FINGERPRINTING BASED ENHANCED INDOOR POSITIONING SYSTEM USING MULTIWALL PROPAGATION MODEL

MS THESIS

Naureen Mujtaba

170357

A Thesis Submitted to the Faculty of Electrical Engineering Department, Military College of Signals, National University of Sciences and Technology, Islamabad, Pakistan, in partial fulfilment of the requirements for the degree of MS in Electrical (Telecommunication) Engineering

January 2019

بسم الله الرحمن الرحيم

Thesis Acceptance Certificate

Certified that final copy of MS/MPhil thesis written by Miss **Naureen Mujtaba,** Registration No. **00000170357** of **Military College of Signals** has been vetted by undersigned, found complete in all respect as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial, fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the student have been also incorporated in the said thesis.

Signature: ______________________________

Name of Supervisor: **(Asst. Prof. Mir Yasir Umair, PhD)**

Date: ______________________________

Signature (HoD): ______________________________

Date: ______________________________

Signature (Dean): ______________________________

Date: ______________________________

Abstract

Nowadays, demands for location-based services (LBS) are going to rise day by day for indoor environments. Due to the unavailability of the extensively used Global Positioning System (GPS) in indoor environments, many other technologies and techniques are followed to fulfil these demands. Amongst all Indoor positioning techniques, Wi-Fi fingerprinting technique has appealed significant interest due to its potential to achieve maximum accuracy at minimum cost.

Efficient systems are important to minimize delays, complexity and the associated additional costs. Positioning accuracy of Wi-Fi indoor positioning systems highly depends upon offline databases. Therefore, Development of robust Wi-Fi fingerprints is performed to improve the positioning results. For this reason, improved Motely Keenan propagation model is used. Additionally, based on original implementation, an Indoor Positioning System (IPS) is developed by utilizing different matching algorithms e.g. k-Nearest-Neighbors (kNN) algorithm, Feature Scaling k-Nearest-Neighbors (FS-kNN) algorithm. And then performance of all matching algorithm is compared with each other. Moreover, a new positioning algorithm is proposed to improve the localization accuracy. The development of the proposed algorithm is mainly based on environmental dispersion. To mitigate the effect of that dispersion, a procedure named as 'Deleting Outlier' is used with existing technique i.e. FS-kNN. The simulation results show that proposed method is superior to some previous methods and achieves an exceptional accuracy with an average positioning error of approximately 1.45m in 30m × 33m area using an up-to-date fingerprint database.

On the basis of the results of this research, it can be concluded that it is possible to use Wi-Fi fingerprinting for indoor positioning to obtain a state-of-the-art accuracy.

Acknowledgments

First, thanks to almighty ALLAH and His Prophet (P.B.U.H) for helping me during my whole life and to finish this work.

I would like to express my sincerest gratitude and appreciation to Dr. Mir Yasir Umair, Assistant Professor in Military college of signal, NUST, for his supervision, careful guidance, great support, and prestigious contributions during the completion of this work. Without his guidance and support, the completion of this research work was not possible.

Special thanks to my Parents, my sisters and my brothers, for their countless prayers, love, continuous encouragement, kind support and providing the stress-free environment during my whole study period.

I would also like to thank my best friend Aisha Javed for her advice and excellent support she provided throughout the simulation and analysis of this work.

In last but not least, I would like to thank the Military College of Signal NUST for providing me such a research-oriented environment, labs and skilled faculty who guided me during my whole MS research period.

Table of Contents

List of Figures

List of Tables

Acronyms

GPS – Global Positioning System

WLAN – Wireless Local Area Network

APs – Access Points

RPs – Reference Points

RSSI – Received Signal Strength Indicator

LOS – Line of Sight

NLOS – Non- Line of Sight

UWB – Ultra-Wideband

IR – Infrared

vSLAM – Visual Simultaneous Localization and Mapping

RFID – Radio Frequency Identification

AOA – Angle of Arrival

TOA – Time of Arrival

TDOA – Time Difference of Arrival

TOF – Time of Flight

ILBS – Indoor Location-Based Systems

IPS – Indoor Positioning System

kNN – k-Nearest Neighbor

FS-kNN – Feature Scaling K-Nearest Neighbor

Cos-kNN – Cosine K-Nearest Neighbor

SA – Simulated Annealing

Chapter 1 - Introduction

This chapter provides introduction of the purpose of this thesis. Further in this chapter, problem statement, motivation to choose this area and applications are discussed. The objectives and methods to achieve these objectives and the research approach are introduced to the end of the chapter. The last section provides an outline of the chapters to come.

1.1. Indoor Localization

In the recent years, the increase in demand of location-based services and Positioning systems have been seen for both people and objects. Due to this reason, the development in the field of positioning and localization has evolved immensely in the last decades. Localization is the process to find the position or location of a specific target based on some observable phenomena.

Positioning services have become widespread with the expansion of modern communication technologies. The growing diversity of commercial uses has conventional demand for indoor position. Global Positioning System (GPS) provides adequate services for positioning and localization in outdoor environments, but the performance of GPS degrades in indoor environments and urban areas because of the reception of weak signal and the lack of communication based on line-of sight between the target and the satellites, thus researchers have proposed and developed different localization systems for localization in indoor environments [1].

Multiple solutions and systems have been developed for indoor localization; each has its own advantages and limitations. The most widely adopted is Fingerprinting technique as it provides better accuracy, although the practical implementation is quite laborious, but the working is very simple or least complex as compared to other localization techniques. Also, it does not require any additional equipment and can be implemented using existing infrastructure.

