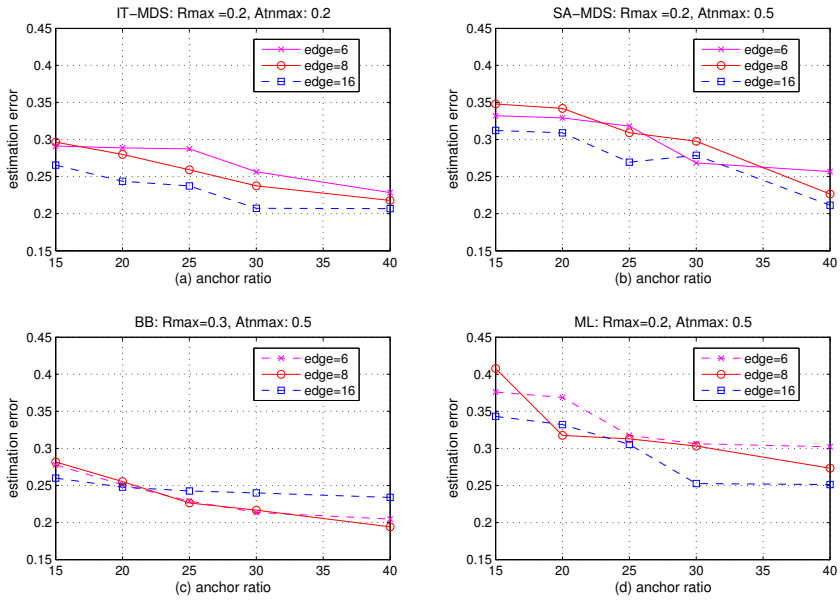

Figure 6.9: Number of edges on the performance of localization

As it can be seen from the charts, when the maximum attenuation is smaller, larger edge number will generally gives better location estimation performance (Figure (a)). The reason is straightforward in that larger number of edge will give almost regular circular radio transmission shape, and if attenuation is also small, then the radio coverage is larger and it is closer to the ideal radio transmission.

On the contrary, when the attenuation is larger, the trend is not so clear (figure (b), (c), and (d)). And the results in figure (c) show that the bounding-box method is relatively stable and insensitive to the edge number.

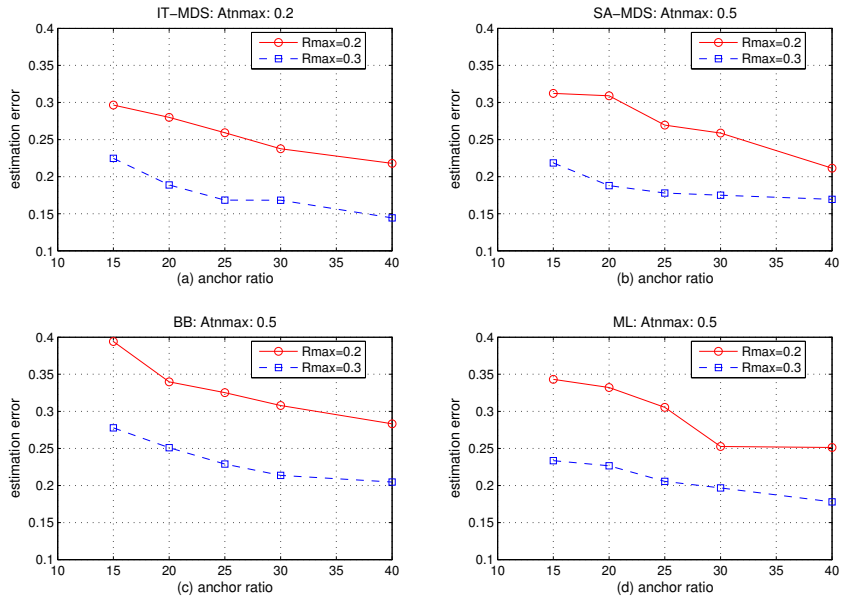

Figure 6.10: Maximum radio range on the performance of localization

Maximum radio range on the estimation performance

Figure 6.10 gives the location estimation of the four algorithms at two different maximum radio transmission ranges. The x-axis represents the anchor ratio, and the solid line with circle, and the dotted line with rectangle denote, respectively, the results at maximum radio of 0.2 and 0.3. We use 16 edges in this simulation. It is clear that larger maximum radio range will result smaller estimation error.

