

INDOOR POSITIONING TECHNIQUES AND APPROACHES
FOR WI-FI BASED SYSTEMS

by

Ayah Mahmoud Abusara

A Thesis Presented to the Faculty of the
American University of Sharjah
College of Engineering
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Electrical Engineering

Sharjah, United Arab Emirates

June 2015

Approval Signatures

We, the undersigned, approve the Master's Thesis of Ayah Mahmoud Abusara.

Thesis Title: Indoor Positioning Techniques and Approaches for WI-FI Based Systems

Signature

Date of Signature

(dd/mm/yyyy)

Dr. Mohamed Hassan
Associate Professor of Electrical Engineering
Thesis Advisor

Dr. Hasan Mir
Associate Professor of Electrical Engineering
Thesis Committee Member

Dr. Mamoun Abdel-Hafez
Associate Professor of Mechanical Engineering
Thesis Committee Member

Dr. Nasser Qaddoumi
Head of Department of Electrical Engineering

Dr. Mohamed El-Tarhuni
Associate Dean of College of Engineering

Dr. Leland Blank
Dean of College of Engineering

Dr. Khaled Assaleh
Interim Vice Provost for Research and Graduate Studies

Acknowledgements

Prima facie, I thank Allah for the endless blessings and the care that was poured out upon me through the people who have made this experience of mine smooth and fruitful.

Among these people is my advisor, Dr. Mohamed Hassan, who I am indebted to for his sincere guidance, support, and unceasing encouragement extended to me throughout the course of the last three years. I also acknowledge Dr. Aboelmaged Noureldin and Dr. Mohammed Maher Atia for sharing expertise that greatly assisted this research. In addition, I thank Dr. Mahmoud Ibrahim for lending a helping hand and for his valuable input.

I am also immensely grateful to my loving parents, who without their encouragement, support, and prayers, this thesis wouldn't have been possible. A special thank you to my sisters Sara and Sadeen and my brothers Baker, Ali, Hasan, Mohammed, and Zaid. I am also thankful to my grandparents and all my other family members. Thank you for being by my side all along the way.

I want to take the opportunity to thank my colleagues and friends for their consistent help and support. I specially thank Menatallah Shehabeldin and Rabiya Momin for being by my side throughout this experience, making it uniquely amusing. I am also grateful to Sanaa Hamid for her continuous guidance and encouragement.

I also place on record, my gratitude towards the Department of Electrical Engineering at the American University of Sharjah for providing me with the graduate assistantship and conducive conditions that were necessary to complete this thesis. I am thankful to all the professors who taught me graduate level courses that enhanced my knowledge. Many thanks also to Mr. Abusaif and Mr. Narayanan for helping me throughout my teaching experience.

To my family ...

Abstract

The rapid expansion of smartphones' market coupled with the advances in mobile computing technology has opened up doors for new mobile services and applications. Quite a few of these services require the knowledge of the exact location of their handsets. Although, existing global positioning systems (GPS) perform best in outdoor environments, they have poor performance indoors. This has initiated the need for a new generation of positioning systems. In this thesis, we focus on wireless local area networks (WLAN)-based indoor positioning systems to act as GPS counterpart indoors. More specifically, we study two received signal strength (RSS)-based positioning techniques, fingerprinting and propagation models. We shed light on the advantages of each technique and propose different methods to counteract their shortcomings. Namely, we propose a hybrid solution of clustering and fast search techniques to reduce the computational requirements of fingerprinting. In addition, we propose a dimensionality reduction technique to restrict the location fingerprints to signal strength values received from only informative Access Points (APs), hence to further reduce fingerprinting complexity. For this purpose, we implement a modified fast orthogonal search method to choose the most informative APs from the set of all hearable APs in the region. Finally, we propose an indoor localization system that integrates the RSS correction methods to enhance the positioning accuracy of propagation models. This proposed system aims to achieve accurate modeling of signals' propagation inside buildings without the need for expensive site surveys required for fingerprinting. Our experiments were conducted inside the engineering building at our university, using real RSS data. The obtained results show that the aforementioned first two proposed methods enhance fingerprinting techniques by reducing their computational complexity, while the third enhances the accuracy of propagation models.

Search Terms: *fingerprinting, propagation models, KNN, clustering, fast orthogonal search, Kalman filtering, Gaussian process regression*

Table of Contents

Abstract	6
List of Figures	10
List of Tables	11
List of Abbreviations	12
1. Introduction	13
1.1 Problem Statement	14
1.2 Contributions	14
1.3 Thesis Outline	15
2. Background	16
2.1 Wireless Positioning Systems	16
2.1.1 GPS-based	16
2.1.2 Cellular-based	17
2.1.3 WLAN-based	18
2.1.4 Bluetooth-based	18
2.2 WLAN Positioning System Components	19
2.3 WLAN Positioning Topologies	19
2.4 Wireless Positioning Algorithms	20
2.4.1 Triangulation	21
2.4.2 Proximity	23
2.4.3 Fingerprinting	23
2.5 Fingerprinting Localization Algorithms	25
2.5.1 Deterministic Algorithms-Nearest Neighbor	25
2.5.2 Probabilistic Algorithms-Maximum Likelihood	26
2.6 RSS Propagation Models	27
2.7 RSS Properties	27
2.8 System Performance Metrics	29
2.9 WLAN Positioning Systems Design Challenges	30

3. Literature Review	32
3.1 Fingerprinting Computational Complexity Reduction	32
3.1.1 Clustering and Search Strategies	32
3.1.1.1 Incremental Triangulation Clustering	32
3.1.1.2 k-Means Clustering	33
3.1.1.3 Fuzzy k-Means Clustering	33
3.1.1.4 The Sierpinski Triangle Search Strategy	34
3.1.2 Fingerprints Dimensionality Reduction	34
3.1.2.1 Principal Component Analysis	34
3.1.2.2 Fast Orthogonal Search	35
3.1.2.3 APs Significance Measures	35
3.2 Dynamic Propagation Models	36
3.3 RSS Correction Methods	36
3.3.1 Feature Extraction	37
3.3.2 Time Series Analysis	37
3.3.3 Kalman Filtering	37
3.4 Online Calibrated Radio Maps	38
3.5 WLAN-based positioning systems	38
4. Enhanced Fingerprinting using Hybrid Search Techniques	40
4.1 Overview of The Proposed Positioning Approach	41
4.1.1 Fingerprinting	41
4.1.2 Location-based Clustering	42
4.1.3 Fast Search Strategies	43
4.1.4 The Proposed Hybrid Search	45
4.2 Experimental Setup	47
4.3 Experimental Results	48
5. RSS Fingerprints Dimensionality Reduction	50
5.1 Overview of The Proposed Approach	51
5.1.1 Fast Orthogonal Search	52
5.1.2 FOS Implementation in Positioning System	54
5.1.2.1 Traditional FOS	54

5.1.2.2	Modified FOS	54
5.2	Experimental Results	56
6.	Error Reduction in Distance Estimation of RSS Propagation Models using Kalman Filters	61
6.1	Overview of Proposed System	62
6.1.1	RSS Measurements Correction	62
6.1.1.1	Kalman Filter Algorithm	62
6.1.2	Probabilistic Propagation Modeling	63
6.1.2.1	Gaussian Process Regression	64
6.2	The Proposed System	65
6.3	Experimental Results	67
7.	Conclusions	71
	References	73
	Vita	77

List of Figures

Figure 1:	Position estimation using GPS.	17
Figure 2:	Indoor positioning algorithms taxonomy.	20
Figure 3:	ToA technique for WLAN-based positioning.	22
Figure 4:	AoA technique for WLAN-based positioning.	23
Figure 5:	Proximity for WLAN-based positioning.	24
Figure 6:	k-Means clustering algorithm.	33
Figure 7:	Illustration of fingerprinting technique.	41
Figure 8:	Illustration of location-based clustering.	43
Figure 9:	Illustration of three step search.	44
Figure 10:	Illustration of orthogonal search.	44
Figure 11:	Illustration of diamond search.	45
Figure 12:	Proposed system model of the hybrid solution.	46
Figure 13:	Top view of the experimental area.	47
Figure 14:	CDF error curve of different clustering and search strategies.	48
Figure 15:	The algorithm to eliminate nonhearable APs.	54
Figure 16:	Flowchart representing the modified FOS procedure.	55
Figure 17:	Positioning error CDF of FOS and modified FOS.	57
Figure 18:	Positioning error CDF of MFOS compared to all techniques.	57
Figure 19:	Proposed system model of propagation models when integrated with Kalman filters.	66
Figure 20:	Illustration of propagation modeling for one AP in the experimental area.	67
Figure 21:	Distance estimation using GPR and Pathloss.	68
Figure 22:	GPR performance using different RSS correction methods.	69

List of Tables

Table 1:	Evaluation of the different clustering and search strategies.	49
Table 2:	Radio map for M gridpoints and N access points.	51
Table 3:	Processing time and complexity for different reduction methods.	58
Table 4:	Performance of different reduction methods for different levels of reduction.	58
Table 5:	Accuracy of clustering and search strategies with and without MFOS.	59
Table 6:	Computational requirements of clustering and search strategies with and without MFOS.	60
Table 7:	Accuracy of GPR distance estimation using different correcting methods.	69

List of Abbreviations

AoA Angle of Arrival.

AP Access Point.

CDF Commulitve Density Function.

FOS Fast Orthogonal Search.

GPR Gaussian Process Regression.

GPS Global Positioning System.

kNN k-Nearst Neighbor.

LoS Line of Sight.

MSE Mean Squared Error.

NIC Network Interface Card.

NLoS None Line of Sight.

PCA Principal Component Analysis.

RF Radio Frequency.

RSS Recieved Signal Strength.

ToA Time of Arrival.

WLAN Wireless Local Area Networks.

Chapter 1: Introduction

Indoor positioning has gained a remarkable interest in the last few years. The increasing human activities inside buildings demanded indoor localization that cannot be achieved solely by the existing Global Positioning Systems (GPS). The reason is that GPS has poor indoor positioning coverage due to the poor penetration of GPS signals through construction materials and the lack of line of sight conditions indoors [1]. Therefore, research efforts are dedicated to find GPS alternatives that can provide seamless indoor and outdoor positioning coverage. The motivation behind these efforts is the demand for ubiquitous location-based services coupled with the increasing trend of mega constructions. Indoor positioning promises a wide range of services and applications, such as autonomous object tracking, impaired vision aid, asset tracking, etc. In addition, indoor positioning contributes greatly to the development of applications on smart phones, enabling context aware applications. According to a recent project established by Google Company [2], 3-D space sensing is one of the features intended for future smart phones.

Various systems have been investigated in the literature to support indoor positioning. This includes infrared-based systems, cellular-based systems, WLAN-based systems, etc. [3]. However, wireless local area networks (WLAN) are the most widely investigated systems to be the indoor counterparts of GPS. Although positioning is not the main application of WLAN, using these networks for indoor positioning is promising due to their ubiquitous availability inside buildings. In addition, using WLAN for indoor positioning is an inexpensive choice for two reasons. First, WLAN offers a stand-alone positioning capability, as it does not need any extra hardware. Second, WLAN is a universal and wide-spread technology, that is nowadays deployed in almost every building, hence no further infrastructure associated cost is incurred [1]. Not only will equipping WLAN with positioning capabilities help provide indoor positioning coverage, but also it will provide WLAN with an important safety public feature [4]. This is in response to the order issued by the federal communication commission in 1996 to enhance the positioning capabilities of WLAN, since it is very difficult to track the exact location of a user in WLAN compared to other Networks [5]. WLAN-based

indoor positioning can help detect the exact location of emergency callers, as well as intruders. In this thesis, wireless local area network (WLAN)-based indoor positioning is investigated. Specifically, Wi-Fi received signal strength (RSS) indoor models are improved and fingerprinting based localization algorithms and techniques are proposed.

1.1. Problem Statement

The complex structure and the dynamic nature of indoor environments pose several challenges to the design of indoor positioning systems [6]. These challenges, which are discussed in more details in Section 2.9, prevent achieving the sub-meter level accuracy required by the majority of indoor positioning applications. In addition, many difficulties involved in the design of indoor positioning systems are intrinsic to WLAN [4]. For example, in a typical WLAN-based positioning system, the computations of the mobile user location take place on a battery-powered device with limited processing capability and limited power supply. Consequently, using algorithms of low computational complexity is mandatory for indoor positioning. Moreover, WLAN-based positioning systems should be designed to concurrently support a large number of users to be compatible with the huge user base of WLAN. Lightweight algorithms accelerate the location finding process and hence enable accommodating a larger number of users. In this thesis, we are mainly concerned with two design aspects of indoor positioning systems; accuracy and complexity. Our objective is to balance the two conflicting goals of high accuracy and low computational complexity in the design of the cost effective WLAN-based positioning techniques.

1.2. Contributions

In this section we explain the major contributions of this thesis.

- A new search method is proposed and implemented that improves the search process for the location best match in fingerprinting. The method enables doing a faster search with better accuracy performance than the conventional clustering and search techniques. The method is based on a hybrid solution of location-based clustering and fast search strategies.

- A new implementation of a fast orthogonal search (FOS) is proposed to measure an Access Point (AP) contribution to the positioning system. The method helps to find the most informative APs in the region and to restrict the radio map to only those points. The modified implementation of FOS was found to result in a better performance than the original implementation.
- The modified FOS algorithm is combined with the hybrid search methods which resulted in an ultimate reduction in the computations of fingerprinting. In addition, a better positioning performance of the hybrid search methods is achieved when the latter is combined with the modified FOS algorithm.
- A system that integrates dynamic propagation models with RSS correction techniques is also proposed to enhance the accuracy performance of the propagation models, as they are a cheaper alternative to fingerprinting. Kalman filters are used to eliminate the environmental noise contaminating the RSS measurements hence improving the performance of the positioning.

1.3. Thesis Outline

The rest of the thesis is organized as follows: Chapter 1 presents the significance and challenges of this research field, as well as the objective and contribution of this thesis. In Chapter 2, the background to this research topic is presented. The latest advances, and the progress achieved in this field so far, are stated in Chapter 3. The methods and positioning approaches proposed by this thesis and their evaluation are explained in Chapters 4, 5 and 6. Finally, Chapter 7 outlines the conclusion and the future work.

Chapter 2: Background

2.1. Wireless Positioning Systems

There are several potential wireless localization systems, besides WLAN-based systems, each of which has its advantages and shortcomings. These systems can be categorized into systems with a non-dedicated infrastructure, or with a dedicated infrastructure [7]. Systems with a dedicated infrastructure are advantageous for having wide specification options and a control over the quality of positioning, while systems with non-dedicated infrastructures are restricted to certain specification and hold less control over the quality of positioning. Despite this, we choose systems with non-dedicated infrastructure because establishing a dedicated infrastructure, only for indoor positioning, has a substantial cost. This cost is a result of a required frequency band dedicated for the new infrastructure to operate on, base stations to provide geographical coverage, new built-in mobile devices' hardware that corresponds to the new infrastructure, and so on [1]. Therefore, systems with non-dedicated infrastructures are still preferable, due to their low cost of implementation and their resources efficiency [8]. A list of systems that utilize existing infrastructure is examined below:

2.1.1. GPS-based

Global positioning system (GPS) was proposed by the US Department of Defense in the 1970s. It is the most accurate available positioning system in outdoor environments. GPS consists of three segments; space, ground, and user segments. The space segment is a constellation of 24 satellites distributed on 6 orbital planes surrounding Earth. Such structure allows at least 4 satellites to be visible anywhere on earth at all times. The ground segment consists of stations to control and monitor the satellites movements, as well as the satellites clocks. The user side on the other hand, consists of all GPS enabled receivers that receive GPS signals for location determination by employing the concept of signal time of arrival (ToA) [9]. The distance of the user receiver from each visible satellite and each satellites longitude, latitude, and altitude are used to estimate the receiver's position [10]. As shown in Figure 1, the user location is estimated as the point of intersection of the satellites signal propagation spheres, using

lateralization techniques. Consequently, a minimum of 4 satellites are needed for location determination, three for location calculation and a fourth one for time synchronization purposes [9]. However, as mentioned before, GPS require LoS conditions which makes it impractical for indoor positioning. Furthermore, satellite signals might undergo atmospheric delays according to weather conditions, which highly affect the distance calculations accuracy [9]. Therefore, network assisted GPS is suggested to compensate for the poor indoor GPS coverage, and GPS atmospheric layer delays.

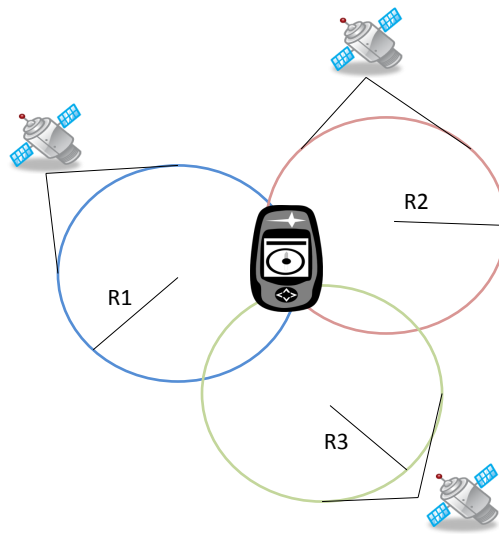


Figure 1: Position estimation using GPS.