Localization can be used for numerous applications like localizing soldiers in combat, collecting marketing data, tracking endangered species, navigating self-driving cars, robot movement, indoor localization for firefighters, hospitals and malls, to provide navigation aid. While others demand indoor localization to make better markets for the customers, to find the exact location of products placed in the warehouses, for the automatic detection of object location, detection of medical equipment in hospital, detection of fireman in the fired building, location detection of dog trained by police to find explosives in building as well as finding tagged maintenance equipment which is scattered all over the plant, given that position aware tour guide systems and many more.

Indoor localization systems also have a key role in automatic detection of objects or tracking products based on their locations. Location detection of a baggage in a large indoor environment, location detection of any lost product in a store, equipment scattered over a large plant or factory. Similarly, location tracking and updates of firefighters in a building caught by fire, keeping track of a patient's movement in intensive care, tracking activities of a suspicious person and many more. Several techniques of localization have been introduced for the purpose of positioning and navigation [2, 3].

The technique of Fingerprint positioning is Received Signal Strength (RSSI) based positioning that uses location dependent characteristics with a location and uses these characteristics to estimate the position. Fingerprint based indoor positioning consists of two phases named as offline and online phases. Fingerprint database is established in offline phase and positioning is performed in online phase. RSS values that are obtained from access points on a specific time period at the predetermined reference points are stored in a database that is called as fingerprint database. Each fingerprint in this database contains basically location information and RSS values obtained from surrounding access points at that location. The basic lay out of fingerprinting based indoor positioning is represented in Figure 1.1.

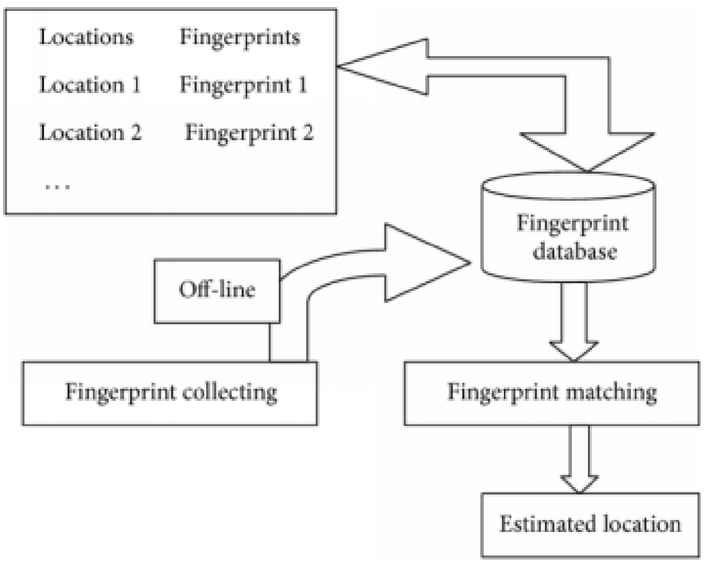


Figure 1.1: Location fingerprinting method [2]

In the offline phase the RSS fingerprints are sampled at each reference position from a number of access points (AP). APs are usually fixed transmitter such as Wi-Fi routers. A reference position (RP) is the position in the indoor environment that needs to be tracked. And the signal fingerprint at each RP are updated or inserted in the survey database. The received signal strength (RSS) can be defined as the measure of signal power from an AP to a receiver which can be sampled without any additional requirement in WLAN environment. Single samples of RSSI recorded from the nearby access point are not sufficient to characterize a fingerprint. Due to presence of noise in the environment it is necessary to obtain an average of the readings to successfully identify a fingerprint. In the online phase the background service running on the user's mobile measures a vector of RSS value at an unknown location and then compares the RSS value received in the online determination phase with the training database. With the help of positioning algorithm the most likely location of mobile user is finally calculated. Fingerprinting helps to significantly improve the accuracy and precision of conventional signal strength lateration techniques.

1.2. Statement and Problem

➢ Indoor Positioning System using Wi-Fi Techniques are very popular in recent years as the buildings are now equipped with the low-cost Wi-Fi APs that can be used for indoor positioning references. However, the accuracy and robustness of this technique is a challenge because the radio signals of Wi-Fi equipment are non-linear, non-line of sight and Multipath in nature.

➢ Many different techniques are proposed to improve the accuracy and robustness of the indoor positioning system in literature [4-14] but they depend on the cost of equipment, computation loss of microprocessor, plan of the building and the implementation.

➢ Due to the geographical uncertainty, environmental dispersion causes an unexpected error in localization.

1.3. Motivation

If there is any unpleasant situation, emergency situation, fire in a building or any other emergency situation, the security and law enforcement personnel need to know the number of people that are affected exactly and their exact locations in order to rescue them safely. Having simple knowledge of number of people in the building or a large number of travelers at an airport are not enough to tackle such situation efficiently. The need is to know the exact locations of the individuals as this can be a matter of life and death.

So, radio systems which have capability of positioning have evolving applications in homeland security, law administration, emergency retort, security command and control and battleground command and control. Indoor location and positioning systems have become very popular in recent years because of the emerging technologies and inventions also. Internet of things, automation, directions and navigations, robotics, self-driven vehicles etc., all of these technologies need localization systems to be implemented within.

During the literature review [25-30], it has been observed that each localization technique has its own advantages and satisfies the demands but on the other hand also has some drawbacks and limitations. Limitations can be in case of accuracy, range, system cost, complexity or other. So, in the development of a new system or technique the pros and cons have to be considered, and characteristics or parameters are very carefully chosen.