6.4.3 Optimized Propagation on Real Networks

In this section, we show the location estimation results using original and optimized propagation methods for the four methods. In the simulation, we fix the anchor ratio at 20%, and change the randomness from 0.2 to 1.0 (x-axis). The maximum radio transmission is set to 0.2, and the

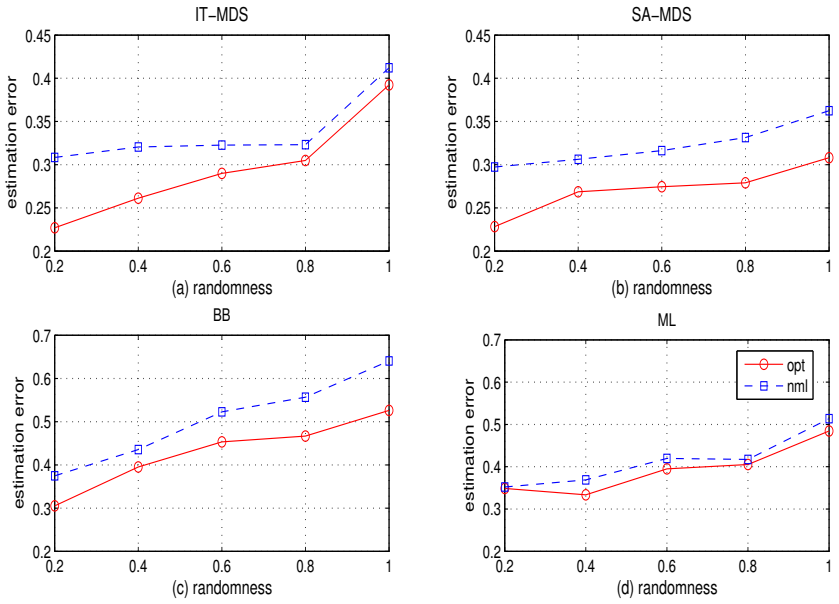

Figure 6.11: Radio propagation optimization on the performance of localization

maximum attenuation is 0.5. The results are given in Figure 6.11. The results using the optimized propagation method are given in solid line with circles (legend *opt*) and the results using the DV-Hop method is given in dotted lines with rectangles (legend *nml*).

The results indicate that the localization performance degrades when the deployment randomness increases. If the optimized propagation method is used, the estimation performance for all four algorithms is improved. And the average improvements for these four algorithms are 17.72%, 24.69%, 10.67% and 5.33%, respectively.

CHAPTER 7

CONCLUSIONS

This dissertation researches the wireless localization problem, and proposes realistic localization mechanisms for both the indoor and outdoor environments. For the indoor localization research, the dissertation introduces a new and automated localization tool called ARIADNE. For the outdoor localization analysis, the dissertation presents two multidimensional scaling (MDS) based algorithms: the iterative MDS and simulated-annealing MDS. Additionally, a realistic radio range irregularity (RRI) model is proposed in order to provide real operating test-beds for the wireless research community.

7.1 Indoor Localization Research Conclusion

The dissertation presents a new and automated indoor localization system: ARIADNE. The system contains two modules: Signal strength map construction and location search.

In the signal strength map construction module, a new radio propagation model is derived to enable the creation of the signal strength map for an entire building with minimal manual intervention. The scalable algorithm generates a signal strength map with high accuracy and thus can be easily deployed to construct these maps for indoor premises. The time varying nature of the propagation characteristics of the wireless channel poses problems to a signal strength table created manually. This is because, even though such a table is accurate at a given instant of time, it can be rendered useless at another instant. The map generation module presented in this dissertation

enables the creation, on a demand basis, of a signal strength map automatically and almost instantaneously. The resulting map is comparable in accuracy to that of a signal strength map manually collected at that instant of time.

In the localization search module, a clustering-based localization algorithm is developed to search the inaccurate signal strength map . Simulation results at two different buildings validate the algorithms and procedures used in ARIADNE. In addition, if position history information is to be used, the localization performance for a mobile user is significantly improved.