2.1.2. Cellular-based

Cellular network is a wireless communication network that covers geographical areas to provide communication for mobile units. Areas that have cellular service are divided into small sections called cells. Each cell is monitored by one or two cellular transceivers called base stations. The working principle of a cellular network is that it tracks callers, and dedicates channels for calls through the base stations closest to the users. The network has a mobile switching center (MSC) that is used to handover channels as the user moves from one cell area to another [11]. Thus, a cellular network

user can be located using the Cell-ID and the geographical sector that are serving it. Such that, the location of the mobile unit is estimated to be the centre of the serving cell. This positioning technique is called proximity, and is discussed further in section 2.4.2. The advantages of using Cell-ID method is that this method is already in use nowadays, and it is unified for all mobile phones [8]. However, this method has very poor accuracy performance, since the coverage area diameter of a cell ranges from 2 Km to 20 Km [7]. Some researchers suggest using this network to enhance GPS performance in urban areas [9].

2.1.3. WLAN-based

Wi-Fi is a technology that provides network access for wireless enabled devices. This technology is governed by a family of standards established by The Institute of Electrical and Electronics Engineers (IEEE). The most popular standard amongst this family is the 802.11b which operates on 2.4 GHz with a data rate of 11 Mbps, depending on the standard [12]. Positioning can be simply implemented in WLAN using RSS-based techniques or triangulation techniques. Moreover, WLAN have a standardized 50 m range, which makes it appropriate enough for indoor positioning [13]. In addition, the costs for the positioning system components are modest, which makes this system an inexpensive choice [13].

2.1.4. Bluetooth-based

Bluetooth is wireless technology with positioning capabilities, it was introduced by Ericsson Company. It is an ad hoc network that functions on the 2.4 GHz Industrial, Scientific, and Medical (ISM) band. This technology is governed by IEEE 802.15 standards which belongs to wireless personal area networks (WPAN). Bluetooth is functional only on a room or hall level, with a range of 10 to 15 m, and a bit rate of 1 Mbps. The main use of Bluetooth technology is to connect devices of different functions, such as tablets, mobile phones, and personal computers [11]. Not only is this technology ubiquitous and embedded in almost all mobile devices, it also has lighter standards than WLAN. Positioning can be achieved using Bluetooth tags, which are small transceivers with unique IDs. In a similar manner to WLAN, the strength of the Bluetooth signal can be used to find the position of a Bluetooth enabled device relative to another [8].

2.2. WLAN Positioning System Components

The positioning system architecture consists of the following units:

- **Network:** The underlying wireless network, IEEE 802.11, utilizes radio frequency (RF) signals to establish communication between the users of the same wireless local area network (WLAN) and the network access points (APs) [3]. The network APs are configured to broadcast beacon packets that contain different types of system information. These packets are transmitted every 100 ms or so. Any network can be used for positioning, even the closed ones, since the beacons are not encrypted and can be received by any WLAN-enabled device [1]. Each AP has a unique MAC address with a known fixed location [9].
- **Communication medium:** In WiFi standards, 11 channels are specified for the 2.4 GHz. Typically, 3 non overlapping channels are occupied at the same time in the same geographical area. If interference is expected, APs can transmit beacon packets on different channels. The mobile devices scan all the channels to find the APs.
- **Mobile unit:** Mobile devices that are WiFi enabled can receive the beacon packets and the system information conveyed on them are extracted using the so-called network interface cards (NICs). The NIC can measure the RSS value from the packets. NIC can measure all the RSS values from all the hearable APs at the location of the mobile handset [1].
- **Server:** This unit is optional. It is implemented in mobile-assisted systems. The location information are transmitted through a TCP/IP link to a central unit, where the position calculation are done [14].

With this system breakdown, the position of a WLAN mobile user can be estimated using the signals it receives from the fixed APs in the surrounding region [1].

2.3. WLAN Positioning Topologies

The system topology is the method by which the system components are arranged and distributed to achieve their intended tasks. For indoor positioning systems, there are three main positioning topologies. The first one is remote positioning, whose

localization units are fixed receivers that sense the transmitted signal from a mobile device and remotely calculates its position. It is the opposite of self positioning, where the mobile target estimates its own position based on the signals it receives from fixed transmitters. If a wireless data link is available in the system, the mobile device can sense the signals from the transmitters and then communicates the measurements to a central unit to estimate the mobile device position. This is the third topology which is called hybrid positioning [1].

2.4. Wireless Positioning Algorithms

Indoor positioning algorithms are divided into three main categories: triangulation, proximity, and scene analysis. Figure 2 summarizes all the techniques used in indoors location estimation [12]. Each of these techniques is discussed in details below:

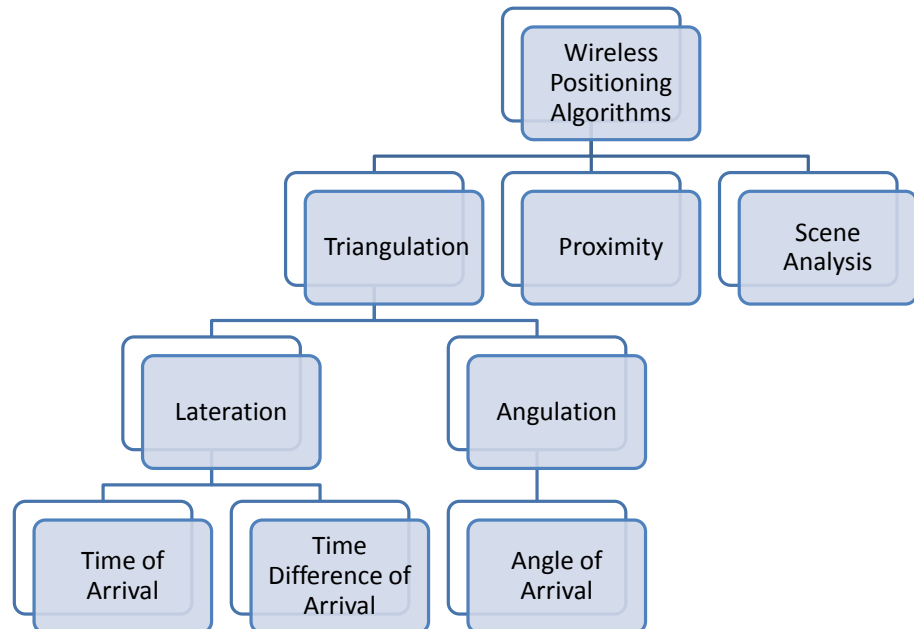


Figure 2: Indoor positioning algorithms taxonomy.

2.4.1. Triangulation

In this method, geometry principles along with signal parameters are used to estimate the position of a mobile unit. Triangulation technique is further divided, based on the used signal parameter, into two subcategories; lateration and angulation [8]. In the lateration technique, distance measurements from fixed reference points (Access points), are used in to locate a mobile device. The most famous lateration technique is known as time of arrival (ToA). This technique utilizes the time that the signal takes to travel from a reference point to the receiver of the target to be detected. The distance traveled by an RF signal is directly proportional to its propagation time, given that an RF signal propagates with the speed of light. Therefore, the velocity of the transmitted signal and its travel time are used to estimate the distance between the transmitter of the reference point and the receiver of the mobile unit [9]. When using lateration techniques, at least three reference points are required to achieve two dimensional (2-D) positioning, which is positioning in reference to the xy plane. Having three access points (APs) available, the xy coordinates of a mobile device position is estimated, as shown in Figure 3. ToA is used to estimate the distance between each AP and the mobile user, such that the calculated distances serve as the radii of the three propagation circles of the signals transmitted from the APs. Finally, the location of the user is estimated to be the intersection point of the three propagation circles [7].

Although the lateration technique is used widely in GPS, it has poor performance indoors. The reason is that it requires line of sight (LoS) conditions, which is hardly the case in indoor environments. Moreover, ToA technique requires time stamps to indicate the departure time of the signal from the transmitter, and the arrival time of the signal to the receiver. It also requires time synchronization between the base stations and the mobile device. These two requirements introduce less system hardware flexibility, as well as limitations in the type of mobile device to be used [12]. Fortunately, synchronization limitation can be eliminated using the time difference of arrival (TDoA) method. In TDoA, the differences in time of arrival of the signals from different APs at the receiver are used as distance measurements, instead of using the absolute travel time [8]. However, this technique still assumes LoS condition, which makes it inappropriate for indoor positioning.

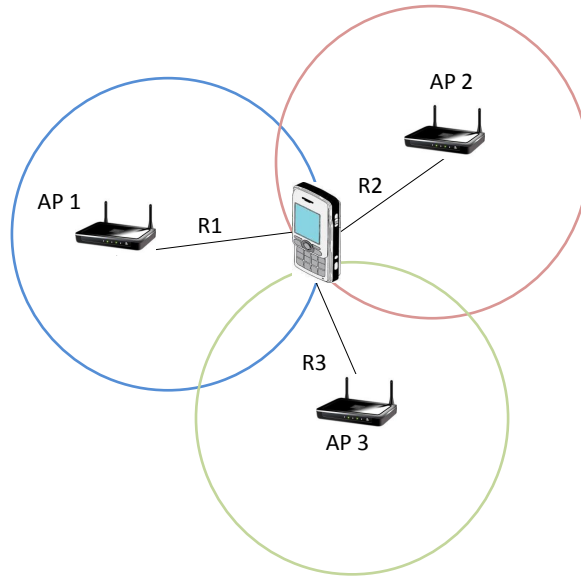


Figure 3: ToA technique for WLAN-based positioning.

The second triangulation technique is angulation. In this method the angle of arrival (AoA) of the signal is used to estimate the position of the target. The 2-D position of the mobile unit is the intersection point between the lines of bearing or the direction lines at which the signals arrive from the reference points, as in Figure 4. In this case, only two reference points or measuring units are needed to help estimate the position of the target. One advantage of this method is that the system can be easily extended to find 3-D positions, which is achieved only by adding a third reference point to the system. Moreover, this method does not require any synchronization between the transmitters of the APs [8]. On the other hand, AoA method requires more complex hardware than ToA and TDoA, as it requires an array of antennas to determine the angle at which the signal arrives at the receiver [10]. Furthermore, AoA technique requires the LoS condition which is almost not available indoors due to the harsh propagation environment. In both triangulation techniques, the system always assumes that the transmitted signal is in direct LoS with the receiver. Namely, the triangulation system does not differentiate between a LoS signal and a None Line of Sight (NLoS) signal. Therefore, it results in erroneous distance estimation.

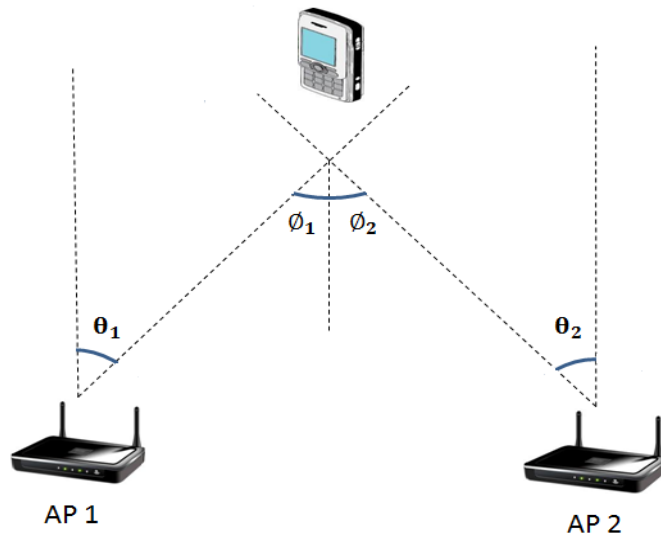


Figure 4: AoA technique for WLAN-based positioning.

2.4.2. Proximity

This algorithm collocates the target with the antenna or the reference point that is closest to it, as depicted in Figure 5. For this case, a dense grid of antennas, with known coordinates, is used. The mobile unit location is assigned to be the coordinates of the antenna that receives the strongest signal from the mobile device [8]. This algorithm is characterized by its simplicity and that it can be implemented on different physical media, such as radio frequency identification (RFID) and base stations (Cell-ID) [8]. The drawback of this method, as previously mentioned in Section 2.1.2, is that it has low resolution and hence poor accuracy.

2.4.3. Fingerprinting

Fingerprinting, which is also referred to as scene analysis, is the algorithm in which locations in an environment are associated with a unique signal parameter to that location. Any signal feature that is location dependent can be used, such as the multipath structure per location, the coordinates of the AP with the strongest signal per location, or the strength of the signal received (RSS) by the user per location [1].

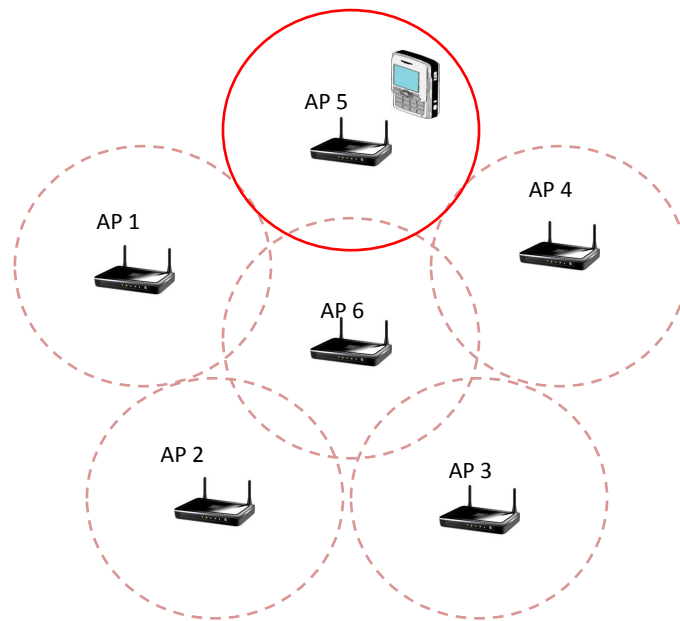


Figure 5: Proximity for WLAN-based positioning.

The most commonly investigated is the RSS-based fingerprinting, which uses the RSS to fingerprint locations in the region of interest. This method is performed in two stages:

- **Offline stage:** This is a calibration stage where fingerprinting actually occur. During this phase, the RSS values received from the visible AP in the region are collected at known grid points, forming an RSS model for that area [7]. In the case where more than one AP are visible, the fingerprint becomes multidimensional. The collected fingerprints are then saved to a database called a radio map [1]. The pre-stored radio map that results from the offline stage is used in the online stage to estimate the user location.
- **Online stage:** A positioning algorithm is used in this stage to match the RSS online measurement of the mobile target with the prestored radio map for that area [8]. There are many positioning techniques that can be used to estimate the position of the mobile user. These methods are implemented and tested in section 2.5.

The RSS-based fingerprinting method have been proven to overcome other methods in terms of accuracy; it results in the highest positioning accuracy indoors. The rea-

son is that in fingerprinting techniques the multipath and fading contaminating the RSS measurements are reflected on the real RSS values recorded in the radio map. Therefore, fingerprinting is considered as an effective method to handle the disturbance in RSS measurements. Furthermore, fingerprinting techniques are of a low cost of implementation compared to the triangulation techniques, since they utilize the existing WLAN infrastructure. For these reasons, fingerprinting techniques are employed by many WLAN-based positioning systems [15], [16]. It is also for these reasons that this thesis is employing fingerprinting techniques.

2.5. Fingerprinting Localization Algorithms

As described in the previous section, localization algorithms are needed to determine the mobile device position during the online phase. There are two types of fingerprinting positioning algorithms; deterministic algorithms and probabilistic algorithms. The first type employs feature matching between the online measurements and the stored training data i.e. Radio Map. The second type uses feature probability distributions to find the grid point with the maximum likelihood of resulting in the online measurements. It is not the objective of this thesis to list all the existing algorithms, only the most popular examples of each type are explained bellow [1].

2.5.1. Deterministic Algorithms-Nearest Neighbor

To demonstrate deterministic positioning algorithms, k -nearest neighbor (KNN) is reviewed in this section. KNN is a simple deterministic positioning algorithm [17] that employs feature matching to find the closest location points in the radio map (neighbors) to the target. This algorithm sets the RSS vector measured by the target device against the n location fingerprints saved in the database to find the k closest points to the target, where $2 \leq k \leq n$. The Euclidean distance $d(i, t)$ between the target t and the pre-stored location points in the radio map is calculated as:

$$d(i, t) = \sqrt{\sum_{z=1}^M (RSS_{iz} - RSS_{tz})^2}, \quad (1)$$

where RSS_{iz} is the RSS measurement received from the z th AP at the i th location point in the radio map. The xy position of the target are found by averaging the x and y coordinates of the k closest fingerprints to the target.