Therefore, this research concerns with two main strategies: 1- To build the radio map, a recently developed propagation model for wireless environment is used to calculate total path loss and provide exact RSSI vector to every reference point. This model is named as Motely Keenan propagation Model and it not only calculates number of walls but also examines the wall types between transmitters and receivers. 2- The development of a Wi-Fi fingerprinting technique for indoor positioning. Second objective contains a proposal of new matching algorithm for positioning accuracy. For this reason, deleting outlier procedure is merged with already existing matching algorithm. Although various studies have currently dealt with Wi-Fi fingerprinting for indoor positioning, the novelty of this research is that it will endeavor to obtain better accuracy resulting from a novel implementation methodology. The widely used k-Nearest-Neighbors (kNN) algorithms will be adjusted with a prior calculation. The algorithm also takes the reliability of the measurements into account by using heat map of databases. Some more advance matching algorithms (FS-kNN, Cosine kNN) are used to obtain high accuracy regarding localization. A new proposed algorithm has developed to obtain more accurate positioning results. The prototype could be used to navigate from one place in a building, to another place in the building in the future.

1.4. Objectives

Development of a Wi-Fi-based fingerprint indoor positioning system is the main focus of this research. The main requirement for getting the aim is to achieve the following individual objectives:

- Explain why Wi-Fi is chosen for the indoor positioning system.
- Search and implementation of the best location algorithms for fingerprinting.

- To propose an indoor localization technique with better accuracy.
- To achieve good results and maximum accuracy.
- To propose a technique that is not complex to implement using existing infrastructure and is easy to understand.

1.5. Summary of Contributions

The contributions of the work in this thesis are as follows:

- Referencing of the whole floor plan.
- Deployment of access points of user defined coordinates.
- Implementation of a modified propagation model in wireless environments.
- Verification of RSSI vectors by propagation map regarding all access points.
- Implementation of KNN algorithm.
- Implementation Feature Scaling KNN algorithm.
- Development Feature Scaling KNN algorithm with deleting outlier.

1.6. Outline

The Chapter 1, (current chapter) gives an introduction to the research conducted. It also justifies the importance of the research, motivation and its relevance to national needs. Level of research already carried out in the field. The objectives of the research are addressed. The Chapter 2 consists of all possible detail about indoor positioning systems i.e. Literature Review. It gives detail about the RSSI based fingerprinting and introduction of the fingerprinting technique. All stages of the fingerprinting technique are presented in detail. Chapter 3 explains the proposed technique, System model of the simulations, mathematical models and equations used in the simulation. Chapter 4 Results and Discussions, presents the results of the simulations, the basic RSS based technique and proposed techniques are tested in indoor scenarios and compared. Chapter 5 is Conclusion and Future work. Chapter 6 contains References and Bibliography.

Chapter 2 - Literature Review

This chapter is mainly based on literature review of the research. In this part of the thesis, an overview of previous developments of indoor positioning has done, which covers the most famous technologies, techniques, propagation model and matching algorithms used for localization. Performance matrices, their weaknesses are discussed as well with a specific focus on fingerprinting-based localization techniques.

2.1. Previous Developments

An IPS is a system to track an object or person situated within a building where the GPS is inadequate. An IPS is often implemented by the use of a portable, sometimes wearable, device. There are a few non-radio [4, 5, 7-14] and wireless [15-22] advancements that have been considered over the most recent couple of years that could be utilized for localization. Current and progressing research focuses for the most part around wireless indoor localization techniques. The vast majority of these methodologies utilize Wi-Fi or Bluetooth signals, exploiting advantage wireless infrastructures as of now deployed in the building.

2.1.1. Non-radio technologies

In the beginning of indoor positioning, when no wireless technologies were available in access, then some non-radio technologies were used in indoor positioning systems. Some of them are still serving according to the requirements. The detail of these technologies is as follows:

2.1.1.1. Magnetic Positioning

It is a general practical method which is based on building's iron structure. These vibrations are sensed by smartphones in the earth's magnetic field as dedicated chips inside. For the mapping of building and achieving the localization, these vibrations are used [5, 6].

2.1.1.2. Acoustic/Sound Positioning

Use of acoustic or sound signals is another non-radio technology for indoor localization. In indoor positioning system, Guoguo [7] achieves positioning accuracy at centimeter-level for several indoor environments. Therefore, extent over the space, an anchor network of low cost devices is utilized by the systems. Modulated localization beacon signals are transmitted by each node using high bandwidth acoustic signal. These signals are processed by a smartphone and localization is performed using a banked server [8].

2.1.1.3. Inertial Measurements

Dead reckoning systems mostly use this method with the help of a device with Inertial Measurement Unit (IMU) which a tracked object or a person carries with him. In modern smartphones the availability of different sensors give benefits to these systems. Tracing the movement of an object or a person is possible with the help of suitable technique on the sensors data. This system is the main focus of recent researches for the pedestrian tracking systems on a commercial off the shelf smartphone. Inertial Measurement Unit which is built-in smartphone is mainly used by tracking systems. During walking, three typical carrying modes of smartphone are identified by system which is utilized to optimize the tracing accuracy [2]. If the mobile user's initial position, step length, step detection and the way of holding a cell phone, is known then he would able to locate himself in the given scenario. PDR system at the same time is robust as well as accurate for the people with their different height, gender and walking speed [3]. For concurrent tracing and positioning of the mobile user, sub-meter error accuracy is obtained while walking a pathway of 28m distance [9].

2.1.1.4. Dead Reckoning

By deliberate the previous location of the target and speed by which it travels, in dead reckoning, present location of a target can be roughly identified. The dead reckoning is a positioning technique which has a need of a known position to start its working. Then it will add and trace changes. These progressions can be as Cartesian axis or speed. With adequately numerous complete location updates, dead reckoning straightly developing location inaccuracies can be delimited inside pre-characterized bounds [10]. Dead reckoning must to utilize other techniques to modify the location afterward every break to enhance the accuracy as well as decrease errors, [11]. Pedestrian dead reckoning is a case of dead reckoning that basically evaluates steps extent and route of a moving individual [10].