During the research in the indoor localization system, it was found that many parameters could affect the localization performance. These parameters include: **1)** sniffers deployment configuration, **2)** number of available sniffers, **3)** signal strength map resolution, **4)**furniture modelling, and **5)**special building structure modelling. We have found that the following mechanisms could potentially improve the localization performance:

1. Triangular (i.e. not aligned) deployment of the sniffers *maximizes* the discrimination of the signal strength triplet in the signal strength map table;

2. The deployment of extra sniffers improves the localization performance;

3. The sniffer deployment should consider the maximum signal strength coverage for the estimated location. Additionally, it is better to cover the important position within the triangular surface of an equilateral triangle formed by three sniffers. Additionally, the *semi-grid* sniffers deployment for large building is optimal;

4. Higher resolution of the signal strength map table improves the localization precision;

5. Large furniture, like bookshelf, affects the indoor radio propagation; thus, a simple method is to model them as additional walls. This will improves the localization performance if no other mechanisms is available;

6. To model the construction columns as additional walls, in two directions and according to the thickness of the columns, helps the localization performance; however, the material property of the column may be different from general indoor partitions.

Thus, the localization precision of the indoor system really depends on many parameters. It is generally not appropriate to evaluate or to compare the performance of various indoor localization systems only based on the best reported performance, which may depend on some test-bed with particular sniffer configurations.

7.2 Outdoor Localization Research Conclusion

For the outdoor localization research, this dissertation proposes two algorithms, IT-MDS and SA-MDS, to enable nodes in a wireless network to estimate their positions. Both methods are based on a combination of multidimensional scaling method and lateration method. Specifically, IT-MDS algorithm considers the radio communication constraints, and dynamically adjust node's positions during each iteration. SA-MDS algorithm, on the other hand, mimics the way of a metal cooling down procedure, and optimally locates the nodes in their most likely positions.

If *DV-Hop* is used, both algorithms are insensitive to the 1-hop range measurement errors. With extensive simulation runs, results show that both algorithms provide accurate and consistent estimates no matter how precise each single estimate is. The average estimation errors are roughly bounded within 40% if a network contains more than 10% anchors and 0.2 radio range is used.

With same network settings, the research compares the proposed algorithms with other existing representation algorithms, the results indicate that the proposed methods yield better estimation performance.

Together with the localization research, this dissertation also introduces a realistic radio irregularity model in order to provide a realistic test-bed for the wireless research. With the model, the research analyzes the real irregular routing properties as well as the potential effects of the model on the distributed localization algorithms. To improve the localization performance, the dissertation proposes a constrained-greedy forwarding radio propagation method to identify a regular routing path under complex operating environments. Through extensive simulation, the dissertation points out that although realistic (irregular) radio greatly degrades the performance of most localization algorithms, optimal localization is still possible if more anchors are deployed, and/or if the constrained-greedy forwarding radio propagation method is adopted.

BIBLIOGRAPHY

[1] N. Bulusu, J.Heidemann, and D. Estrin. GPS-less Low Cost Outdoor Localization for Very Small Devices. *IEEE Personal COmmunications Magazine*, 7(5):28–34, Oct. 2000.

[2] P. Bahl and V. Padmanabhan. RADAR: An In-Building RF-Based User Location and Tracking System. *Proc. IEEE Infocom 200*, pages 775–784, 2000.

[3] Yongguang Chen and Hisashi Kobayashi. Signal Strength Based Indoor Geolocation. In *Communications, 2002. ICC 2002. IEEE International Conference on*. Communications, 2002. ICC 2002. IEEE International Conference on, May 2002.

[4] Didier Chincholle, Mikael Eriksson, and Alex Burden. Location-sensitive services: it's now ready for prime time on cellular phones! In *Proceedings of the conference on Designing interactive systems*, pages 331–334. ACM Press, 2002.

[5] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki. Practical Robust Localization over Large-Scale 802.11 Wireless Networks. In *MobiCom*, Sept 2004.

[6] Dragos Niculescu and Badri Nath. VOR Base Stations for Indoor 802.11 Positioning. In *MobiCom'04*. ACM, Oct. 2004.

[7] Chris Savarese. Robust positioning algorithms for distributed ad-hoc wireless sensor networks. Master's thesis, Berkeley, 2002.

[8] A. Savvides, H. Park, and M. B. Srivastava. The n-hop multilateration primitive for node localization problems. *Journal of Mobile Networks and Applications (MONET)*, 8:443–451, 2003.

[9] Xiang Ji and Hongyuan Zha. Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling. In *IEEE INFOCOM*, 2004.

[10] Lingxuan Hu and David Evans. Localization for Mobile Sensor Networks. In *MobiCom'04*. ACM, Sept 2004.