2.5.2. Probabilistic Algorithms-Maximum Likelihood

In probabilistic algorithms, the maximum a posteriori (M.A.P) principle is applied to find the location of the target mobile device. To be able to apply this principle, the joint distribution of the RSS from all the APs visible to a certain grid point, is required. In other words, the radio map for this case is composed of the RSS joint probability density functions (PDF) from each AP for each location point in the map. Unfortunately, the correlation between the RSS values from different APs at a certain location is not clear. However, researchers still assume the PDFs from different APs at certain location points to be independent to simplify the analysis. Therefore, the joint PDF can be simply obtained by the product of the individual RSS PDFs from each AP [1]. During the online phase of this algorithm, the location of the target is approximated to be the fingerprint that has the maximum probability of resulting in the RSS vector. Such probability is found using Bayes rule, where the conditional probability of a certain grid point corresponding to the location of the target, given the observed RSS vector, is given by:

$$\max_m (P\{p_m|r\}) = \max_m \left(\frac{P\{r|p_m\}P\{p_m\}}{P\{r\}} \right). \quad (2)$$

The grid point that maximizes the above probability is approximated as the position of the target [1]. This optimization problem can be simplified, since the probability of an observed RSS vector is constant for all grid points. Moreover, the grid points are assumed to be equiprobable. Thus, the problem can be simplified to the maximum likelihood problem given by:

$$\max_m (P\{p_m|r\}) = \max_m (P\{r|p_m\}). \quad (3)$$

2.6. RSS Propagation Models

Propagation modeling is another RSS-based positioning technique that is considered as a cheaper alternative to fingerprinting. Such models are capable of replacing the radio maps employed in fingerprinting with theoretical formulas that express the space propagation of an RF signal. Therefore, propagation models do not require the expensive site surveys needed by fingerprinting to calibrate radio maps. The simplest model according to which a radio signal propagates in free space, is the path loss model. In this model, the power density of a radio signal logarithmically decays with the distance traveled from the transmitter, this model is given by:

$$P_r = P_0 - 10n \log_{10} \left(\frac{d_r}{d_0} \right), \quad (4)$$

where P_0 is the initial transmitted power at a reference distance given by d_0 , d_r is the distance between the transmitter and the receiver, n is the path loss exponent which is an environment-dependent parameter. The path loss exponent, in open space, is known to be $n = 2$ while in indoor environments it takes a value between $n = 4 - 6$. On the other hand, the loss exponent in indoor environments can be less than $n < 2$ in narrow areas and corridors, as such narrow structures can act as waveguides [18]. Having these parameters and the measured RSS, the distance from the receiver to the transmitter can be extracted and the position of the device can be estimated using triangulation techniques. However, even if indoor losses are included in the formula, this model unfortunately does not hold in indoor environments. This is due to the fact that indoor environments are characterized by dense multipath and NLoS conditions, which cause the RSS power patterns to deviate from the path loss model. This results in high distance errors, causing the path loss model to be incapable to stand alone for indoor position estimation [1].

2.7. RSS Properties

Many factors, beside the distance from the AP, appear to have an effect on the power levels indicated by the RSS measurements. These factors are mainly divided into

two groups, hardware-related and environment-related factors. The hardware-related factors are due to the method by which RSS is reported by WLAN cards. In WLAN, the RSS measurement that is visible to the receiving device is not the actual received strength, rather it is an indication of the strength level. In essence, this RSS indicator is the average of the RSS values observed over a certain sampling period [18]. Different WLAN vendors, which support different implementation of hardware and chipsets, are expected to result in different RSS measurements at the same location [19]. The reason is that the RSS measurement can vary substantially depending on the quality of the implemented receiver and the type of receiving antennas. In addition, RSS values are measured differently by different vendors. For instance, the power levels are usually specified between 0 and a maximum level known as RSSI-Max. This RSSI-Max varies from a vendor to another, for example the RSSI-Max specified by Cisco is based on 100 levels while Atheros chipset bases its RSSI-Max on 60 levels [19]. Moreover, the quantization step of the real RF signal can differ depending on the used WLAN card. In general, the card that uses more quantization levels result in a better positioning performance since it provides a better representation of the real signal. On the other hand, there are environment-related factors that affect RSS measurements, such as building materials, user's presence and orientation and environmental noise. The effect of environmental noise on RSS measurements is typically modeled as a combination of the small scale fading and the large scale fading. The large scale fading is owed to the signal attenuation due to the absorption of the signal by large structures and it has a log-normal distribution. Whereas, the small scale fading is a result of multipath and it has a Rayleigh distribution under NLOS conditions and a Rician distribution otherwise [20]. It is worth mentioning, that those models are used for communication purposes, mainly to model the Wi-Fi coverage indoors. However, the aforementioned models cease to fit real RSS data resulting in these models to be insufficient for positioning. The effect of human presence on the RSS signals is mainly attributed to the absorption of the RF signals by the water molecules constituting the majority of human bodies [20]. In addition, the wavelength of the Wi-Fi signal is much smaller than the average human trunk, causing the signals to diffract around the human bodies resulting in further losses [13]. The orientation of the measuring device in reference to the AP has also a major effect on

the RSS values. This problem can be avoided using WLAN cards with omni-directional antennas.

2.8. System Performance Metrics

Although the accuracy of positioning systems is the most important metric, it is not sufficient to assess the overall performance of such systems. Therefore, researchers have established certain performance criteria by which positioning systems can be evaluated and compared against each other [8]. The most important metrics are defined below:

- **Accuracy:** The accuracy of a positioning system is usually specified by the location error. Namely, it is the mean distance error between the estimated and the true position. This performance metric is highly desirable, but typically it is acquired at the expense of other metrics
- **Precision:** the precision of the system measures the coherence and consistency of the results. Such metric can be calculated in different ways. The most popular way is the distance error standard deviation. Small standard deviation indicates high precision, thus the smaller the result error deviation from the mean value the better the system performance.
- **Complexity:** The complexity of the system can be investigated in terms of hardware and software. In this thesis, only software complexity is considered, which translates into the computational complexity of the positioning algorithm. This metric is important, especially when the computations are performed on a mobile device, where the processing capability and the power supply are limited. Unlike previous metrics, there is no specific method to measure the complexity of the system. Fortunately, it can be attributed to the computational time of the algorithm. Hence, the system that takes less time to complete the computations is considered of lower complexity.
- **Robustness:** Robustness is the ability of the system to resume functioning under perturbations. That is in the case of a node failure or the addition of a node to the system. The system is preferred to be persistent under unusual circumstances.

- **Scalability:** Scalability in principle is similar to robustness, but it deals with positioning scope changes. Such changes can be on a geographical scale or a density scale. The first scope change is associated with the changes in the system coverage area, whereas the second is related to the changes in the number of positioning units per area. It is very often that the systems accuracy degrades when the distance between the locator and the target is increased. The same applies when the number of units (positioning nodes) per area is reduced. The positioning system must ensure enough number of localization units to accommodate for increasing the space domain, provided that the number of units does not exceed the required quantity. Otherwise, increasing the number of units/area will result in communication channels' congestion, more complex computations, and obviously higher cost.
- **Cost:** There are many underlying factors other than the financial cost that govern the overall cost of the positioning system. These include time, space and energy costs. The space cost is related to the density of units per area, as explained previously. The time cost is related to the time needed for installation and maintenance. Finally, the systems energy consumption is also classified as a cost, and it is required to be fair, especially for mobile devices.

2.9. WLAN Positioning Systems Design Challenges

As previously mentioned, WLAN-based positioning systems are potential candidates to be the indoor counterparts of GPS. In such systems, the strength of the signal received by the mobile unit can be used as a measure of distance. However, positioning is not the main application of Wi-Fi technology. Thus, some challenges arise when using WLAN infrastructure for localization. These challenges are discussed below:

- **Temporal variation of RSS:** RSS measurements are usually contaminated with environmental noise causing the RSS values to fluctuate even at the same location [9]. The environmental noise results from the NLoS conditions and the dynamic environmental variations characterizing the indoor environments. The NLoS conditions are considered the main source of errors in localization sys-

tems [21]. To elaborate on this, the NLoS conditions occur when the RF signals travel through obstructed paths causing them to get reflected, refracted, and diffracted by obstacles. As a result, different rays of the same signal, arriving from different paths, add up with the original signal at the receiver in a phenomenon called multipath. This prevents raw RSS measurements from being a reliable distance measure. On the other hand, the dynamic environment variations are introduced inside buildings due to moving objects, and in general, due to people's presence [12]. Consequently, the multipath pattern is continuously altered making it very difficult to be stochastically modeled and compensated for [9].

- **Interference:** Wi-Fi operates on the 2.4 GHz band which is an unlicensed band that is also used by Bluetooth technology, microwave ovens, and cordless phones. These devices cause interference to mobile devices, hence affecting the RSS measurements and position estimation results [4]. In addition, water molecules are sensitive to the 2.4 GHz frequency, and since 70 % of the human body is composed of water, people's presence and movement inside buildings highly deteriorate the signal strength, and hence, affect the RSS model [12].
- **Latency and throughput:** A mobile device takes some time to scan for Wi-Fi signals, coupled with the time needed by the system to provide location information, is considered a source of latency that becomes crucial if there are many users' positioning requests. In addition, the continuous scanning for Wi-Fi signals if the user is moving causes data flow interruptions which in turn results in a degraded throughput [1].

Chapter 3: Literature Review

This chapter outlines related previous work and the literature that covers certain aspects of RSS-based positioning systems investigated in this thesis. It also discusses the latest advances in the indoor positioning field and some existing WLAN-based systems.

3.1. Fingerprinting Computational Complexity Reduction

Fingerprinting based algorithms require exhaustive matching between the RSS measured by the mobile target and the pre-stored RSS samples to find the best match. This process is computationally expensive. To reduce the number of needed operations by fingerprinting to locate the user, clustering and search strategies are employed. In addition, the dimension of a location fingerprint can grow to include all the hearable access points in the region. However, not all the APs actually contribute to positioning; the majority are just redundant. Therefore, including all APs in the positioning system results in a superfluous computational cost, if not also, a deterioration in the positioning accuracy.

3.1.1. Clustering and Search Strategies

The complexity of fingerprinting algorithms can be reduced by minimizing the number of operations needed to find the best match. For the purpose of reducing the computational complexity of fingerprinting, different clustering techniques are proposed in the literature [15, 22–24]. Clustering is defined as the process of grouping the elements in a data set according to some feature called cluster key, such that the elements in the same cluster show more resemblance to each other than to the elements outside the cluster. There are various techniques with which clustering can be performed; all vary with respect to the feature that classifies the clusters.

3.1.1.1. Incremental Triangulation Clustering

Incremental Triangulation (IT) clustering technique is employed in [15], where the fingerprints that share a common set of visible APs constitute a cluster. In other words, the cluster key in this case is the set of APs that cover the location points. The

multi-level clustering approach proposed by [15] works as follow: at first, the system takes the first AP visible to the target mobile device and scans for the pre-identified clusters that contain this AP. Then, it takes the second AP visible to the target and scans for the clusters that are covered by the first and second APs, and so on.

3.1.1.2. *k*-Means Clustering

k-Means clustering is another clustering technique that is used in [22]. This technique uses minimum distance to assign elements to *k* clusters in the data set iteratively. Firstly, *k* clusters are initialized with *k* centre points, and then the points in that data set which are closest to a certain center are assigned to the cluster which that center belongs to. After that, the centroids for each cluster are recalculated, and the pervious steps are repeated to find the new centroids. The same steps are repeated until there are no changes in the centre points of each cluster. The algorithm for *k*-means clustering is described below:

k-Means Clustering

- 1: Given *n* data points to be clustered
 - 2: **begin** initialize *k* centers
 - 3: **do** classify *n* points to nearest clusters
 - 4: recalculate the centers for each cluster
 - 5: **until** there is no change in the centers
 - 6: **return** new *k* centers
 - 7: **end**
-

Figure 6: *k*-Means clustering algorithm.

3.1.1.3. *Fuzzy k*-Means Clustering

On the other hand, fuzzy *k*-Means clustering is proposed in [23] to reduce fingerprinting complexity. This technique relies on fuzzy logic to assign elements to clusters, where one data point can be assigned to more than one cluster. Namely, each data element has a sort of fuzzy membership to all *k* clusters, such that the membership is expressed in a form of a probability [25]. This technique is used to minimize the

following objective function:

$$J = \sum_{i=1}^n \sum_{j=1}^k u_{ij} d_{ij}^2, \quad (5)$$

where n is the number of data points, u_{ij} is the degree of membership of element i in cluster j and d_{ij} is the distance between the data point i and centre of cluster j . Clustering is carried out iteratively until Equation 5 is minimized.

3.1.1.4. The Sierpinski Triangle Search Strategy

The Sierpinski triangle search strategy is implemented in [26] to optimize the fingerprints matching process in a Wireless Sensor Network (WSN). The working principle of this strategy is to divide the search region into sub-areas, and find the closest sub-area to the new measurement. As though, polygonal regions are covered by reference nodes placed exactly at the vertices. Then, each region is divided to 4 triangles and each triangle is divided further to 4 sub-triangles. During localization, the search for the target location is restricted to the sub-regions of the triangle with the reference nodes that are closest to the target.

3.1.2. Fingerprints Dimensionality Reduction

If it is possible to reduce the computational complexity of fingerprinting algorithms by reducing the number of search operations to find the best match, then it is legitimate to consider reducing the number of APs involved in the calculations too. Since, each dimension of a location fingerprint corresponds to RSS measurements from a certain AP in the region, then dimensionality reduction techniques can be used to reduce fingerprints dimensions from M to C , where $C < M$.

3.1.2.1. Principal Component Analysis

Principal Component Analysis (PCA) is a famous dimensionality reduction method that copes with the high dimensionality problem by linearly combining features into a lower dimensional space [25]. PCA uses the knowledge of the training data covariance matrix to decorrolate the features and to project the data in the direction of the largest variance. Therefore, PCA can be used to choose the most informative APs in any region given its radio map. However, PCA has a major drawback when used for classification, which is due to the fact that PCA preserves the features with the maximum variance, but

not the most discriminative features. Therefore, PCA dimensionality reduction might be at a cost of degradation in the positioning accuracy which is against our objective of maintaining high positioning accuracy.

3.1.2.2. *Fast Orthogonal Search*

Fast Orthogonal Search (FOS) is used to reduce the dimensionality of RSS fingerprints, as described in [27]. This approach is shown to outperform conventional dimensionality reduction techniques, such as PCA in terms of speed and accuracy. Namely, the FOS algorithm employs a Gram-Schmidt orthogonalization procedure to obtain the set of APs that, if used as basis functions, would minimize the total mean square error over all the APs in the global radio map. Therefore, an AP significance is measured in FOS by its contribution to other visible APs. If an AP is found to be contributing to most of the APs in the radio map, it is then marked as significant. Namely, FOS finds the subset of APs that form together the best replica of the original radio map.

3.1.2.3. *APs Significance Measures*

Other approaches are proposed in the literature to eliminating redundant APs. This is achieved by measuring the AP significance or contribution to the positioning performance. In [28], the strength of the signal received from a certain AP is used to measure the significance of that AP at each location fingerprint in the region. Each AP is therefore assigned a weight proportional to the strength of the signal and those with minimum weights are dropped from each fingerprint. A preliminary study of the various AP significance measures such as, average RSS, entropy, variance, maximum RSS is presented in [29]. The different measures are examined during both phases of fingerprinting. The obtained results show that if RSS-based measures are used to eliminate redundant APs, the accuracy of positioning will improve. Furthermore, the study in [30] suggests storage reduction for systems employing on-device stored radio maps for mobile-based positioning. The authors propose two modes for choosing the most informative APs, which are batch and continuous modes. In the batch mode, the APs that are mostly visible at all the location fingerprints are marked as significant. While in the continuous mode, the algorithm updates the most informative APs every

time a new fingerprint is recorded, in the sense that APs generating low RSS values are dropped.

3.2. Dynamic Propagation Models

Propagation models, such as pathloss, are static and deterministic which means they do not take into account the randomness in the RSS measurements when calculating the distance. Therefore, such models do not suit the indoor environments and are expected to result in poor positioning performance. In [31], the authors suggest an alternative probabilistic model to the commonly encountered path loss model; that is the Gaussian process regression model (GPR). GPR models are used to estimate the distance of the target from any AP. GPR is a powerful tool that can handle noisy RSS measurements. In addition, the probabilistic nature of GPR models allow them to capture the dynamic changes in indoor environments, as well as the multipath pattern. Moreover, the system suggested in [31] is capable of performing the estimation of the signal propagation model at runtime without the need for offline training. In addition, the authors in [32] suggest updating the conventional path-loss model into a path-loss log normal shadowing model that takes into account the shadowing effect in the environment. This model is justified by the fact that the shadowing effect in indoor environments is considered as a source of the temporal variations characterizing RSS. Therefore, it can reflect the noise in indoor environments. Recursive least squares (RLS) algorithm is suggested to estimate the parameters of this model [32]. The RLS algorithm is used to filter out noise and provide a better estimate of the log normal path-loss model.