2.1.1.5. Ultrasonic

A mechanical wave which is a wavering of compression that is transferred by the medium is for the most part known as Ultrasonic wave. It has very short range and doesn't interfere with electromagnetic waves. Air and structure substantial is utilized as transmission media in ultrasonic localization systems. Time of Arrival (TOA) estimations of Ultrasound beats is utilized to gauge the relevant distance or range between gadgets. To settle collectors deployed at known positions, an approximation of directions of the emitter is conceivable by trilateration using three series [12].

2.1.1.6. Visual Positioning

Visual positioning has two directions. First is to determine the position of a user with the help of a camera enabled mobile device. Visual makers are used in most of the methods in this environment. It is conceivable to measure the position of the gadget by finding and translating these visual makers and computing point to the marker from the gadget. In the second direction, database of images is used by visual positioning systems [13]. Interpolation on the image database is used to estimate its position using mobile devices. Simultaneous Localization and Mapping (SLAM) is a typical usage of visual positioning. Problem of creating databases of a new scenario by a mobile robot, is directly concerned by the SLAM while at the same time that map is using for navigating the environment. SLAM always tries to use numerous altered forms of antennas plus optical antennas, called as Visual Simultaneous Localization and Mapping (vSLAM) [14]. vSLAM algorithms are typically, vision- andodometry-based. Low cost navigation in disorderly and colonized environments is enabled by these systems. vSLAM robot has the ability to search its location without user-intervention as there is no initial map is required. Due to this ability, the robot is in a position to build a reliable map and localizes itself from the map [15].

2.1.2. Wireless Technologies

A standout amongst the most well-known methods for indoor localization is utilization of radio-frequency indications. These frameworks normally utilize idea of RSS, a sign of acknowledged signal's strength level estimated with the help of recipient, for localization. The inverse square law relates to radio waves transmission; hence distance assurance is conceivable in light of the connection among transmitted and received signal power. This makes radio- frequency signal especially beneficial for positioning, generally Bluetooth [16, 17] or Wi-Fi based localization systems [18, 19] are utilized.

2.1.2.1. Radio Frequency Identification (RFID)

An arrangement which transfers distinctiveness of the object or individual wirelessly by means of radio waves is portrayed by Radio Frequency Identification. In immense systems RFID technology is normally used to naturally distinguish entities. There are two main components of RFID, on which different frequencies of radio signals are exchange i.e. RFID readers and RFID tags. These Radio signals are emitted by RFID tags and received by RFID readers and vice versa. Objects that need to be traced are attached to RFID tags. For sending and receiving the data amongst them, both use predefined radio frequency and protocol. RFID mainly consists of a microchip with memory storage up to 2kB as well as a radio antenna. There is an active and passive tag, whereas RFID reader is consisted on various components i.e. power supply, antenna, transceiver, processor, and an interface in order to connect to a server [20, 21].

The RFID reader is placed in the building to a fix position and the passive RFID tag is attached to user cell phone or name badge etc. location is estimated when the passage of RFID tag with this badge. This system needs to be close passage for the implementation to prevent it to go out of the reader's range. The operation principle is same for localization using Bluetooth or Wi-Fi techniques. These both wireless technologies are well established in modern smartphones. 86% accuracy has been achieved by Bluetooth technology [22]. The basic advantage of these technologies is that the network infrastructure is present already and various access points are placed in fixed position. So there is no need for an extra investment in specialized hardware.

2.1.2.2. Ultra-Wideband (UWB)

Bandwidth greater than 500MHz of a RF signal is defined as UWB by the FCC (Federal Communications Commission). UWB is a communication channel which uses a wide portion of frequency to broadcast information. UWB emitter transmits a large amount of data by consuming very low amount of energy [23]. This technology is likewise utilized for localization by using time of arrival (TOA) and time difference of arrival (TDOA) of RF signs to acquire the distance among the entity and the reference point.

Millimeter accurate positioning by consuming low power is only allowed by RF technology [24, 25]. But most of the mobile devices are not equipped with this technology up till now. But it is expected that in near future mobile devices will be embedded with this technology.

In June 1997 the IEEE 802.11 WLAN standard was approved. Standard characterizes decorum and good inter-connection of data communication hardware via air in a local area network (LAN) utilizing transferor sense numerous access decorum with collision avoidance (CSMA/CA) medium sharing system. With a run of a range of 50 to 100 meter and the mill net piece rate of 11, 54, or 108 Mbps, nowadays IEEE 802.11 is the most leading local wireless networking standard [20]. Utilizing Wi-Fi in indoor localization and triangulation framework relies upon deliberate a list of wireless routers that are accessible in a zone in which the framework works in. Most famous WLAN localization technique is RSS (Received Signal Strength) that is informal to abstract in 802.11 systems and can keep running on off the-shelf WLAN equipment. Time of Arrival (ToA), Time Difference of Arrival (TDoA), and Angle of Arrival (AoA) techniques are fewer widespread in WLAN because of the many-sided quality of time suspension and angular estimations. Accuracy of normal WLAN localization technique utilizing RSS is around 3 to 30 meters, with an up to dated rate in the series of a small number of seconds [21].

2.1.2.3. Zigbee

Functioning on topmost of the IEEE 802.15.4 description, the Zigbee standard offers networks, privacy and application support services. This is the wireless personal area network with very short distance and low rate even that a basic Zigbee hub has least cost with low complexity [29]. It depends on two parts which are multichannel dual technique radio and a microcontroller on solitary bit of silicon. Zigbee is basically considered for those requests which oblige less power utilization and less data material. Localization is achieved by coordination and communication with neighboring nodes by using this technology. There are two physical devices utilized for Zigbee: Full Function Device (FFD) and Reduced Function Device (RFD) [12].