[11] Dragos Niculescu and Badri Nath, editors. *Ad Hoc Positioning System (APS) using AoA*, San Francisco, CA, 2003. IEEE INFOCOM.

[12] Alex Hills, Jon Schlegel, and Ben Jenkins. Estimating Signal Strengths in the Design of an Indoor Wireless Network. *IEEE Trans. on Wireless Communications*, 3(1), Jan. 2004.

[13] P. Krishnan, A.S. Krishnakumar, Wenhua Ju, C. Mallows, and S. Ganu. A System for LEASE: System for Location Estimation Assisted by Stationary Emitters for Indoor Wireless Networks. In *Proceedings of IEEE Infocom 2004*, Hong Kong, March 2004.

[14] Ahmad Hatami and Kaveh Pahlavan. In_building Intruder Detection for WLAN Access. *Position Location and Navigation Symposium, 2004, PLANS 2004*, pages 592–597, April 2004.

[15] S. Meguerdichian, F. Koushanfar, G. Qu, and M. Potkonjak. Exposure in wireless ad hoc sensor networks. In *In International Conference on Mobile Computing and Networking (MobiCom'01)*, pages 139–150, Rome, Italy, July 2001.

[16] T. Yan, T. He, and J.A. Stankovic. Differentiated surveillance service for sensor networks. In *Proceeding of First ACM Conference on Embedded Networked Sensor Systems (SenSys 2003)*, Los Angeles, CA, 2003.

[17] M. Ancona, S. Locati, and A. Romagnoli. Context and location aware textual data input. In *Proceedings of the 2001 ACM symposium on Applied computing*, pages 425–428. ACM Press, 2001.

[18] Keith Cheverst, Nigel Davies, K. Mitchell, and A. Friday. Mobile-awareness: designing for mobile interactive systems. *ACM SIGGROUP Bulletin archive*, 22:8–11, April 2001.

[19] Alan Dix, T. Rodden, N. Davies, J. Trevor, A. Friday, and K. Palfreyman. Exploiting space and location as a design framework for interactive mobile systems. *ACM Transactions on Computer-Human Interaction (TOCHI), in press.*, 7(3):285–321, Sep. 2000.

[20] Eija Kaasinen. User needs for location-aware mobile services. *Personal Ubiquitous Computing*, 7(1):70–79, 2003.

[21] K. Cheverst, N. Davis, K. Mitchell, and A. Friday adn C. Efstreatiou. Developing a context-aware electronic tourist guide: some issues and experiences. In *Proceedings of the conference on Handheld and Ubiquitous Computing (HUC'00)*, 2000.

[22] D.C. Steere, A. Baptista, D. McNamee, C. Pu, and J. Walpole. Research challenges in environmental observation and forecasting systems. In *Proceedings of the sixth annual international conference on Mobile computing and neting*, pages 292–299. ACM Press, 2000.

[23] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao. Habitat monitoring: application driver for wireless communications technology. In *Proceedings of the ACM SIGCOMM Workshop on Data Communications*, pages 20–41, 2001.

[24] H. Wang, J. Elson, L.Girod, D. Estrin, and K. Yao. Target classification and localizaiton in habitat monitoring. In *Proceedings of the IEEE ICASSP 2003*, Hong Kong, April 2003.

[25] P. Boettcher, J. A. Sherman, and G. A. Shaw. Target localization using acoustic time-difference of arrival in distributed sensor networks. In *Proceedings of SPIE 47th Annual Meeting*, 2002.

[26] E. Howden. Networked sensors for the objective force. In *Proceedings of SPIE 47th Annual Meeting*, 2002.

[27] R.L.Moses, D. Krishnamurthy, and R. Patterson. An auto-calibration method for unattended ground sensors. In *Proc. ICASSP*, pages 2941–2944, May 2002.

[28] M.R. Perlman and Z.J. Haas. Determining the optimal configuration for the zone routing protocol. *IEEE Journal on Selected Areas in Communications*, 17(8):1395–1414, Aug. 1999.

[29] Y.B. Ko and N.H. Vaidya. Geotora: A protocol for geocasting in mobile ad hoc networks. *Proceedings of IEEE ICNP 2000*, pages 240–250, Nov. 2000.

[30] Chris Savarese, K. Langendoen, and J. Rabaey. Robust positioning algorithms for distributed ad-hoc wireless sensor networks. *Proceedings of the General Track: 2002 USENIX Annual Technical Conference*, pages 317–327, 2002.