3.3. RSS Correction Methods

One way of handling the temporal variations in the RSS measurements challenge discussed in Section 1.2 is to use RSS correction methods. The objective of such methods is to obtain location only dependent information from the RSS measurements by eliminating any environmental noise.

3.3.1. Feature Extraction

A commonly used RSS correction method is the feature extraction based method that considers statistical features of the RSS measurements such as the mean, mode, standard deviation, etc [33]. The correction is achieved by taking samples from the RSS measurements at a certain location and then replacing the samples by one RSS value. This value can be taken as one of the statistical features of the samples, such as average or mode.

3.3.2. Time Series Analysis

In [34], time series analysis is used to analyze the correlation between consecutive RSS samples received from one AP. This paper investigates the temporal variation of the RSS values at a fixed location. By experimental tests, it is shown that the fluctuations in the RSS values, at a fixed point, can be as large as 10 dBm. Therefore, considering one sample in the position estimation, might result in low accuracy. In addition, simple averaging improves accuracy but the experiments show that consecutive samples have high correlation. Hence, assuming independence of samples in averaging is misleading. To account for the samples correlation, the authors of this paper treat the RSS samples as a discrete time series that is expressed by a first order autoregressive model. The mean and variance of n correlated samples are obtained, and it is shown that the variance of n correlated samples is further reduced resulting in less fluctuations and hence better expected positioning accuracy.

3.3.3. Kalman Filtering

The accuracy performance of RSS-based fingerprinting technique is improved in [35] and [36] using the Kalman filtering method. Kalman filter is used to eliminate the temporal variations of RSS signals. Typically, the RSS measurements, at the same location, exhibit fluctuations that can be as large as 10 dBm. Using Kalman filter, the authors claim that the RSS fluctuations are reduced to 1 ~ 3 dBm. The filtered RSS values are then averaged to find the true RSS value at a certain location.

3.4. Online Calibrated Radio Maps

Although fingerprinting outperforms other indoor positioning algorithms, it is not commercially adopted due to the many impracticalities coupled with it. This is because fingerprinting require extensive site surveys to calibrate the radio map, which are costly and time consuming. In addition, any constructed radio map runs outdated if there is a variation in the environment, compelling new up to date site surveys. To avoid this tedious process, attempts were made to find a method to automatically construct radio maps without the need for offline data training. In [37], the authors suggest a method to offset the offline constructed radio maps to adapt for the environmental changes by using a system of reference points. The presented model applies regression analysis to predict new radio map values using retrieved relationship between the RSS values measured by a mobile device and those measured by the reference points. The same principle is applied in [38], except that an artificial neural networks based model is used to offset the environmental factors. In [39], the authors exploit manifold alignment and a Hidden Markov Model to update outdated radio maps of a wireless-based localization system without the need for any extra hardware. In all the previous models, at least an initial offline data training is required to calibrate the radio map for the region of interest. However, the authors in [14] propose a system that estimates and calibrates radio maps automatically for indoor positioning with zero-configuration needed. The system proposed in [14], consists of a network of APs that are equipped with transceivers that allow them to measure RSS values received from neighboring APs. With a modification to the APs software, the power recordings, along with the APs MAC addresses, are carried on managerial packets to be sensed by Wi-Fi enabled devices. Thereafter, the mobile devices communicate the received data to a central unit where the radio map is constructed using regression algorithms.

3.5. WLAN-based positioning systems

Many implementations of WLAN-based positioning systems are proposed in the literature. In [16], an RF-based positioning system for indoors localization is presented. The RSS values collected, at known locations in the region of deployment, along with

the RSS propagation models are used to find the mobile user location. Namely, the Floor Attenuation Factor (FAF) propagation model is used. Triangulation technique and KNN are employed to find the user location using both empirically and theoretically determined information. The system is implemented on two phases; an online and an offline phase. During the offline phase, the RSS data are collected and the propagation model parameters are learnt. During the online phase, the RSS values measured by the user are communicated to a central unit where the location calculations take place. Another famous WLAN-based positioning system is Horus system which is proposed in [15]. Horus system is also an RF based positioning system that employs probabilistic positioning techniques. This system is recognized for its ability to achieve high accuracy results using reduced complexity algorithms. The reason for the achieved high accuracy is that the system compensates for sources of wireless channels instability. Moreover, it utilizes clustering techniques to decrease the algorithm computational time. Consequently, this system is suitable for mobile devices with limited power supply. This system uses offline fingerprinting to construct a radio map that is composed of an RSS distribution received from visible APs at each grid point. It has different components starting with the clustering module that is used to reduce the complexity of calculations by dividing the fingerprints into groups, according to the common visible APs by them. It uses a discrete space estimator to return the fingerprint that has the maximum likelihood of resulting in the measured RSS vector. A correlation handler is used to correct the discrete estimator output by taking the average of n RSS samples. Then, it forwards the output of the correlation handler to the continuous space estimator to refine the location estimate.

Chapter 4: Enhanced Fingerprinting using Hybrid Search Techniques

In this chapter, we propose a hybrid search technique to enhance fingerprinting localization algorithms. This proposed search technique employs selective matching between the received signal strength (RSS), measured by the target, and the pre-stored fingerprints to reduce fingerprinting computational requirements. The reduction is achieved by minimizing the number of search points needed to find the best match between the target RSS and the pre-stored fingerprints. Although, clustering helps to minimize the search operations needed to find the target location, the amount of its complexity reduction is restricted by the trade-off between the required number of search points and the positioning error. Namely, as the number of used search points reduces, the error performance becomes worse. This has encouraged us to investigate fast search strategies, such as Three Step Search (TSS), Orthogonal Search (OS) and Diamond Search (DS), which are used to avoid exhaustive search in video compression applications [40, 41]. Our investigation concluded that a hybrid solution of clustering and search strategies violates the accuracy-complexity tradeoff allowing more computations reduction.

Therefore, we propose a system that integrates fast search strategies with clustering techniques to further reduce the complexity of fingerprinting algorithms, while maintaining high positioning accuracy and precision. To do so, we have modified the fast search strategies such that they can be easily integrated with clustering techniques. In our system model, a clustering technique is applied to the constructed radio map during the offline phase of fingerprinting, and a fast search strategy is used during the online phase to find the delegate cluster. It is important to note that fast search strategies are used to speed up the process of finding the cluster where the mobile user is located. Once a cluster is found, a fine resolution search within that cluster is used to achieve higher localization accuracy. Such breakdown of the localization process is expected to be of low computational complexity which makes the proposed system more suitable for battery-operated devices with limited processing capability. In addition, achieving accurate positioning makes the system suitable for applications requiring high position-

ing accuracy, such as tracking mobile emergency callers and automated object tracking.

4.1. Overview of The Proposed Positioning Approach

In this thesis, we adopt the location-based clustering technique proposed in [24]. A major advantage of this technique, over other clustering techniques, is its simplicity. This is due to the fact that it is a one step classification that only uses the physical proximity of the grid points. In addition, this technique is not only scalable to changes in the system coverage area, but is also robust to environmental changes [24]. For search strategies, we employ the so-called fast search strategies that are used in applications demanding a fraction of a second processing, such as motion estimation in video compression algorithms [40, 41]. In the following subsections, the breakdown of the proposed system is explained.

4.1.1. Fingerprinting

As explained in Subsection 2.4.3, fingerprinting is performed on two phases; an offline phase and online phase. This technique is illustrated in Figure 7.

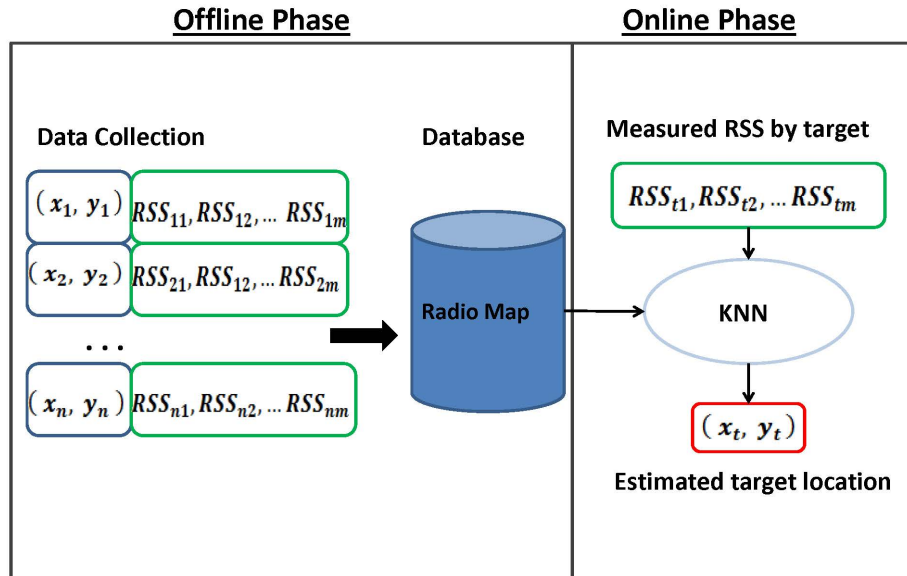


Figure 7: Illustration of fingerprinting technique.

During the offline phase, an $n \times m$ radio map for the area of interest is constructed. The RSS values RSS_{ij} received from m visible APs in the region are collected at n grid points p_i , where $i = 1, \dots, n$, and $j = 1, \dots, m$. The resulting radio map is saved to a database, to be used during the online phase. In the online phase, a matching algorithm is used to compare the RSS measured by the target, against the pre-stored radio map, to find the grid point with the minimum matching error. The matching algorithm adopted in this thesis is the weighted KNN, instead of the standard KNN that is explained in Subsection 2.5.1. The weighted KNN is regarded as a better positioning algorithm, since it gives more importance (higher weight) to the neighbors closest to the target. In other words, it employs a weighted average of the k closest grid points to the target to estimate its location. This algorithm, as clarified by Equation 6, employs the root mean square error to associate the locations p_i of the grid points with weights w_i , where RSS_t is the RSS measured by the target and RSS_i is the RSS at the i th grid point.

$$p_t = w_1 p_1 + w_2 p_2 + \dots + w_k p_k, \quad (6a)$$

$$w_i = \frac{e^{-(RSS_t - RSS_i)^2}}{\sum_{z=1}^k e^{-(RSS_t - RSS_z)^2}}. \quad (6b)$$

Such that, the grid point with the minimum RSS root mean square error is given the maximum weight and vice versa. Eventually, the position of the target p_t is estimated by summing the weighted locations of the k closest grid points, where $2 \leq k \leq n$. [7].

4.1.2. Location-based Clustering

In location-based clustering, the radio map is divided into distinct or overlapping clusters of neighboring grid points, with defined midpoints for each cluster. The clustered radio map resulting from the offline phase is saved into a database. During the online phase of fingerprinting, the predefined midpoints of the clusters, marked by 1 in Figure 8, are examined first. Afterwards, the search for the best match is limited to the cluster with the closest midpoint to the target, marked by 2 in Figure 8.

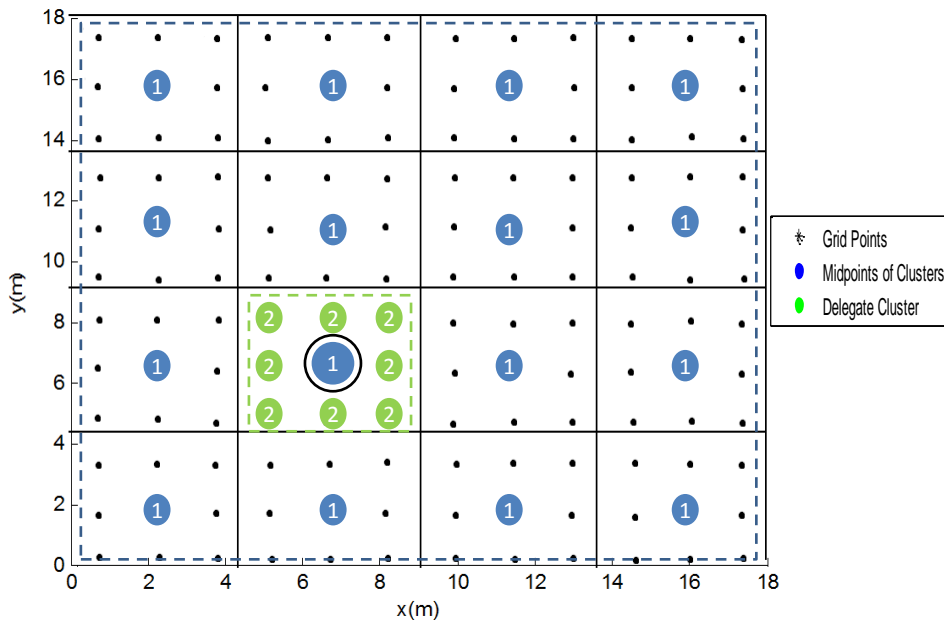


Figure 8: Illustration of location-based clustering.

4.1.3. Fast Search Strategies

Fast search strategies can also be considered for the use in indoor positioning to reduce the number of operations needed to find the best match. In this thesis, strategies, like Three Step Search (TSS), Orthogonal Search (OS), and Diamond Search (DS), are modified to be easily integrated with fingerprinting algorithms and to suit location finding applications.

The first search strategy to be used, is the TSS. This search strategy consists of three levels of hierarchal search with varying step sizes. In a sense that the search resolution increases for every next level, as shown in Figure 9. The first search window, marked by 1 on the figure, is reduced around the search point that results in the best matching. The same is repeated for the remaining steps, until the distance between search points reaches 1 point separation.

The second search strategy is the Orthogonal Search (OS), or sometimes called Directional Search. As the name suggests, the search in this technique takes place on two orthogonal directions, horizontal and vertical.

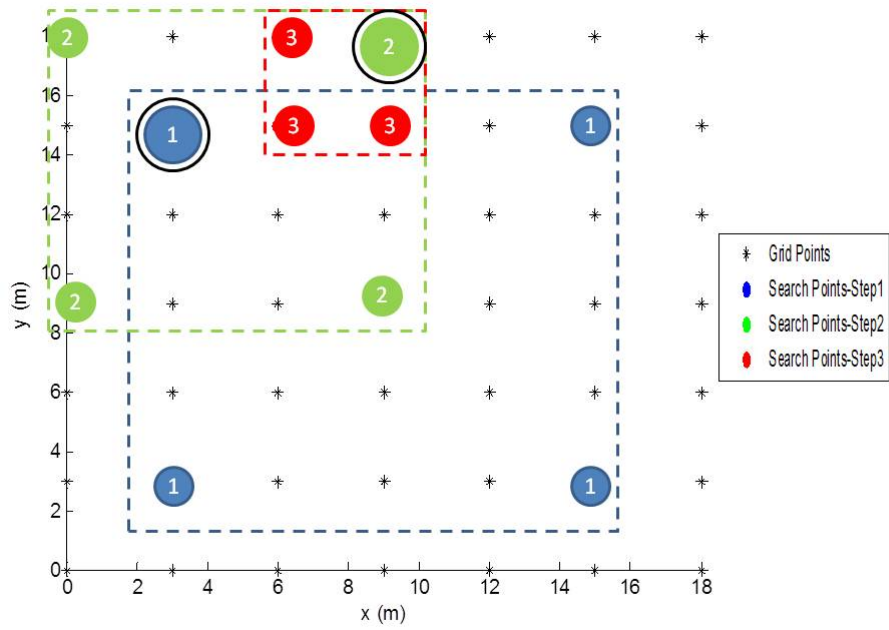


Figure 9: Illustration of three step search.

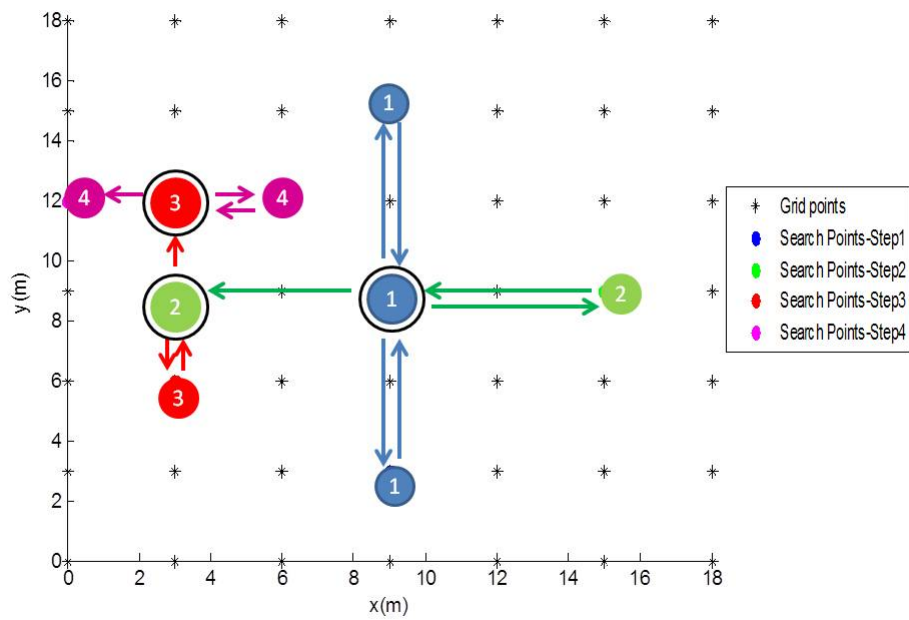


Figure 10: Illustration of orthogonal search.