2.1.2.4. Bluetooth

Radio standard utilized for radio personal area network (WPANs) is Bluetooth. Unlike Zigbee, Bluetooth is an exclusive arrangement accomplished by Bluetooth Special Interest Group (SIG). To work in the 2.4 GHz band, Bluetooth is supposed to design as a low power technology for shared communication [16].

Its gross bit rate is also very low (1Mbps) compared to WLAN as well as shorter range say 10 cm to 10 meters. For investigation of localization via Bluetooth wireless technology, there exists a local positioning group from various Bluetooth SIG groups [12].

2.1.2.5. Cellular Based

With lower positioning accuracy, Global System for Mobile Communications (GSM) frameworks is offered in many republics which outreach the exposure of WLAN. In contrast with WLAN, for the prevention of interference from other devices using analogous frequency [32], GSM operates in licensed bands. Using cellular network, indoor positioning is possible but there is a condition that building must cover b number of base stations or one base station along with enough RSS strength expected by indoor mobile users. Fingerprinting created by power level (RSS) is the most commonly used method of GSM indoor positioning [20].

2.1.2.6. FM

This technology also used for indoor positioning system. Almost every mobile phone is equipped with FM receivers and one get FM station easily in urban area so it does not require extra infrastructure. FM radio also uses the same techniques for positioning as other do but this system has low accuracy for localization.

2.1.2.7. Infrared Signals

Active badge system is one of the first IPS that is introduced by R. Want et al. in 1992. Similar to the name badge, they developed a wearable badge which emits IR signals. For making the possibility to locate the person, receivers are placed at different specified positions in the building. Of course, there are number of limitations. First of all, there must be line of sight among receivers and the badge. Short range of transmission signal is another limitation of infrared, that's why a number of receivers are, needed by which the cost is bring upward while the accuracy of the location is directly related to number of receivers [4].

Techniques	Technology	Localization
Fingerprinting	RSS	Central
Signposting	RSS	Central or Distributed
Trilateration	RSS, TOA, AOA	Central or Distributed

Table 2.1: Overview of technologies and techniques.

2.1.3. Approaches for Indoor Positioning

From the writing on indoor localization an order can be made: from one viewpoint, the amounts that are estimated (the technology), and then again, the way in which these amounts are utilized to figure a particular area (the techniques). Table 2.1 depicts a diagram of this order. Fingerprinting, signposting and trilateration etc. are the utmost significant techniques. Received Signal Strength (RSS), Time of Arrival (TOA) and Angle of Arrival (AOA) are few technologies that are utilized for the purpose of indoor localization. The conceivable localization methods cannot just contrast in technologies which is utilized, yet additionally in manner by that the information is processed. In this way, alternative arrangement is made among central and distributed techniques. Localization is centrally measured or scattered on every cell phone individually [26, 27]. For indoor localization, the operation of the most significant technologies and techniques is described in the following section.

2.1.3.1. RSS

It takes extent of power, existing in expected radio signal. And that power depends upon the distance among the TX and RX. A huge no. of test estimations of the gotten signal are required to this received power determination. During disjoint time interval, sampling is taken place. Unit of RSS is dBm of the measured power. The advantage of RSS is that it does not consume extra bandwidth. When using RSS it does not matter whether the distance is measured at the versatile station or the base station. Building structure highly influences on like dividers, entryways, floors, solid, steel, furniture, machines, and so forth and additionally reflections, multi-way by these structures [28].

2.1.3.2. TOA

Localization performed by TOA innovation depends upon time taken by the signal to spread from versatile station to base station. For this purpose, it utilizes the limited speed of spread, approx. the speed of light (300 meters per microsecond) 1 as shown in Fig. 2.1. Actually, this time is directly related to the distance traveled by the signal. But the flaw is that there must be a clock with very fine resolution. Second is that, to achieve an accurate distance measurement it requires complete synchronization. For example, a very small error of time measurement say 100 nanoseconds leads to the localization error of 30 meters shown in figure 2.1. Now a days UWB technology is being used in cost effective and accurate way.

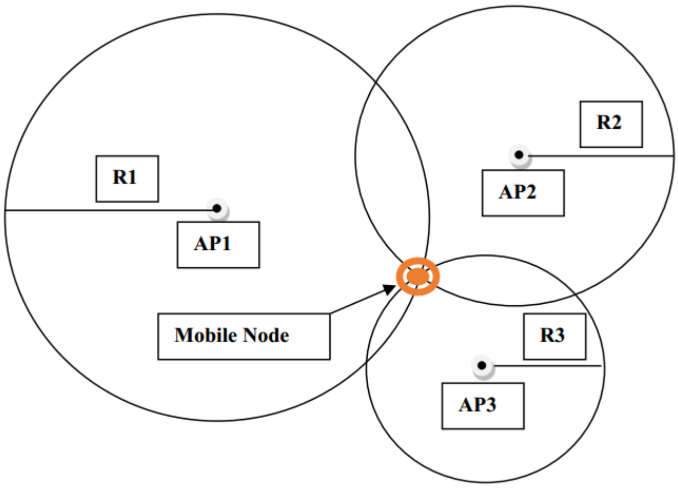


Figure 2.1: Location sensing Time of Arrival (ToA) [29].

2.1.3.3. Time Difference of Arrival (TDOA)

This is simply a variant of TOA. TDOA actually skips the synchronization amongst receiver and transmitter. At receiving node, localization is performed with the help of difference in time of arrival. Therefore, there must be at least three nodes for basic operation (transmitter node and minimum two receiver nodes). Because at the receiving nodes, the difference in arrival time is actually independent of the time base of the transmitter hence there is no need to be synchronization between transmitter and receiver. But there is always need of synchronization between all receiving nodes [28]. Figure 2.2 gives the location sensing strategy of TDOA.