[31] K. Premaratne, J. Zhang, and M. Doguel. Location Information-Aided Task-Oriented Self-Organization of Ad-Hoc Sensor Systems. *IEEE sensors Journal*, 4(1), Feb. 2004.

[32] M. L. Sichitiu, V. Ramadurai, and P. Peddabchagari. Simple algorithm for outdoor localization of wireless sensor networks with inaccurate range measurements. In *Proceedings of the International Conference of Wireless Networks*, Las Vegas, Nevada, USA, June 2003. CSREA Press 2003.

[33] Tian He, Chengdu Huang, Brian M. Blum, John A. Stankovic, and Tarek Abdelzaher. Range-free location schemes for large scale sensor networks. *MobiCom'03*, Sep. 2003.

[34] Radhika Nagpal, H. Shrobe, and J. Bachrach. Organizing a global coordinate system from local information on an ad hoc sensor network. *IPSN'03*, April 2003.

[35] Radhika Nagpal. Organizing a global coordinate system from local information on an amorphous computer. *A.I. Memo 1666*, Aug. 1999.

[36] D. Niculescu and B. Nath. Localized positioning in ad hoc networks. *IEEE SNPA 2003*, May 2003.

[37] S. Capkun, M. Hamdi, and J.-P. Hubaux. GPS-free positioning in mobile ad-hoc networks. *Cluster Computing*, 5(2), April 2002.

[38] Vaidyanathan Ramadurai and M. L. Sichitiu. Localization in wireless sensor networks: A probabilistic approach. In *Proc. of the 2003 International Conference on Wireless Networks (ICWN 2003)*, pages 275–281, Las Vegas, NV, June 2003.

[39] Gang Zhou, Tian He, S. Krishnamurthy, and J. A. Stankovic. Impact of Radio Irregularity on Wireless Sensor Networks. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, June 2004.

[40] K. Pawlikowski, H.-D.J. Jeong, and J.-S.R.; Lee. On credibility of simulation studies of telecommunication networks. *Communications Magazine, IEEE*, 40:132–139, Jan 2002.

[41] David Kotz, C. Newport, R. S. Gray, J. Liu, Y. Yuan, and C. Elliott. Experimental Evaluation of Wireless Simulation Assumptions. In *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, October 2004.

[42] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer-Verlag, 4th edition, 1997.

[43] W. Kumm. GPS Global Positioning System. *Bielefeld, Germany:Klasing*, 1995.

[44] Jacek Czajewski. The accuracy of the global positioning systems. *IEEE Instrumentation & Measurement Magazine*, 2004.

[45] Robin Strahan. Location sensing technologies. In *Report No. e-mc2.2.1.1.2002, Department of Computer Science, University*, 2002.

[46] web. http://www.tycoelectronics.com/gps.

[47] web. http://europa.eu.int/comm/dgs/energy_transport/galileo/index_en.htm.

[48] web. http://www.gps-practice-and-fun.com/a-gps.html.

[49] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The Active badge Location Systems. *ACM Transactions on Information Systems*, 10(1):91–102, Jan. 1992.

[50] R. Ward, A. Jones, and A. Hopper. A New Location Technique for the Active Office. *IEEE Personal Communications*, 4(5):42–47, Oct. 1997.

[51] A. Ward. *Sensor-driven Computing*. PhD thesis, Cambridge University, U.K., May 1999.

[52] N.B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. *Proceedings of MOBICOM'00*, Aug. 2000.

[53] Ascension technology corporation, http://www.ascension-tech.com.

[54] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer. Multi-camera multi-person tracking for EasyLiving. In *Third IEEE International Workshop on Visual Surveillance*, Dublin, Ireland, July 2000.

[55] Homayoun Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7), July 1993.

[56] D. Molkdar. Review on radio propagation into and within buildings. *Microwaves, Antennas and Propagation*, 138:61–73, Feb 1991.

[57] Theodore S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 1996.

[58] M. Hassan-Ali and K. Pahlavan. A new statistical model for site-specific indoor radio propagation prediction based on geometric optics and geometric probability. *Wireless Communications, IEEE Transactions on*, 1:112–124, Jan. 2002.

[59] Supachai Phaiboon. An Empirically Based Path Loss Model for Indoor Wireless Channels in Laboratory Building. In *Proceedings of IEEE TENCON'02*, 2002.