This strategy is illustrated in Figure 10, starting from a centre point (labeled by 1 in Figure 10), the two vertical nearby points at a distance of 2 points separation are examined. The point that results in minimum error, when compared to the target, is chosen as the centre for the horizontal search. The same is repeated for steps 3 and 4 but with a distance separation of 1.

The third strategy examined is the Diamond Search (DS). This technique is similar to the TSS, but diamond-shaped search windows are used instead of square-shaped windows. The DS strategy is elaborated in Figure 11. As shown in the figure, the centre point and its four surroundings, forming a diamond shape, are tested in the first step. Later on, the point that results in the minimum distance error is compared against its neighbors to find the best match. The same is repeated for the following steps until the smallest possible search window is reached.

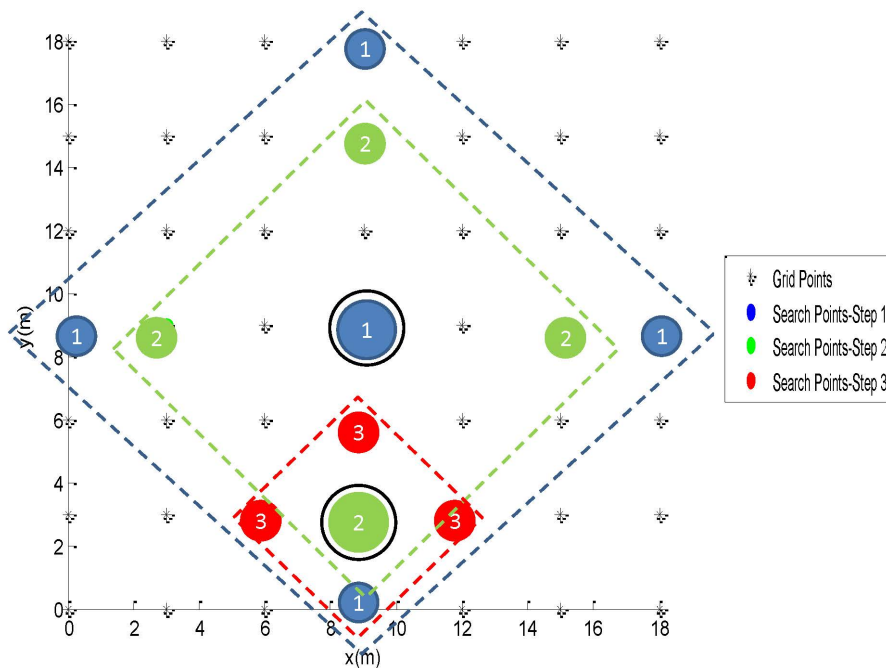


Figure 11: Illustration of diamond search.

4.1.4. The Proposed Hybrid Search

In the proposed hybrid search, we investigate the performance of clustering when integrated with one of the search strategies proposed in this chapter. If a suffi-

cient number of search points is used, clustering techniques can result in a comparable accuracy to that of full search. To further reduce the number of search points, while maintaining the good performance of clustering, fast search strategies can be used. Fast search strategies converge faster to the minimum error matching, requiring a minimum number of search points when compared to clustering. However, they are less accurate and less precise. Since fast search strategies result sometimes in a local best match, instead of an absolute best match, we accordingly integrate both clustering and fast search strategies into the system to capitalize on the advantages of both techniques. The proposed hybrid solution is shown in Figure 12. In the proposed hybrid system, location-based clustering is performed during the offline stage immediately after data collection. Whereas, any of the three suggested search strategies is implemented, during the online phase, to search through the midpoints of the clusters. In other words, our proposed system uses search strategies to reduce the number of search points needed to implement clustering.

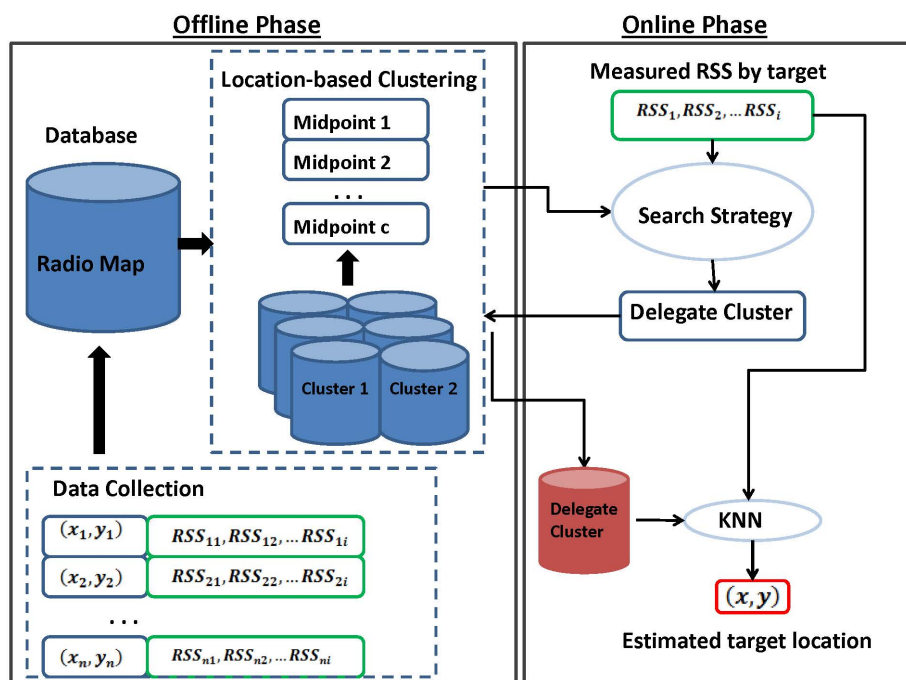


Figure 12: Proposed system model of the hybrid solution.

4.2. Experimental Setup

A test bed was set up to collect RSS values at the first floor of the engineering building I rotunda at the American University of Sharjah, Sharjah, UAE. The RSS values from 15 visible access points, over an area of 324 m², were collected. The RSS values were measured using a personal laptop with the Intel Wireless N 2230 WLAN card. The measured values were extracted from the specified card using the Vistumbler Software. The testing area was divided into 49 grid points with 3 meters separation between each. Then, approximately 300 RSS samples were recorded at each location point within a 2 minutes period. The samples were measured with 4 different orientations at each grid point. The average of the RSS values from each AP at each location point is saved into a radio map. The radio map construction is illustrated in Figure 13. The black dots in the figure indicate the fingerprints, while the stars indicate the location of the access points. Finally, the obtained radio map is tabulated and exported to MATLAB. The RSS values for 15 targets distributed over the area were also recorded for testing.

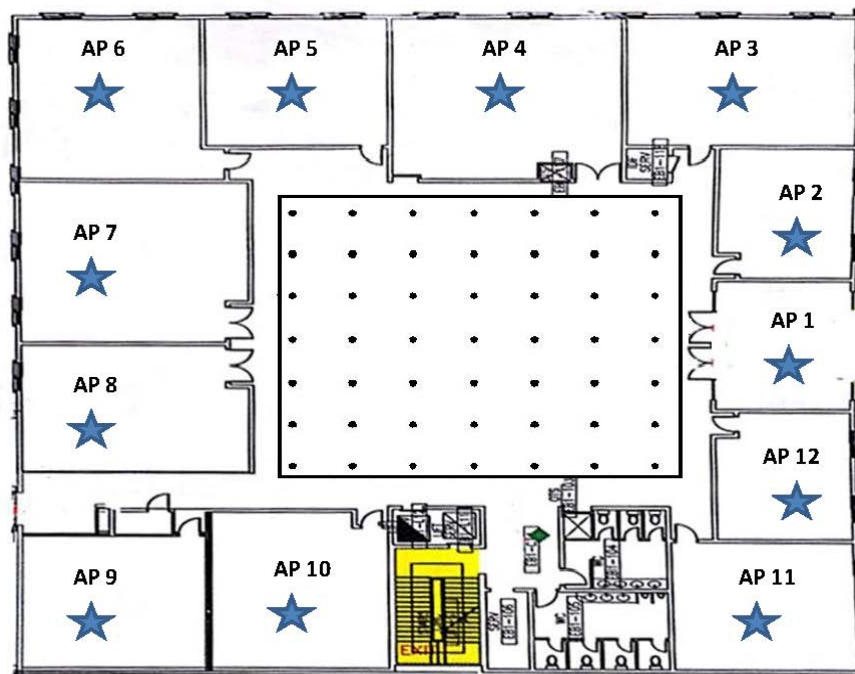


Figure 13: Top view of the experimental area.

4.3. Experimental Results

To evaluate the performance of the proposed hybrid search, the geographical map with the 49 grid points is divided into 9 overlapping location-based clusters. Afterwards, the hybrid search strategies are used during the online phase of fingerprinting to find the location of the 15 targets. For the sake of comparison, the localization error performance is investigated for different search schemes. Full search is intuitively the best when it comes to error performance. Therefore, the error cumulative density function (CDF) obtained by full search is used as a reference to assess other schemes. More specifically, we have studied the performance of full search against the performance of location-based clustering, the three fast search strategies when deployed independently, and the performance of the proposed hybrid search. Figure 14 depicts the error CDF curves for all search schemes.

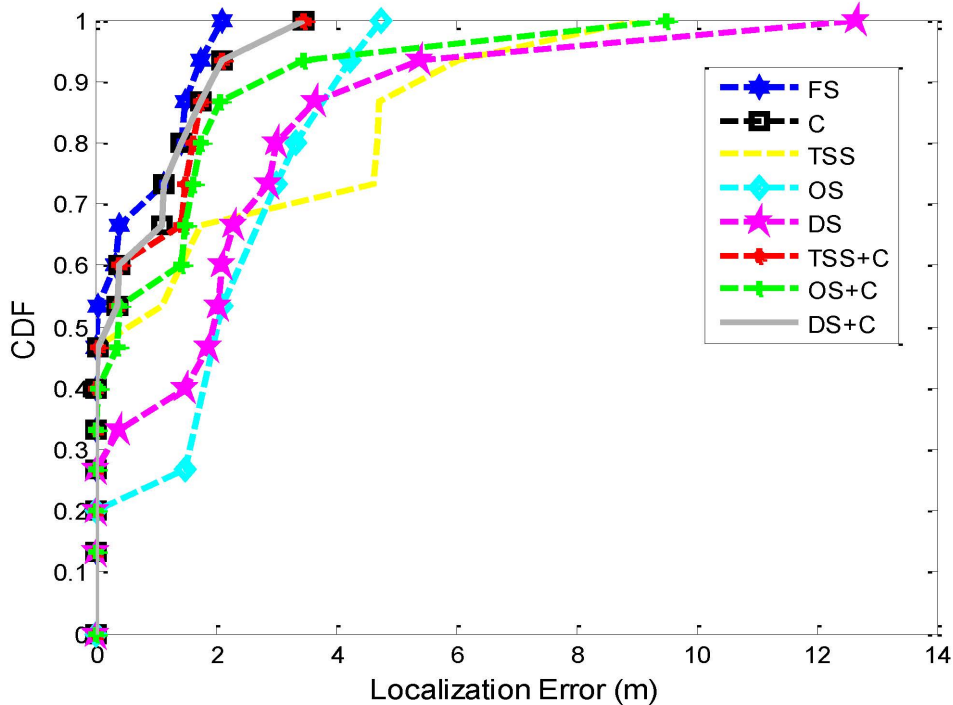


Figure 14: CDF error curve of different clustering and search strategies.

In the figure, (TSS+C), (OS+C) and (DS+C) stand for the hybrid schemes, where the three step search is integrated with clustering, the orthogonal search is integrated with

clustering, and the diamond search is integrated with clustering, respectively. Figure 14 shows that the error CDF curves of the different combinations of the hybrid search overlap with that of clustering. It also shows that those curves are very close to the error performance curve of the full search, achieving the second best performance. The probability of error for the fast search strategies, however, add up to one at a high distance error value, indicating poor performance.

Table 1, summarizes the accuracy, precision and percentage of used search points for each scenario. The mean distance error is used as a metric to judge the accuracy of positioning, while the distance error standard deviation is attributed to the positioning precision. The results in Table 1 show that the full search results in the highest accuracy and precision at the cost of high computational complexity, as it requires searching the entire map. When using location-based clustering techniques, it can be seen from the results that the accuracy and precision are slightly reduced, although the required number of search points is less than half of the points required for full search. While the three suggested search strategies have the lowest required number of search points, their positioning accuracy and precision are the worst. Combining search strategies with clustering techniques result in an improved performance. As it can be seen from the last three hybrid solutions results, the accuracy and precision of which are comparable to the full search performance. Although, they require a reduced number of search points.

Table 1: Evaluation of the different clustering and search strategies.

Scenario	Accuracy (m)	Precision (m)	Search Points (%)
Full search	0.5622	0.7455	100
Clustering (C)	0.7982	1.0390	36.37
Three-Step Search (TSS)	2.2302	2.8727	24.49
Orthogonal Search (OS)	2.3712	1.5274	24.49
Diamond Search (DS)	2.5250	3.2176	30.61
Hybrid Solution 1 (TSS+C)	0.8506	1.0710	34.69
Hybrid Solution 2 (OS+C)	1.4829	2.4476	30.61
Hybrid Solution 3 (DS+C)	0.7921	1.0390	33.47

Chapter 5: RSS Fingerprints Dimensionality Reduction

In this chapter, a dimensionality reduction method is proposed to eliminate redundant or non-informative access points (APs) from the positioning system. Reduction methods can be used to restrict the location fingerprints only to RSS values measured from informative APs. In doing so, the performance of fingerprinting is improved from two aspects. First, the number of RSS values involved in the position calculation is reduced, hence reducing the computational requirements of fingerprinting. Second, including only informative APs can even improve the positioning accuracy. Although, as discussed earlier, the multidimensional RSS vectors provide unique identification of position points in the region of interest, it is known in practice that a larger number of features can lead to worse positioning results [25]. This happens because involving too many APs in the location computation causes confusion to the system especially in indoor environments. This is due to the fact that the signals transmitted by those APs undergo spatially variant environmental noise that can be highly more severe in some places compared to others. Therefore, it is not at all unusual to have very close APs that result in very different RSS values at the same location point. In addition, one should keep in mind that mobile devices are involved in location computations in WLAN-based positioning which makes high computational costs unaffordable. Whether the positioning system is totally mobile based or assisted by a server, the high dimensional RSS vectors are costly. High dimensionality infers computational and storage costs in mobile based systems, while in mobile assisted systems, high dimensionality induces transmission costs and time delays. Reducing the dimensionality of RSS fingerprints helps in avoiding all these costs and it could improve the accuracy of positioning if some of these APs are misleading. Therefore, the objective of this study is to employ a feature reduction technique to reduce the dimensionality of the RSS vector while maintaining good positioning performance.

The proposed dimensionality reduction method, in this thesis, is based on the fast orthogonal search method proposed in [27]. Specifically, a modification to the FOS implementation is proposed. The objective is to benefit from the lightweight of the original FOS algorithm and modify it to further improve the resulting positioning accu-

racy. According to the modified algorithm, the most informative APs to the positioning system are those providing unique signal strength values that enhance the regional discrimination. To elaborate on this, for an AP to be marked informative, it should provide regionally discriminative RSS values, which if added to an RSS vector, will result in a higher correspondence to the associated location point. To the best of our knowledge, there is no obvious measure for the truly most informative APs to the positioning system. However, we argue that the APs resulting in divergent RSS values, that are more unique than the rest of the APs in the radio map, are more informative. Therefore, we modified the FOS algorithm to choose the AP with the minimum contribution to the rest of the APs. However, before we implement the modified FOS, it is needed to ensure that all the RSS values chosen as discriminant are valid RSS values and not zero i.e., the AP is not hearable. Therefore, an algorithm is proposed to remove all the APs that are not hearable at all location points.