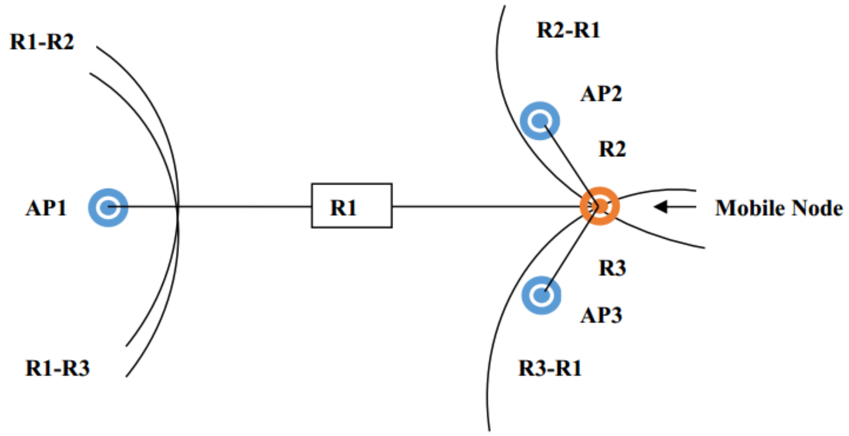


Figure 2.2: Location sensing Time Difference of Arrival (TDoA) [28].

From the figure it is clear that line of sight is compulsory between transmitter and receiver for the distance measurement. Otherwise it will choose most brief landing time which will specifically identify with mistaken calculations of the distance. These types of reflections clearly additionally impact the exactness of RSSI based strategies.

2.1.3.4. AOA

Technology could be utilized for positioning in a hub network where all hubs are equipped with ordinance reception antennas. It is possible by making use of the reception angle of the signal. AOA with the help of reception angle of the signal, AOA technology could be castoff for localization in a node network where nodes are equipped with ordinance reception antennas.

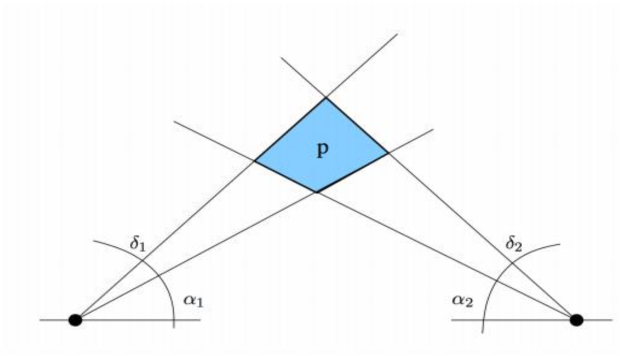


Figure 2.3: AOA visualization [26]

In the given figure 2.3, it is shown that two reference hubs are used to decide an unequivocal area territory (p) at the accepting zone. For smaller angles that are distinguished by the receiving antenna, the area zone (p) will be lesser. It is conceivable to utilize AOA in blend with RSS or TOA to do genuinely correct localization. Remember that this technology additionally gives improper outcomes if there is no LOS between the hubs [29].

2.1.4. Techniques

Number of techniques are used to perform localization. Most of them are discussed here.

2.1.4.1. Fingerprinting

A standout amongst the most widely recognized indoor positioning methods is fingerprinting [6, 17]. This procedure makes utilization of a database of signal values of every hub in the system (generally RSS esteems) estimated on a few particular areas. Accordingly, it is conceivable to construct a radio map of the building for every hub's estimations [30, 31].

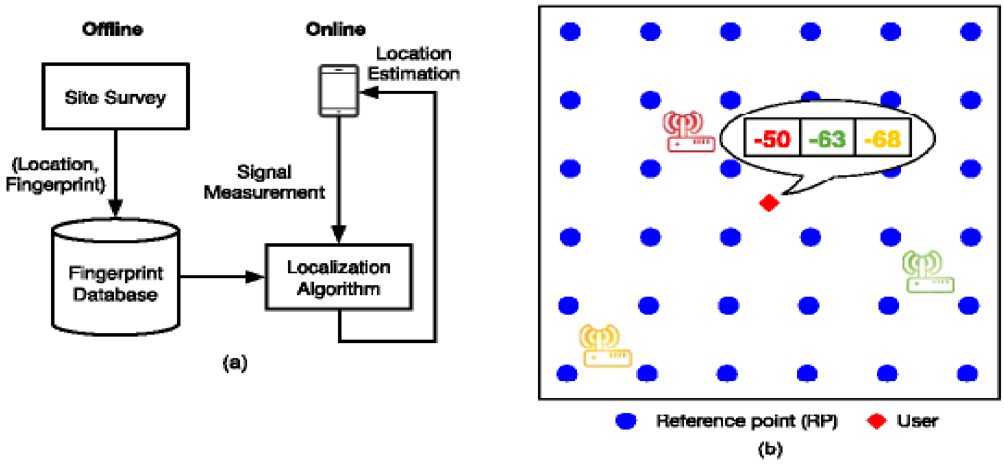


Figure 2.4: Fingerprinting workflow of training and online phase [30]

To decide the location of the blind hub in a building, one contrasts the gotten signal and vector saved in radio map. The best match (along with all reference hubs) gives the location. From the best comparability (with all reference hubs), the situation of the blind hub can be figured.