[60] M. Lott and I. Forkel. A multi-wall-and-floor model for indoor radio propagation. *Vehicular Technology Conference, IEEE*, 1:464–468, May 2001.

[61] Paramvir Bahl, V.N. Padmanabhan, and A. Balachandran. Enhancements to the RADAR User Location and Tracking System. Technical report, Microsoft Research, WA 98052, Feb. 2000.

[62] P. Prasithsangaree, P. Krishnamurthy, and P.K. Chrysanthis. On indoor position location with wireless LANs. In *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, pages 720–724, Sept. 2002.

[63] Santosh Pandey, B. Kim, F. Anjum, and P. Agrawal. Client assisted location data acquisition scheme for secure enterprise wireless networks. In *IEEE WCNC 2005*, 2005.

[64] Ankur Agiwal, Parakram Khandpur, and Huzur Saran. Locator: location estimation system for wireless lans. In *Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*, Oct. 2004.

[65] Andrew M. Ladd, Kostas E. Bekris, Algis Pl Rudys, Dan S. Wallach, and Lydia E. Kavraki. On the Feasibility of Using Wireless Ethernet for Indoor Localization. June 2004.

[66] Y. Oshman and P. Davidson. Optimization of observer trajectories for bearings-only target localization. *IEEE Transactions on Aerospace and Electronic Systems*, 35:892–992, 1999.

[67] D. Niculescu and B. Nath. Ad-hoc positioning system (APS). In *GLOBECOM 2001*, San Antonio, Nov. 2001.

[68] Andreas Savvides, Chih-Chieh Han, and Mani B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *MOBICOM*, Rome, Italy, 2001.

[69] A. Savvides, H. Park, and M. Srivastava. The bits and flops of the n-hop multilateration primitive for node location problems. *First ACM International Workshop on Wireless Sensor Networks and Application (WSNA)*, pages 112–121, 2002.

[70] Jeffrey Hightower and Gaatano Boriello. Localization systems for ubiquitous computing. *IEEE Computer Magazine*, 34(8):57–66, August 2001.

[71] Joshua A. Tauber. Indoor location systems for pervasive computing. Technical report, MIT, August 2002.

[72] Lance Doherty, Kristofer S. J. Pister, and Laurent E. Ghaoui. Convex position estimation in wireless sensor netowrks. *Proc. of IEEE Inforcom*, 3:1655–1663, 2001.

[73] A. Savarese, J. Rabaey, and J. Beutel. Locationing in distributed ad-hoc wireless sensor networks. *In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2037–2040, May 2001.

[74] X. Ji and H. Zha. Multidimensional scaling based sensor positioning algorithms in wireless sensor networks. In *Proceedings of the 1th Annual ACM Conference on Embedded Networked Sensor Systems*, pages 328–329, Nov. 2003.

[75] T. He, J. A. Stankovic, and T.F. Abdelzaher. Speed: A Stateless Protocol for Real-Time Communication in Sensor Networks. *Proceedings of IEEE ICSCS'03*, May 2003.

[76] Saad Biaz and Yiming Ji. A Survey and comparison on localisation algorithms for wireless ad hoc networks. *International Journal of Mobile Communications (IJMC)*, 3(4):374–410, 2005.

[77] Saad Biaz and Yiming Ji. Precise Distributed Localization Algorithms forWireless Networks. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, (IEEE WoWMoM 2005)*. IEEE, June 2005.

[78] K. Langendoen and N. Reijers. Distributed location in wireless sensor networks: A quantitative comparison. *Computer Networks*, 43:499–518, 2003.

[79] K. Kleinrock and J. Silvester. Optimum tranmission radii for packet radio networks or why six is a magic number. In *in National Telecommunications Conference*, pages 4.3.1–4.3.5, Birmingham, Alabama: IEEE, Dec. 1978.

[80] G. Young and A.S. Householder. A note on multidimensional psycho-physical analysis. In *Psychometrika, 6*, pages 331–33, 1941.

[81] L. Guttman. A new approach to factor analysis: the radex. in p. lazarsfeld (ed.). In *Mathematical thinking in the behavioral sciences*, New York:Free Press, 1954.

[82] T.F. Cox and M.A.A. Cox. *Multidimensional scaling*. London: Chapman Hall, 1994.

[83] I. Brog and P. groenen. *Modern multidimensional scaling: Theory and applications*. New York: Springer, 1997.