5.1. Overview of The Proposed Approach

During the Offline phase of the fingerprinting technique, a radio map is obtained by collecting real RSS samples at known locations. The obtained radio map is represented as an $M \times N$ matrix, see Table 2, where each of the M rows corresponds to a known location and the columns represent the signal strength measurements from N APs. Having the radio map in this format, dimensionality reduction techniques will be implemented to reduce $M \times N$ matrix to a $M \times C$ matrix, where $C < M$.

Table 2: Radio map for M gridpoints and N access points.

	AP_1	AP_2	\dots	AP_N
P_1	RSS_{11}	RSS_{12}	\dots	RSS_{1N}
P_2	RSS_{21}	RSS_{22}	\dots	RSS_{2N}
\vdots	\vdots	\vdots	\vdots	\vdots
P_M	RSS_{M1}	RSS_{M2}	\dots	RSS_{MN}

5.1.1. Fast Orthogonal Search

The fast orthogonal search method considers the following model:

$$y(n) = \sum_{m=0}^{c-1} a_m p_m(n) + e(n), \quad (7)$$

where $y(n)$ is every data column of the radio map, modeled using a small subset of the other $N - 1$ columns. Furthermore, p_m are the C basis functions and a_m are the coefficients that are calculated using optimization techniques in order to minimize the error $e(n)$, which is the error between the actual output and the estimated $y(n)$, as described in Equation 7. Essentially, Orthogonal Search (OS) techniques study each term $a_m p_m(n)$ in order to determine its contribution in modeling the desired output. FOS is an iterative algorithm that searches for C columns from among N columns ($C < N$), that if used as basis functions, would minimize the total mean squared error (MSE) over all columns. These C columns are equivalent to the most informative APs in the radio map. The functions $p_m(n)$ in 7 are replaced using the Gram-Schmidt procedure, with a set of orthogonal basis functions $w_m(n)$ that are represented in the following model [42]:

$$y(n) = \sum_{m=0}^{c-1} g_m w_m(n) + e(n). \quad (8)$$

The parameters g_m minimize the MSE are modeled as [43]:

$$g_m = \frac{\overline{y(n)w_m(n)}}{\overline{w_m(n)^2}}, \quad (9)$$

and the MSE is calculated as:

$$\overline{e^2} = \overline{\left(y(n) - \sum_{m=0}^{c-1} g_m w_m(n) \right)^2}. \quad (10)$$

With some mathematical manipulations, the above equation can be simplified to:

$$\overline{e^2} = \overline{y(n)^2} - \sum_{m=0}^{c-1} Q_m, \quad (11)$$

where,

$$Q_m = \frac{\overline{[y(n)w_m(n)]^2}}{\overline{w_m(n)^2}}. \quad (12)$$

It is clear that there is reduction in error by an amount Q_m due to the addition of the term $g_m w_m(n)$ in the model. Therefore, it is required to calculate Q_m for each candidate and choose the one for which Q_m is found to be greatest. Furthermore, the construction of orthogonal basis functions w_m is found to be computationally intensive and can be avoided in FOS, where only the correlations of $w_m(n)$ with $p_m(n)$ and $y(n)$ are required [43]. These can be obtained through the following set of equations:

$$g_m = \frac{C(m)}{D(m,m)}, \quad \forall m \in \{1, \dots, M\} \quad (13)$$

$$a_{mr} = \frac{D(m,r)}{D(r,r)}, \quad \forall r \in \{1, \dots, m\} \ \& \ m \in \{1, \dots, M\} \quad (14)$$

and where

$$D(0,0) = 1, \quad (15)$$

$$D(m,0) = \overline{p_m(n)}, \quad \forall m \in \{1, \dots, M\} \quad (16)$$

$$C(0) = \overline{y(n)}. \quad (17)$$

Thus, in general,

$$D(m,r) = \overline{p_m(n)p_r(n)} - \sum_{i=0}^{r-1} a_{ri}D(m,i), \quad (18)$$

$$C(m) = \overline{y(n)p_m(n)} - \sum_{r=0}^{m-1} a_{mr}C(r). \quad (19)$$

Finally, Q_m , the amount each orthogonal function deducts from the MSE, is calculated using the following equations:

$$Q_m = g_m^2 \overline{w_m^2(n)}, \quad \forall m \in \{1, \dots, M\} \quad (20)$$

$$Q_m = g_m^2 D(m,m) = \frac{C(m)^2}{D(m,m)}.$$

5.1.2. FOS Implementation in Positioning System

Two different implementations of FOS are tested in this thesis; traditional FOS and modified FOS.

5.1.2.1. Traditional FOS

Each column of the radio map is treated as a potential basis function $p_m(n)$. At each iteration, the significance of each column is evaluated by calculating the reduction in error caused by adding the basis function $p_m(n)$ to the current model of $y(n)$. The basis function with the largest reduction in error (maximum Q_m), when added to the group of already selected basis functions from previous iterations, is stored. The iterations, for the current model of $y(n)$, are carried out until C data columns, causing the largest reduction in error, are found. Consequently, the same procedure is carried out for every $y(n)$, after which the C columns for each model are tallied in order to identify the APs (columns) that result in the highest Q_m values [9].

5.1.2.2. Modified FOS

To ensure that all the APs are hearable at all location points, a simple algorithm is used. The algorithm basically eliminates all APs that don't cover the entire location area. This algorithm is explained below:

Eliminate nonhearable APs

```
1: for  $i = 1$  to  $N$  do
2:  $y(i)$ ; (Radio Map Columns)
3:   for  $z = 1$  to  $M$  do
4:     if any of  $y(i)$  values = 0
5:       Remove the  $i$ th column from the radio map
6:   end
7: end
```

Figure 15: The algorithm to eliminate nonhearable APs.

Afterwards, the radio map becomes ready to be compressed using the modified FOS (MFOS). In the implementation of MFOS, the search is after the unique APs. Therefore, The FOS algorithm is altered such that Q_m value for each $p_m(n)$ is calculated

for each model $y(n)$. For every model, the C top p_m s with the lowest Q_m s are stored. Finally, C columns for each model are tallied in order to identify the APs (columns) that contributed the least to the models indicating the most unique APs. A flowchart describing the above process is shown in Figure 16.

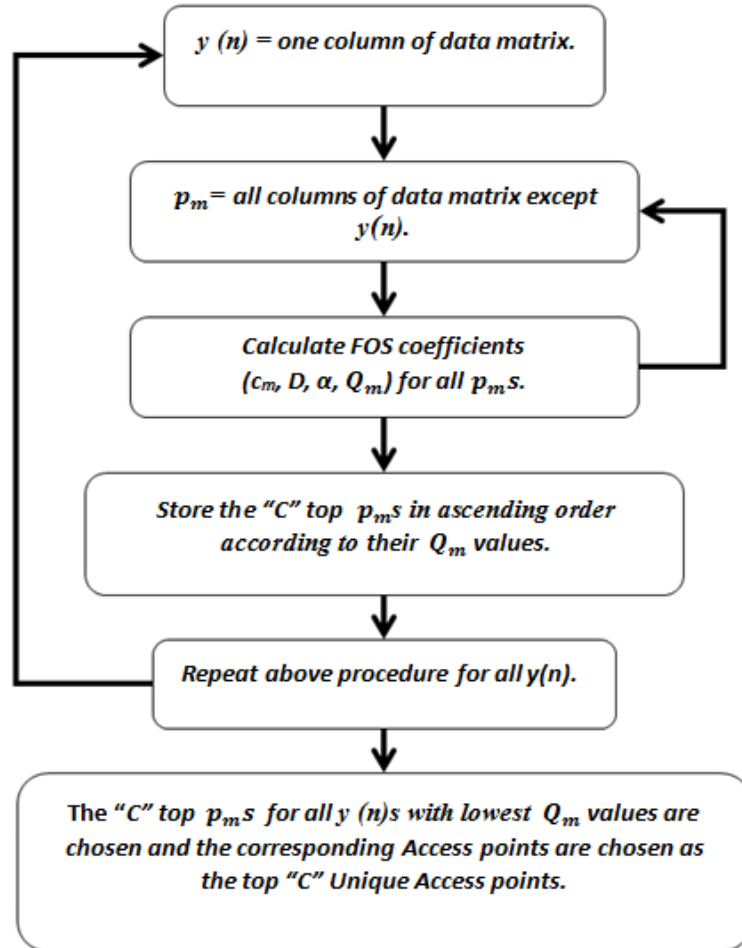


Figure 16: Flowchart representing the modified FOS procedure.

The reduced radio map is saved so that it will be used during the online phase of fingerprinting. Positioning is carried out using one of the matching algorithms, such as KNN, to find a new RSS measurements' location. Particularly, weighted KNN is adopted in this part as the matching algorithm which is explained by Equation 6 in Subsection 4.1.1.

5.2. Experimental Results

To evaluate the performance of the proposed hybrid search, the experimental setup and the radio map obtained in Section 4.2 were used in this part as well. Using the nonhearable APs elimination algorithm, 12 out of 15 APs in the region were found to be hearable to all location fingerprints. Therefore, the obtained radio map is reduced to 12×49 . The dimensionality reduction methods were used to reduce the 12×49 radio map to a 4×49 . Where, $M = 49$ is the number of location gridpoints, $N = 12$ is the number of visible APs, and $C = 4$ is the number of chosen APs. In other words, the 4 most informative APs, out of the visible 12 APs, in the region were determined using dimensionality reduction techniques. The feature reduction on the radio map was performed separately for PCA, traditional FOS and the altered FOS. For all methods, the execution halts after adding 4 data columns. The three algorithms are compared to a reference solution of 4 APs chosen heuristically (try all possibilities to choose the APs that result in minimum error). The reduced radio maps obtained by all the above mentioned techniques were used to find the location of 15 targets and the positioning accuracy was used to indicate the quality of the dimensionality reduction technique. The KNN algorithm is used and the $k = 2$ closest neighbors were found for each case. The performance evaluation of the three dimensionality reduction techniques, such as the PCA, the FOS and the modified FOS, are all compared to the positioning performance of full radio map and the heuristic search. The performance of the FOS and the modified FOS are shown in Figure 17. As shown in this figure, the modified FOS algorithm provides a lower error when compared to the original FOS. Namely, it can be seen that for high probabilities modified FOS result in a better error performance, achieving higher accuracy.

In Figure 18, the positioning error performance is shown for all scenarios. As it is shown in the figure, the full radio map gives the best performance along with the heuristic search, while the modified FOS results in the second best performance. The original FOS method performs worse than the former techniques, however it outperforms PCA which matches the results in [27]. In Table 3, the overall performance of the various dimensionality reduction techniques is summarized.

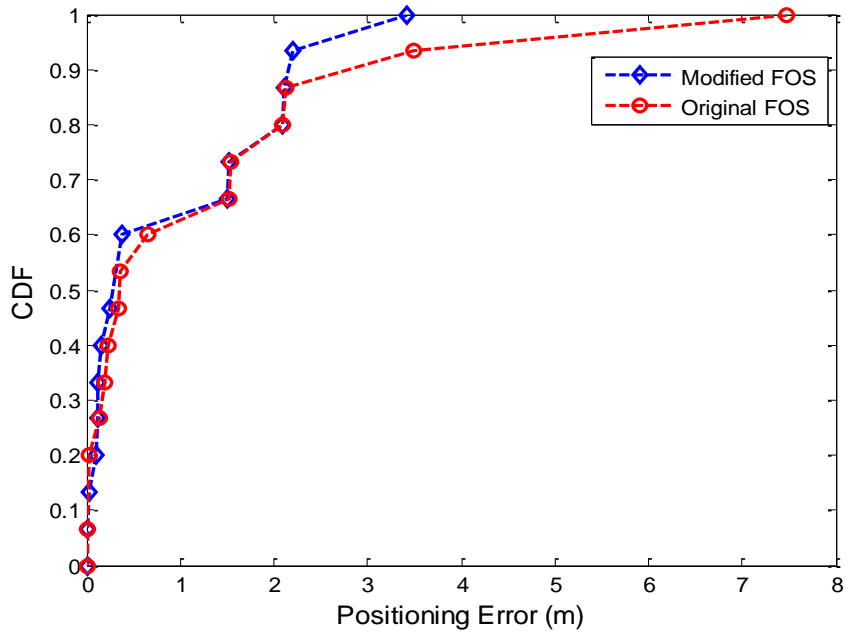


Figure 17: Positioning error CDF of FOS and modified FOS.

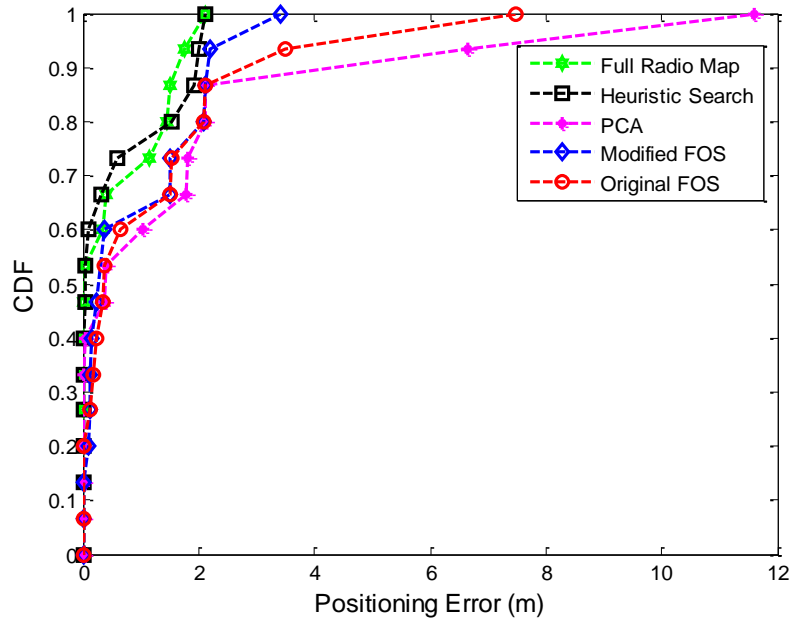


Figure 18: Positioning error CDF of MFOS compared to all techniques.

In terms of computation, the complexity of the FOS, due to the cross-correlations between all pairs of data and applying mean square error reduction N times, is $C_{FOS} = O(MN^2 + N^2C)$ which is small compared to $C_{PCA} = O(MN^2 + N^3 + MCN)$. Further-

more, the MFOS has similar complexity to that of the original. Since, the computational complexity for the FOS is much better than the PCA, it has a faster processing time. The processing time of the algorithms were obtained using MATLAB.

Table 3: Processing time and complexity for different reduction methods.

Metric	PCA	FOS & MFOS
Average error (m)	1.8639	1.5135 or less
Computational complexity	11,136	7,632
Processing times (secs)	2.857	0.094

Afterwards, we investigated the performance of the MFOS compared to the FOS at different reduction levels. The reduction methods were used to obtain different reduced sizes of the same radio map. The chosen APs from each reduction method were used to find the location of 15 targets. The accuracy performance for each method is presented in Table 4.

Table 4: Performance of different reduction methods for different levels of reduction.

Scenario	Resulting accuracy (m)		
	Heuristic search	FOS	MFOS
Reduction by 1 AP	0.5471	0.6012	0.5471
Reduction by 2 APs	0.4955	0.5743	0.6505
Reduction by 3 APs	0.5248	0.6221	0.6820
Reduction by 4 APs	0.5002	0.7382	0.6004
Reduction by 5 APs	0.4886	0.6076	0.6964
Reduction by 6 APs	0.4663	0.9765	0.8571
Reduction by 7 APs	0.4767	1.0773	0.9663
Reduction by 8 APs	0.5796	1.3465	0.9574
Reduction by 9 APs	0.8860	2.3294	1.5692

The heuristic search is used as a reference to define the minimum possible error at every

reduction level. It can be seen that the MFOS mainly results in a better performance than the FOS, especially when a significant reduction is achieved. When only one AP is removed from the system, it can be seen that the MFOS is as good as the heuristic search. In addition, it can be seen that when more than 50 % of the radio map is eliminated, the APs chosen by the MFOS result in a better positioning accuracy than those chosen by the FOS. From the obtained results, the modified FOS algorithm provides the best performance in terms of the average positioning error. In addition, it still maintains the good computational complexity and the processing time performance which is much better than the PCA method.

Additionally, the proposed reduction method in this chapter can be easily combined with the clustering and search strategies proposed in Chapter 4. Consequently, an ultimate reduction in the computational requirements of fingerprinting is achieved. Such that, the hybrid search solution of clustering and search strategies is used to minimize the number of search operations needed to find the best match, while the MFOS technique is used to minimize the number of features used in the matching process. Combining both methods is achieved as following: The reduced radio map attained by the MFOS is divided into clusters during the offline phase of fingerprinting. Then, a fast search strategy is used, during the online phase, to search through the pre-determined clusters. Once a delegate cluster is found, the search for the location best match is executed within that cluster only. The performance of all the clustering and the search strategies investigated in Chapter 4 when combined with the MFOS is presented in Table 5.

Table 5: Accuracy of clustering and search strategies with and without MFOS.