The technique comprises of two stages, a preparation stage (training segment) and an online stage (or arranging segment). Figure 2.4 demonstrates these two stages for Wi-Fi fingerprinting. The region is checked on all sides of Access Points (AP) from the Mobile Unit (MU) [32, 33] in the training stage. Then in the database, the fingerprints from each AP are saved together with its (x,y)- position. This stage utilizes the supposed war-driving technique. In online stage a localization algorithm is utilized which matches online query of the portable unit to fingerprints saved in radio map. And where most ideal match is done, this is the portable units (x,y)- position.

Size of radio map utilized is dependent upon the accuracy of the localization and the quantity of reference points. The primary drawback of the system is the preparation stage. Building a radio map for every hub in the building is extremely labor-intensive and tedious for huge spaces [34]. In addition, these radio maps are a preview of signal power instant of estimation affecting as erroneous measurements, at time signal powers modify or one of the hubs wind up defective.

With radio map based on fingerprints, the indoor positioning is done centrally, all information is handled by the side of a focal area (e.g. remote server with database), in opposition to distributed localization.

2.1.4.2. Signpost Positioning

Signpost localization [35] is the easiest indoor localization procedure, as well as it is least accurate. In other words, it is very simplified type of fingerprinting. The correct position of every access point should be identified before utilizing the signpost procedure. Typically, a characteristic term for access points is used, for example label of a particular room or work area. Position of the blind hub is connected with the characteristic label of reference hub that is best received among all. Thus, it is possible to forecast, in which room the blind hub is located by using signpost algorithm.

There is no requirement for a number of calculations, is the main benefit of this procedure. While low accuracy of the localization (generally room area estimation) is its drawback. Expansion of reference points is directly related to how much accuracy we need to obtain. This additional framework will prompt extra expenses to implement the signpost localization strategy.

By utilizing the signpost localization, positioning is performed centrally or distributed. Just like fingerprinting a focal database is also used for signpost positioning. It is conceivable for the blind hub to decide its own position on the basis of the constrained calculated power required.

2.1.4.3. Angulation

The angulation is like lateration yet it utilizes the angles in estimating the position of the object [36]. The two-dimensional angulation needs one length and two angle calculations, i.e., the reference point distance as appeared in figure 2.5. In 3 axis angulations, one azimuth, one length and two angle calculations are associated with exact localization.

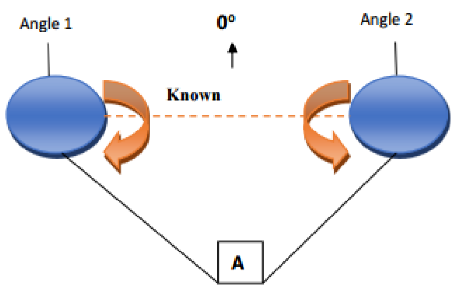


Figure 2.5: 2D position using angulation method requires one distance and two angle measurements. [36]

2.1.4.4. Triangulation

The geometric properties of triangles include in the triangulation technique to estimate location of the entity. This method includes the lateration that uses a few distance calculations among the known points, or by utilizing the angulation technique which processes angle in respect to points with known separation [37].

2.1.4.5. Lateration

The position of the target is estimated by lateration [36] in which distance measured from numerous reference points is used to locate the target. In two- dimensions, the position of object is assessed through 3 non-collinear points by utilizing the distance estimations, as it is appeared in figure 2.6. In three dimensions, the location of an entity needs the distance estimations from 4 non-coplanar points.

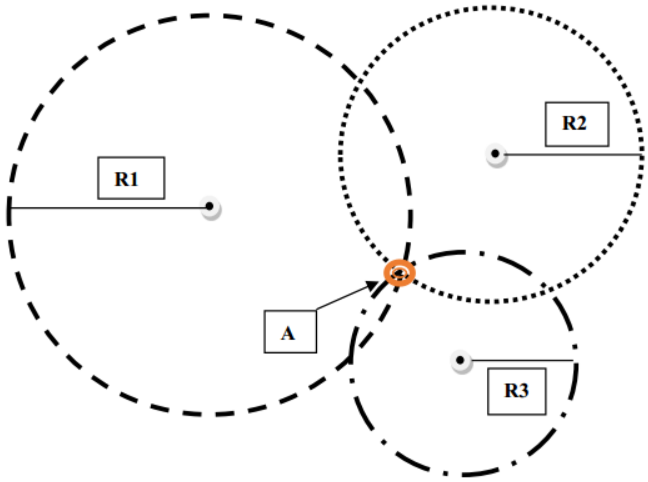


Figure 2.6: 'Determination of 2D position using lateration. It requires distance measurements between the object `A' and 3 non-collinear points. [36]

The measurements of the distance can be calculated by using three techniques, which are as follows.

a) Direct A physical act or movement is used for measurement of distance.

b) Time-of-Flight Estimating distance from the target to nearly point P by means of time-of-flight implies estimating time it revenues to move among target and P point at a celebrated velocity.

c) Attenuation With increase of distance from radiation source, the strength of released signal drops. Decrease correspond to genuine strength is known as offsetting. For example, the radio signal emitted from the target is attenuate by a factor that is directly related to 1 / r when it grasps P point at distance r from the target. Time-of-flight is more accurate than attenuation.

2.1.4.6. Trilateration

Utilization of trilateration is a standout amongst the most conventional methods for indoor positioning. The calculation-based technologies for trilateration are: RSSI, TOA and AOA. Calculations of distances are utilizing by trilateration for estimating the positions with the help of circles, spheres or triangles. Trilateration does not utilize angle measurements rather than triangulation.

In a two-dimensional plane, by making use of signal strength of at least two reference hubs which changed over to distances, the location of a blind hub can be resolved. It is identified that blind hub at that point lies upon dual circles that have their radii equivalent to these distances. As center of these circles shape a triangle organized with blind hub, the position can be estimated.