[84] Trevor F. Cox and Michael A.A. Cox. *Multidimensional Scaling*. Chapman and HallCRC, 2 edition, 2001.

[85] G. Young and A.S. Householder. Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3:19–22, 1938.

[86] I.J. Schoenberg. Remarks to maurice frechet's article "sur la definition axiomatique d'une classe d'espaces vectoriels distancies applicables vectoriellement sur l'espaace de hilbert". *Annals of Maththematics*, 36:724–732, 1935.

[87] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker. Complex Behavior at Scale: An Experimental Study of Low-Power Wireless Sensor Networks. Technical report, Technical Report UCLA/CSD-TR 02-0013, 2002.

[88] Zhong Ji, Bing-Hong Li, Hao-XIng Wang, Hsing-Yi Chen, and Tapan K. Sarkar. Efficient Ray-Tracing Methods for Propagation Prediction for Indoor Wireless Communications. In *IEEE Antennas and Propatation Magazine*, volume 43, April 2001.

[89] H. Kim and H. Ling. Electromagnetic scattering from an inhomogeneous object by ray tracing. *IEEE Trans. Antennas Propagat.*, 40:517–525, May 1992.

[90] Reinaldo A. Valenzuela. Ray tracing prediction of indoor radio propagation. In *Personal, Indoor and Mobile Radio Communications, 1994.* 5th IEEE International Symposium on Wireless Networks - Catching the Mobile Future., Sept. 1994.

[91] Chang-Fa Yang, Boau-Cheng Wu, and Chuen-Jyi Ko. A Ray-Tracing Method for Modeling Indoor Wave Propagation and Penetration. *IEEE Transaction on Antennas and Propagation*, 46(6), June 1998.

[92] A. Falsafi, K. Pahlavan, and G. Yang. Transmission Techniques for Radio LAN's - A Comparative Performance Evaluation Using Ray Tracing. *IEEE J. on Sel. Areas in Comm.*, 14(3):477–491, April 1996.

[93] M. Lott. On the performance of an Advanced 3D Ray Tracing Method. In *Proc. of European Wireless & ITG Mobile Communication*, Oct. 1999.

[94] Henry L. Bertoni, Walter Honcharenko, L.R. Maciel, and H. Xia. UHF propagation prediction for wireless personal communications. *Proc. IEEE*, 82:1333–1359, Sept. 1994.

[95] G. German, Q. Spencer, L. Swindlehurst, and R. Valenzuela. Wireless indoor channel modeling: statistical agreement of ray tracing simulations and channel sounding measurements. *Acoustics, Speech, and Signal Processing, 2001. Proceedings, IEEE Int. Conf. on*, 4, May 2001.

[96] N. Metropolis, A.W. Rosenbluth, M. N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1958.

[97] K.A. Dowsland. Simulated annealing. In C. R. Reeves, editor, *Modern Heuristic Techniques for Combinatorial Problems, chapter 2. McGraw-Hill Book Company, Berkshire*, 1995.

[98] S. Kirkpatrick, C. D. Gelatt, and M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598), May 1983.

[99] S.C. Johnson. Hierarchical Clustering Schemes. *Psychometrika, 2:241-254*, 1967.

[100] J. B. MacQueen. Some Methods for classification and Analysis of Multivariate Observations. In *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pages 1:281–297. Berkeley, University of California Press, 1967.

[101] Vijay Abhijit, Carla Ellis, and Xiaobo Fan. Experiences with an Inbuilding Tracking System. In *PIMRC*, 2003.

[102] Saad Biaz and Yiming Ji. Precise Distributed Localization Algorithms forWireless Networks. Technical Report CSSE05-02, Auburn University, March 2005.

[103] Robin Sibson. Studies in the Robustness of Multidimensional Scaling: Perturbational Analysis Classical Scaling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):217–229, 1979.

[104] R.A. Valenzuela, S. Fortune, and J. Ling. Indoor propagation prediction accuracy and speed versus number of reflections in image-based 3-D ray-tracing. *48th IEEE Vehicular Technology Conferernce*, 1:539–543, May 1998.

[105] S. Ganu, A.S. Krishnakumar, and P. Krishnan. Infrastructure-based Location Estimation in WLAN Networks. *IEEE Wireless Communications and Networking Conference (WCNC 2004)*, 1:465–470, March 2004.