Scenario	Resulting accuracy (m)	
	without MFOS	with MFOS
Full search	0.5622	0.5920
Clustering (C)	0.7982	0.6696
Three-Step Search (TSS)	2.2302	1.7324
Orthogonal Search (OS)	2.3712	2.5304
Diamond Search (DS)	2.5250	2.8335
Hybrid Solution 1 (TSS+C)	0.8506	0.6802
Hybrid Solution 2 (OS+C)	1.4829	1.5105
Hybrid Solution 3 (DS+C)	0.7921	0.6696

It can be seen from the results that the error performance actually improves when using the reduced radio map obtained by MFOS in most of the scenarios. Although, the accuracy deteriorates for some scenarios, the amount of accuracy reduction at those scenarios is insignificant.

In the same manner, the time needed to find the positions of the 15 targets is investigated. An assumption is made that all the RSS measurements for all the targets are available at the processing unit. In other words, the time investigated in this part is just the time needed to search for the best match of the 15 targets in the radio map. Table 6 summarizes the processing time for the different clustering and search strategies studied in Chapter 4 with and without using MFOS. It can be seen from the results that the processing time for the location calculations is reduced significantly when the reduced radio map, obtained by MFOS, is used instead of using the entire radio map. In addition, the results underline the effect of the different clustering and search strategies on the processing time. The results show that using full search and clustering are at the cost of increased time requirements. While the fast search strategies (TSS, OS, DS) are the fastest in location determination. Finally, the hybrid search solution proposed by this thesis is shown to have a relatively small processing time when combined with the MFOS, as well as high positioning accuracy, as shown in Table 5.

Table 6: Computational requirements of clustering and search strategies with and without MFOS.

Scenario	Processing time (sec)	
	without MFOS	with MFOS
Full search	0.029704	0.019688
Clustering (C)	0.026615	0.012736
Three-Step Search (TSS)	0.012011	0.007180
Orthogonal Search (OS)	0.009335	0.007225
Diamond Search (DS)	0.010020	0.008438
Hybrid Solution 1 (TSS+C)	0.014285	0.010776
Hybrid Solution 2 (OS+C)	0.010921	0.009579
Hybrid Solution 3 (DS+C)	0.011619	0.009185

Chapter 6: Error Reduction in Distance Estimation of RSS Propagation Models using Kalman Filters

In this chapter, a different positioning technique is investigated; that is propagation models. Such models are less accurate than fingerprinting-based positioning techniques, however, they don't require expensive calibration. To enhance the positioning accuracy of propagation models, we propose using RSS correction methods. The reason behind using the RSS correction methods is to handle the temporal variations in the RSS measurements. Essentially, correction methods are used to eliminate any environmental noise contaminating the RSS measurements, hence obtain location only dependent measurements. We argue that when RSS correction methods are integrated with propagation models, the performance of such combined models is expected to get enhanced significantly.

A commonly used RSS correction method is the feature extraction based method that considers the statistical features of the RSS measurements, such as the mean, the mode and the standard deviation [33]. However, such correction methods give equal importance to all samples without differentiating between noisy and true RSS values. Therefore, such methods do not suit the noisy nature of the RSS samples. On the other hand, the most famous RF propagation model is the path loss model. However, the path-loss model is deterministic, which means it does not take into account the randomness of the RSS. Therefore, it does not suit indoor environments and is expected to result in poor positioning performance. Instead, we use Kalman filters in the correction process of the RSS measurements. Then, Kalman filters will be integrated with a probabilistic propagation model, such as the Gaussian Process Regression, to estimate the distance at which those filtered RSS values were measured.

Kalman filter is a powerful mathematical tool that is typically used in estimation problems. In the proposed system, Kalman filters are used to estimate the actual RSS value from a set of noisy RSS measurements at a certain location. We adopted Kalman filters in this thesis due to their known optimality in achieving minimum mean square error (MMSE). Hence, Kalman filters are expected to result in better performance in terms of estimating real RSS values, as compared to other RSS correction methods.

Afterwards, the filtered RSS values are forwarded to a GPR model to estimate the distance of the target from the AP. The GPR is another powerful tool that can handle noisy measurements. The probabilistic nature of the GPR models allow them to capture the dynamic changes in indoor environments, as well as the multipath pattern. Therefore, integrating both tools is expected to provide accurate distance estimation and hence enhance the positioning accuracy of propagation models which are cheaper alternatives to the more expensive fingerprinting techniques.

6.1. Overview of Proposed System

In this section we present the different components of the proposed system. First, we present the RSS measurements correction method using Kalman filters. Second, we discuss the usage of the GPR models in distance estimation. Then, we present the coordination between those system components to estimate a user position.

6.1.1. RSS Measurements Correction

Let $\{RSS_{1m}, RSS_{2m}, \dots, RSS_{km}\}$ be a set of k noisy RSS values measured from a certain AP (AP_m) at a certain location. Kalman filter will be used to estimate the true RSS value given the noisy measurements.

6.1.1.1. Kalman Filter Algorithm

Kalman filters are Bayesian filters that provide the state that maximizes the probability of resulting in the online observations $p(x|Y^k)$, where $Y^k = [y(0), y(1), \dots, y(k)]$ are the observations up to time k . As stated previously, Kalman filters provide the MMSE estimate which happens to be the mean of $p(x|Y^k)$, and its covariance is the measure of accuracy of the estimate. With every new observation, the filter propagates the mean and the covariance of a system state to provide a new estimate. Therefore, the Kalman filter is suitable for real time processing of data. In this paper, we implement a discrete time Kalman filter to estimate a time varying state parameter, such as the RSS, that follows a discrete time linear difference equation expressed by

$$x_k = A_{k-1}x_{k-1} + w_{k-1}, \quad (21)$$

while the observed system is given by

$$y_k = H_k x_k + v_k. \quad (22)$$

The random processes w_k and v_k are the system and the observation noises, respectively, which are assumed to be independent of each other with covariance matrices Q_k and R_k . A is a matrix that relates the previous to the current estimates, while the H matrix defines the relation between the measurement and the desired estimate. In our system, both matrices are equal to one, since the estimate here is just a constant that can be measured directly. Based on the system model and the measurements model, the filter is supposed to estimate the true RSS value. A posteriori estimate x_k^+ of the state can be found if all the measurements up to the k th time are available. On the other hand, if all the measurements up to time $k - 1$ are available, a priori x_k^- estimate of RSS value at time k is found. This makes x_k^- a predicted estimate of the state, whereas x_k^+ a smoothed estimate of the state. In addition, Kalman filters provide a measure of the uncertainty in the state estimate at each iteration which is calculated by the estimation error covariance P . The filter algorithms can be summarized using the following equations [44]:

$$x_k^- = A_{k-1} x_{k-1}^+, \quad (23)$$

$$P_k^- = A_{k-1} P_{k-1}^+ A_{k-1}^T + Q, \quad (24)$$

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1}, \quad (25)$$

$$P_k^+ = (1 - K_k H_k) P_k^-, \quad (26)$$

$$x_k^+ = x_k^- + K_k (y_k - H_k x_k^-). \quad (27)$$

6.1.2. Probabilistic Propagation Modeling

Propagation models are used to estimate the distance of a device from the AP, given its measured RSS value. Probabilistic propagation models consider the input and output to be randomly distributed. Hence, they work better in dynamic environments. In the case of indoor positioning, x , the input to the propagation model, is the RSS measurement, while y , the output, is the estimated distance. The input and the output

are related through an underlying function $f(\cdot)$ as presented by:

$$y = f(x) + \varepsilon, \quad (28)$$

where ε is an additive noise.

6.1.2.1. *Gaussian Process Regression*

The GPR is a mathematical tool that is used to estimate the output of a system at a new input, given a set of noisy input/output data. A Gaussian process (GP) is defined in [45], as a set of jointly Gaussian distributed random variables with a mean and covariance that are given by the following equations [46]:

$$\mu_x = E[f(x)], \quad (29)$$

$$k(x, x') = E[(f(x) - \mu_x)(f(x') - \mu_{x'})]. \quad (30)$$

In GPR models, the ε is assumed to be an independent additive zero mean Gaussian noise $\mathcal{N}(0, \sigma_n^2)$. Furthermore the covariance function is chosen to be the squared exponential in

$$k(x_p, x_q) = \sigma_f^2 \exp\left(\frac{-(x_p - x_q)^2}{2l^2}\right). \quad (31)$$

For this covariance function, it can be seen that when the input values are close, the corresponding random variables have high correlation, almost equal to σ_f^2 . Whereas, the covariance approaches zero when x_p is far from x_q . A separation factor, l , is used to define the amount of change in the input that is required to result in a significant change in the output [45]. After adding the noise to the system, the covariance function is given by:

$$\text{cov}(y_p, y_q) = k(x_p, x_q) + \sigma_n^2 \delta_{pq}, \quad (32)$$

where δ_{pq} is the Kronecker-delta function which is zero everywhere, except when $p = q$, resulting in only the diagonal elements of the covariance matrix to be interrupted by

noise [46]. The GPR model covariance matrix is evaluated for n data points as:

$$K = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix}, \quad (33)$$

while the covariance function for a new input x^* is given by

$$\begin{aligned} K^* &= \begin{bmatrix} k(x^*, x_1) & k(x^*, x_2) & \cdots & k(x^*, x_n) \end{bmatrix}, \\ K^{**} &= k(x^*, x^*). \end{aligned} \quad (34)$$

Finally, the estimated output y^* at the new input x^* is given by the estimated mean given by

$$\bar{y}^* = K^* K^{-1} y, \quad (35)$$

and the estimated uncertainty is given by the variance of the estimate which is given by

$$\text{var}(y^*) = K^{**} - K^* K^{-1} K^{*T}. \quad (36)$$

6.2. The Proposed System

In our proposed system, we integrate Kalman filters and Gaussian regression models to estimate a distance traveled by a received Wi-Fi signal. The purpose of such integration is to remove the temporal variations of the RSS measurements and to employ probabilistic propagation models that mimic the dynamic nature of signals propagation inside buildings. A one time confined calibration phase is needed to learn the parameters of the GPR models and the Kalman filter model covariances. The system is then ready to be used in online localization. Figure 19 shows the integration of the proposed system components into an online localization system. First, the mobile device, which is to be located by the system, collects the RSS samples (RSS_{ij}) from visible APs in the region, where $i = 1, \dots, m$ APs and $j = 1, \dots, k$ samples. Those, RSS samples are transmitted to a location server where the location computation is performed. The raw RSS samples (RSS_{ij}) measured by the mobile device are corrected using a Kalman filter

to find the best estimate of the true RSS value received from each AP. As a result, one estimate from the k samples received from each APs is obtained, resulting in a vector of m RSS values $\{RSS_1, RSS_2, \dots, RSS_m\}$. Secondly, the filtered RSS values are forwarded to the previously trained GPR propagation models to approximate the distance of the mobile device from each fixed APs. For each AP in the region, a GPR model is defined, hence, m GPR models are obtained. The input to each GPR model is the RSS value received from the AP corresponding to it, and the output is the distance traveled by the signal from the AP and received by the device ($RSS_i \Rightarrow GPR_i \Rightarrow d_i$). The exact location of the user can then be calculated using triangulation techniques. Our objective is to reduce the error in distance estimation to enhance the positioning accuracy. Kalman filters provide the optimal estimate of the RSS value by minimizing the mean square error (MSE) given the noisy RSS samples. While GPR models are used to estimate the distances, at which the RSS values are measured, through regression without the need for the tedious calibration phase. Our proposed approach is more practical for commercial use than fingerprinting techniques and more accurate than other traditional propagation models.

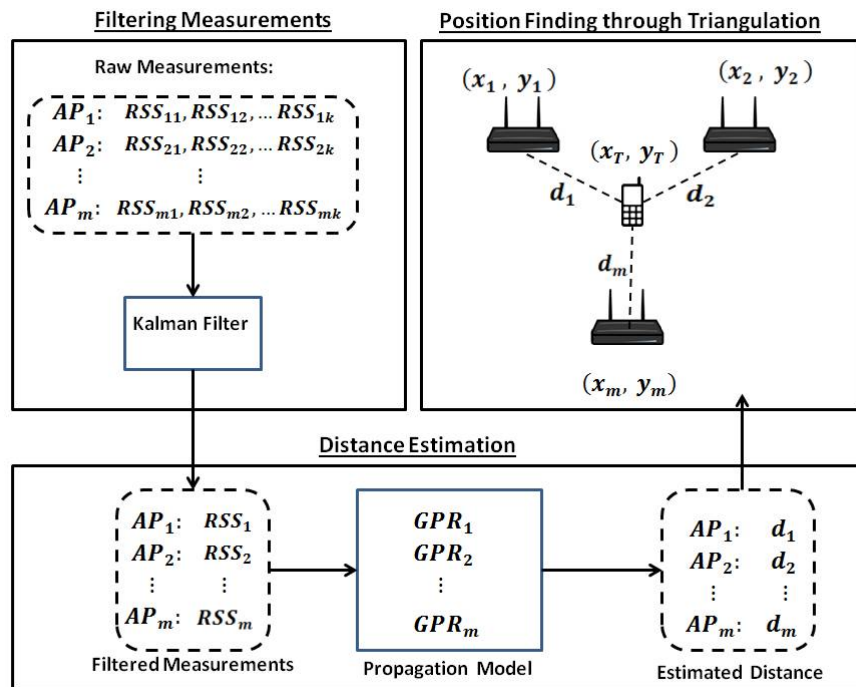


Figure 19: Proposed system model of propagation models when integrated with Kalman filters.

6.3. Experimental Results

To evaluate the performance of the proposed system, 10 RSS measurements and their corresponding distances were first collected to learn the model for each AP in the range, as illustrated in Figure 20. In our experiments, we considered σ_f and σ_n to be 10 dBm which is the amount of RSS fluctuations caused by the temporal variations. In addition, we considered the separation factor (l) to be 18 m, the length of the testing area. The Distance estimation is carried out for 36 targets sparsely distributed over the testing area. During the run time of the positioning system, 50 RSS samples measured from the same AP were collected at the 36 targets locations. Kalman filters were used to correct these measurements. The parameters of the covariance functions of the system and measurement noise, R_k and Q_k , were learnt through tuning and they were found to be 0.1 and 0.001 respectively.

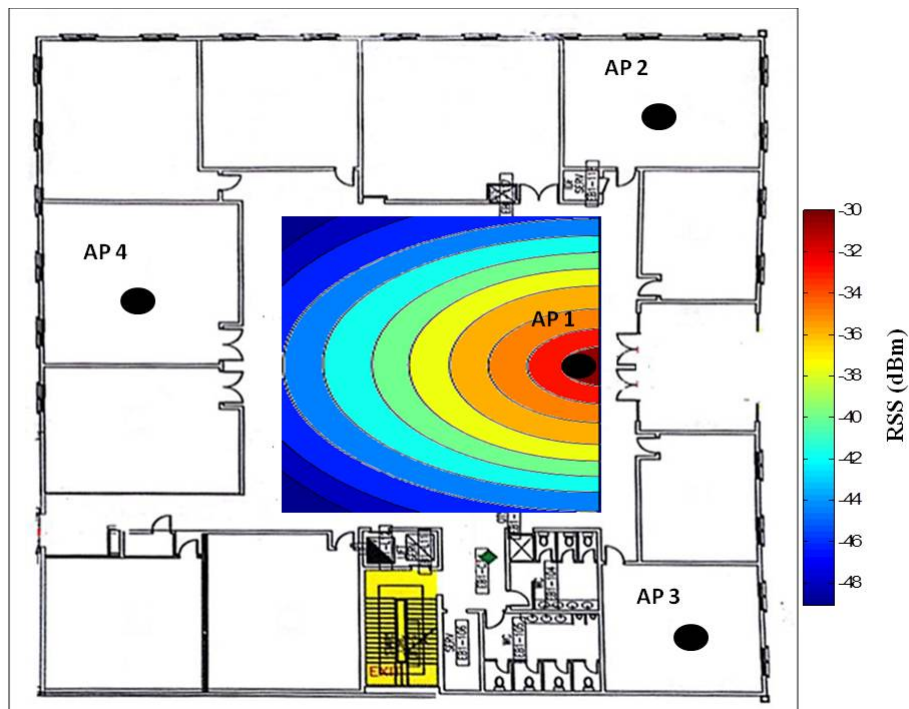


Figure 20: Illustration of propagation modeling for one AP in the experimental area.

Our proposed system is tested on two stages. Firstly, Kalman filters are used to correct RSS values for different propagation models: the PathLoss and the GPR

model. The distance estimation of both propagation models are plotted against the true distances for RSS values ranging from -35 to -65 dBm. The performance is shown in Figure 21 and it can be seen from the figure that the GPR performance is much better than PathLoss especially for low RSS values. The achieved average error using Kalman filters and GPR models is 3.635 m. However, the estimation error of PathLoss models is still very high (8.53 m) even when it is integrated with Kalman filters. Therefore, our results prove that GRP models outperform PathLoss models in distance estimation.

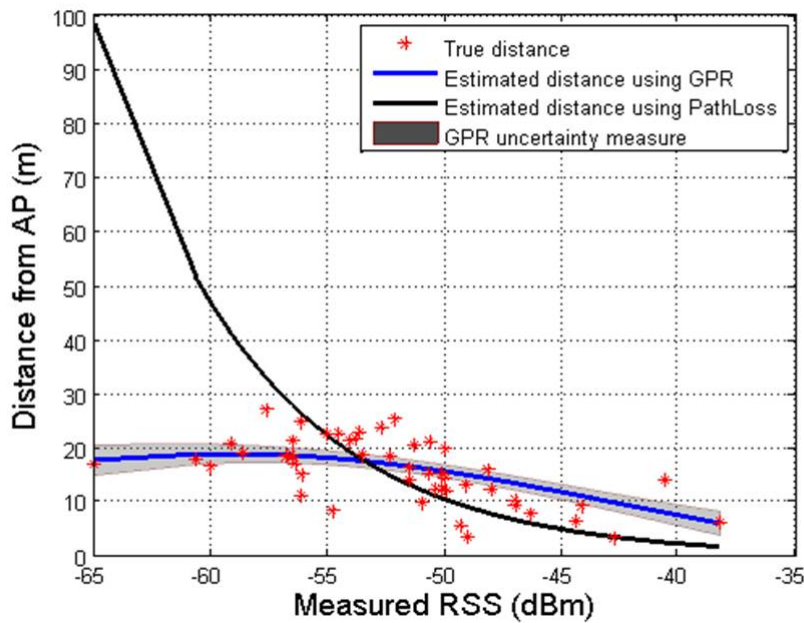


Figure 21: Distance estimation using GPR and Pathloss.

The second part of our proposed system evaluation is to test the GPR models performance when integrated with different RSS correction techniques. The distance estimation error cumulative density function (CDF) of the GPR models integrated with different RSS correction methods is shown in Figure 22. As shown in the figure, Kalman filters result in the best performance achieving a lower distance error for high probabilities. Other RSS correction methods, such as using the statistical mean and median of the noisy measurements, result in a worse performance with higher distance error. The average of distance error estimation for GPR and the three RSS correction

methods is shown in Table 7. As summarized in the table, integrating Kalman filters improves the distance estimation accuracy by almost 2 m when compared to other RSS correcting methods. More specifically, integrating GPR models and Kalman filters result in the best error performance amongst propagation models and RSS correction methods.

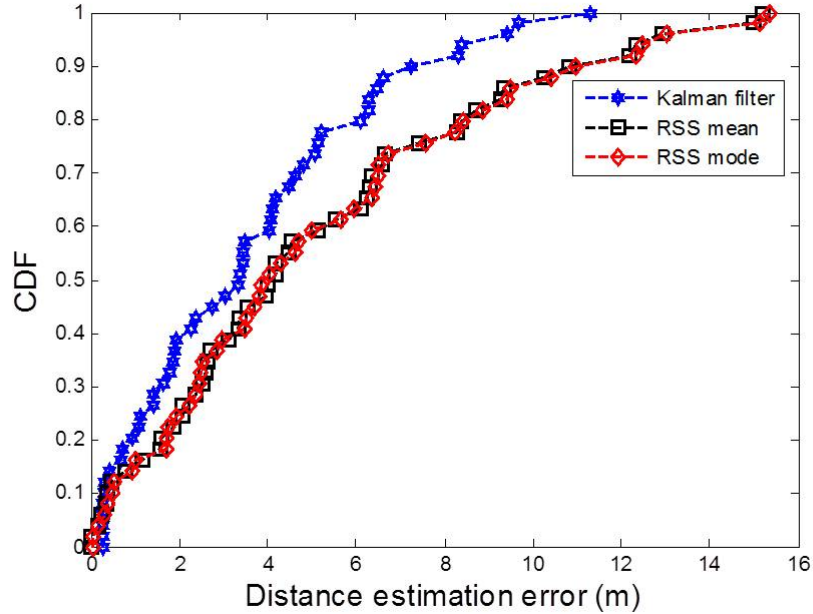


Figure 22: GPR performance using different RSS correction methods.

Table 7: Accuracy of GPR distance estimation using different correcting methods.

Correcting method	Kalman filter	Mean	Mode
Accuracy (m)	3.61	5.17	5.22

It is important to note that propagation models are highly dependent on the observability of the APs that are employed by the positioning system. In other words, informative APs should be used to guarantee the good performance of distance estimation. To illustrate, the integrated Kalman and the GPR models were used to estimate distances from a non informative AP. It was found that the average distance error has

increased to 4.7670 m. This proves that the position of the AP and its observability of the targets are highly important. Therefore, we suggest using the MFOS technique proposed in Chapter 5 to learn the APs that should be used by propagation models to assure accurate distance estimation.

Chapter 7: Conclusions

Advances in smartphone technology, coupled with the increasing demand for context aware and location based services, have initiated the need for a new generation of positioning systems. While existing global positioning systems (GPS) have good performance outdoors, they perform poorly inside buildings. This thesis studied different indoor positioning techniques and approaches which are fingerprinting and propagation models. In particular, we proposed a hybrid solution of clustering and fast search techniques to improve fingerprinting techniques by reducing their computational requirements. Our results showed that the proposed search approach has a comparable performance to the full search with a much reduced number of searching points. Moreover, the study presented in this thesis highlights the tradeoff between the system performance and the required number of search points; the less the number of search points, the worse the performance. With the hybrid solution, the accuracy-complexity tradeoff is fortunately violated, allowing further complexity reduction while maintaining high accuracy.

In addition, the FOS and the modified FOS algorithms were implemented in order to reduce the dimensionality of the radio map to a matrix containing only the most informative APs to the positioning system. Moreover, the modified FOS, which is a variant of the FOS, is proposed in this thesis where the reduction in error is calculated to determine the APs that result in the unique RSS values. Both these algorithms were compared to the PCA reduced radio map, heuristically reduced radio map and the full radio in terms of the positioning error. Results illustrate that the FOS provides lower error when compared to the PCA, however, modified FOS outperforms its conventional FOS counterpart and has a performance very close to the performance of heuristic and full radio map. Furthermore, the FOS and the MFOS are less computationally intensive than the PCA and therefore have lower processing time. Later, the MFOS was combined with the hybrid search solution proposed previously which resulted in a better positioning accuracy and a faster processing.

Furthermore, we have studied propagation models based positioning techniques. Specifically, the GPR propagation models were investigated and their performance was

compared against conventional propagation models, such as PathLoss. In addition, a system that integrates Kalman filters with GPR was proposed in this thesis. The objective of such system design is to enhance the performance of GPR models by eliminating the RSS temporal variations through Kalman filtering. It is found that our proposed system improves distance estimation by almost 2 m. Our objective is hence achieved which is to improve the modeling of signal propagation inside buildings at the lowest cost.

As part of future work, more clustering techniques can be considered for different cases of positioning algorithms. Also, the effect of other involved parameters is to be investigated, such as the number of grid points and the area of the test environment. In addition, the amount of reduction suitable for different positioning cases will be studied. MFOS will also be investigated to be used in radio map partitioning for large scale areas. On the other hand, The effect of the number of samples on the performance of Kalman filters will be studied. Moreover, the latency resulting from samples collection will also be considered. In addition, the distance estimation of the GPR models can be enhanced by assuming a suitable mean for the estimates. Finally, the estimation of an exact position of a user, using the distance estimates and triangulation techniques, will be taken into account.

References

- [1] H. A. Karimi, *Advanced location-based technologies and services*. Boca Raton, FL: CRC Press, Taylor Francis Group, 2013.
- [2] ATAP, “Project tango,” 2013. [Online]. Available: <https://www.google.com/atap/project-tango>[Accessed:Apr-04-2015]
- [3] M. Youssef, A. Agrawala, and A. Udaya Shankar, “WLAN location determination via clustering and probability distributions,” in *PerCom 2003. Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, Mar 2003, pp. 143–150.
- [4] A. Sayed, A. Tarighat, and N. Khajehnouri, “Network-based wireless location: challenges faced in developing techniques for accurate wireless location information,” *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 24–40, Jul 2005.
- [5] A. M. Ladd *et al.*, “Robotics-based location sensing using wireless ethernet,” *Wireless Networks*, vol. 11, no. 1-2, pp. 189–204, 2005.
- [6] E. Mohammadi, “Indoor location based services,” Master’s thesis, University of Calgary, 2011.
- [7] A. Roxin *et al.*, “Survey of wireless geolocation techniques,” in *2007 IEEE Globecom Workshops*, Nov 2007, pp. 1–9.
- [8] H. Liu *et al.*, “Survey of wireless indoor positioning techniques and systems,” *Systems, Man, and Cybernetics, Part C: IEEE Transactions on Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.
- [9] M. M. M. Atia, “Nonlinear estimation techniques for high-resolution indoor positioning system,” Ph.D. dissertation, Queen’s University, 2013.
- [10] J. Coastabile, “Wireless position location,” Virginia Polytechnic Institute and State University, Wireless Symposium Presentation, 2010.
- [11] B. A. Forouzan, C. A. Coombs, and S. C. Fegan, *Data communications and networking*. Boston: McGraw-Hill, 2001.
- [12] E. C. L. Chan, G. Baciuc, and I. ebrary, *Introduction to wireless localization: with iPhone SDK examples*. Singapore: IEEE, 2012.
- [13] K. Yu, I. Sharp, and Y. J. Guo, *Ground-based wireless positioning*. Chichester, U.K; Hoboken, NJ: Wiley, 2009.
- [14] M. Atia, A. Noureldin, and M. Korenberg, “Dynamic online-calibrated radio maps for indoor positioning in wireless local area networks,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 9, pp. 1774–1787, Sep 2013.

- [15] M. Youssef and A. Agrawala, "The horus WLAN location determination system," in *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '05, 2005, pp. 205–218.
- [16] P. Bahl and V. Padmanabhan, "Radar: an in-building RF-based user location and tracking system," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the Proceedings IEEE Computer and Communications Societies*, vol. 2, 2000, pp. 775–784.
- [17] J. Ma *et al.*, "Cluster filtered KNN: A WLAN-based indoor positioning scheme," in *WoWMoM 2008. International Symposium on a World of Wireless, Mobile and Multimedia Networks*, June 2008, pp. 1–8.
- [18] R. Muatz, "Application for venia legendi in positioning and engineering geodesy," Ph.D. dissertation, Institute of Geodesy and Photogrammetry, Department of Civil, Environmental and Geomatic Engineering, 2012.
- [19] K. Kaemarungsi and P. Krishnamurthy, "Analysis of WLAN's received signal strength indication for indoor location fingerprinting," *Pervasive and Mobile Computing*, vol. 8, no. 2, pp. 292 – 316, 2012, special Issue: Wide-Scale Vehicular Sensor Networks and Mobile Sensing.
- [20] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *MOBIQUITOUS 2004. The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, Aug 2004, pp. 14–23.
- [21] D. Liu *et al.*, "Analysis of wireless localization in nonline-of-sight conditions," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 4, pp. 1484–1492, May 2013.
- [22] B. Altintas and T. Serif, "Improving RSS-based indoor positioning algorithm via k-means clustering," in *11th European Wireless Conference 2011 - Sustainable Wireless Technologies (European Wireless)*, Apr 2011, pp. 1–5.
- [23] H. Zhou and N. Van, "Indoor fingerprint localization based on fuzzy c-means clustering," in *2014 Sixth International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Jan 2014, pp. 337–340.
- [24] A. Agiwal, P. Khandpur, and H. 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*, ser. WMASH '04, 2004, pp. 102–109.
- [25] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. New York: Wiley, 2001.
- [26] H. Li, X. Zhao, and M. Tan, "A new search strategy of radio fingerprint matching method in wireless sensor network," in *2011 First International Conference on Robot, Vision and Signal Processing (RVSP)*, Nov 2011, pp. 18–22.

- [27] M. Atia, M. Korenberg, and A. Noureldin, "Fast features reduction of radio maps for real-time fingerprint-based wireless positioning systems," *Electronics Letters*, vol. 47, no. 20, pp. 1151–1153, Sep 2011.
- [28] J. Ledlie, "Method and apparatus for on-device positioning using compressed fingerprint archives," Jun. 9 2011. [Online]. Available: <http://www.google.com/patents/WO2011067466A1?cl=en>[Accessed:Apr-20-2015]
- [29] E. Laitinen *et al.*, "Access point significance measures in WLAN-based location," in *2012 9th Workshop on Positioning Navigation and Communication (WPNC)*, Mar 2012, pp. 24–29.
- [30] L. Wirola, L. Wirola, and R. Piche, "Bandwidth and storage reduction of radio maps for offline WLAN positioning," in *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2013, pp. 1–9.
- [31] M. Atia, A. Noureldin, and M. Korenberg, "Dynamic propagation modeling for mobile users' position and heading estimation in wireless local area networks," *IEEE Wireless Communications Letters*, vol. 1, no. 2, pp. 101–104, Apr 2012.
- [32] R. Mehra and A. Singh, "Real time RSSI error reduction in distance estimation using RLS algorithm," in *2013 IEEE 3rd International Advance Computing Conference (IACC)*, Feb 2013, pp. 661–665.
- [33] M. Zhou *et al.*, "Correlated received signal strength correction for radio-map based indoor Wi-Fi localization," in *2014 International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Jul 2014, pp. 1–6.
- [34] M. Youssef and A. Agrawala, "Handling samples correlation in the horus system," in *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2004.*, vol. 2, Mar 2004, pp. 1023–1031 vol.2.
- [35] Z. Hui-qing, X. Jian-bo, and Z. Jing-li, "Multi-fingerprints joint indoor positioning algorithm in WLAN based on Kalman filter and database of partitioned ap id," in *The 26th Chinese Control and Decision Conference (2014 CCDC)*, May 2014, pp. 4055–4059.
- [36] X. Jian-bo, Z. Hui-qing, and Z. Jing-li, "Self-adapting multi-fingerprints joint indoor positioning algorithm in WLAN based on database of AP ID," in *2014 33rd Chinese Control Conference (CCC)*, Jul 2014, pp. 534–538.
- [37] J. Yin, Q. Yang, and L. Ni, "Adaptive temporal radio maps for indoor location estimation," in *Third IEEE International Conference on Pervasive Computing and Communications, 2005. PerCom 2005.*, Mar 2005, pp. 85–94.
- [38] H. Wang *et al.*, "Dynamic radio map construction for WLAN indoor location," in *2011 International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 2, Aug 2011, pp. 162–165.

- [39] X. Cai, L. Chen, and G. Chen, “Constructing adaptive indoor radio maps for dynamic wireless environments,” in *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing, and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC)*, Dec 2013, pp. 41–47.
- [40] J.-N. Kim and T.-S. Choi, “A fast three-step search algorithm with minimum checking points using unimodal error surface assumption,” *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 638–648, Aug 1998.
- [41] S. Zhu and K.-K. Ma, “A new diamond search algorithm for fast block-matching motion estimation,” *IEEE Transactions on Image Processing*, vol. 9, no. 2, pp. 287–290, Feb 2000.
- [42] H. Abbas, “A novel fast orthogonal search method for design of functional link networks and their use in system identification,” in *ISIC 2007. IEEE International Conference on Systems, Man and Cybernetics*, Oct 2007, pp. 2743–2747.
- [43] L. I. Nahlawi, “Genetic feature selection using dimensionality reduction approaches: A comparative study,” Ph.D. dissertation, Queen’s University, 2010.
- [44] D. Simon, *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J: Wiley-Interscience, 2006.
- [45] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. Cambridge, Mass: MIT Press, 2006.
- [46] M. Abden, “Gaussian process for regression: A quick introduction,” Tech. Rep., Mar 2008.

Vita

Ayah Mahmoud Abusara was born in 1991, in Jordan. At an early age, her family moved to Dubai, United Arab Emirates, where she was raised and educated. In, 2009, she ranked the 4th in city of Dubai and the 9th nationwide in the public secondary exams. She was a recipient of the Abu Dhabi Prince Court Scholarship to pursue her Bachelor of Science degree from the American University of Sharjah (AUS), Sharjah. She joined the electrical engineering program at AUS and graduated in 2013 with a GPA of 3.82 out of 4.0 (magna cum laude). She was also among the top 10 graduate achievers. In the same year, Ayah was admitted to the graduate studies program at the American University of Sharjah where she was awarded a two-year teaching assistantship. During her Masters, she had two conference papers accepted and presented in two different IEEE conferences; ICCSPA15 and ICMSAO15. She simultaneously worked as a teaching and research assistant, where she was responsible for several laboratory, grading, and tutoring tasks, in addition to working toward this thesis. Her research areas of interest include localization systems, mobile and wireless networks, probabilistic modeling, nonlinear estimation and cognitive radio networks.