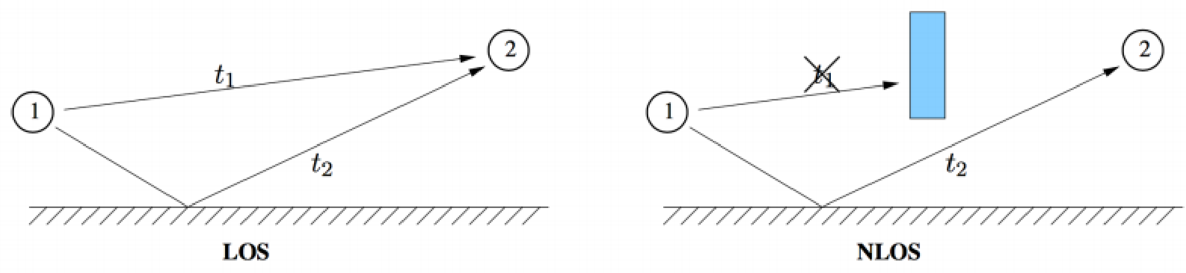


Figure 2.7: LOS/NLOS drawback of TOA [26]

In a three-dimensional plane, somewhere around three reference hubs are required. The location of the blind hub can be estimated by means of conventional geometric techniques. The center area of the three circles or sphere together with their radii is adequate data to decide a positioning area (likewise appeared in figure 2.6). Since these measurements not often give an unambiguous position, the most possible location is often achieved by means of interpolation methods, for example, the least squares method [38].

The Global Positioning System depends on this procedure, with the exception of that GPS is in three dimensions. So, spheres are required rather than circles to evaluate the distances from the position to the satellites. After the geometric triangulation with the three spheres two points of conceivable position are resolved [39]. One of the focuses isn't on the surface of the earth, so it very well may be wiped out, finding the current position. So, on the off chance that it is needed to utilize three-dimensional trilateration or multilateration for indoor localization, one need somewhere around four reference hubs [40, 41].

2.1.4.7. Proximity

A method estimates the closeness of an identified set of points knows as proximity [42, 43, 44]. There are three strategies for the most part used to calculate the proximity, which are as per the following [36].

a) Detection of the physical contact: The technologies such as capacitive field detectors, pressure, and touch sensors are used for the detection of the physical contact.

b) Measurement of cellular access point: The monitoring in the proximity location technique is used to monitor or locate the device when it moves from one access point, in a cellular network, to another.

c) Automatic Identification Systems The third approach of the proximity technique is the use of Automatic Identification System (AIS) e.g., point of sale (POS) terminals, landline telephone records, computer logs, credit card and electronic card-lock logs.

2.1.4.8. Vision Analysis

The vision analysis technique assesses the position from the pictures got at one or more points. To shelter entire space and capture real time metaphors one or numerous cameras are settled in the following region of an IPS. The followed targets are recognized from the images. The experiential metaphors of objectives are observed in pre-estimated fingerprints to perform location assessments. Moreover, by making use of the captured images, vision locating method can give beneficial positioning context for services [36].

Comparison

Table 2.2 summarizes a comparison between these technologies and techniques. As shown in table, it is obvious that fingerprinting is utmost reasonable strategy for the suggested indoor localization framework. Note that when using the other methods one needs to specify the location of the antennas. This is a major drawback in comparison to fingerprinting.

Method	Indoor Accuracy	Coverage	LOS or NLOS	Affected by Multipath	Cost
RSS	Low	Good	Both	No	Low
AOA	Medium	Good	LOS	Yes	High
TOA	High	Good	LOS	Yes	High
Fingerprinting	High	Good	Both	No	Low

Table 2.2: Comparison of indoor position methods. [39]

2.1.5. Matching Algorithms

After accomplishment of offline phase, where a complete site survey is conducted and signal strength values from each access points at all reference points are saved in the radio map, we talk about techniques involved in online stage. The position of an object is assessed from the data of radio map using different algorithms. The two most common algorithms or matching techniques are nearest neighbor (deterministic) and Bayesian (probabilistic). A simple difference between deterministic and probabilistic techniques is in the prior method the RSS values are represented as single values and the later uses probability distributions for representation. The probabilistic methods contain more information about the signal strength range but are much complex, whereas the deterministic methods are simple to implement and process [45, 46].

2.1.5.1. Probabilistic (Bayesian)

The probabilistic method based on Bayes Theorem, the signal strength vector collected in the online phase say $s = (s1, ..., sk)$, an d the set of locations is $X = \{ x1, x2, \ldots, xm\}$, we have to find the location in radio map that maximizes the conditional probability is shown in Eq. 2.1:

$$p(xi / s) \tag{2.1}$$

Mathematically it can be presented as:

$$Arg \max x[p(x / s)] \tag{2.2}$$

By Bayes theorem, we know that

$$P(x / s) = P(s / x) p(x) / P(s) \tag{2.3}$$

From above Eq. 2.3, the location with maximum probability of occurrence of the received signal strength vector will be the estimated location [47].

The Horus system [48] is also a fingerprinting-based WLAN system for location estimation. Unlike RADAR, it is a probabilistic system. The radio map saves the histogram of the received signal strength samples from each AP. Its location estimation phase is combination of two techniques i.e. discrete space estimator and continuous space estimator.

Deterministic In indoor localization systems, the existing work can be categorized in to two classes i.e. 1- particular hardware founded, 2- present structure founded. Like infrared [49], ultrasound [50], or RFID [51], it is always needed for some external deployment of extra location beacons for specialized equipment-based approaches. The main advantage of such approaches is that we can achieve desired localization accuracy by deployment of location beacons at necessary density and optimal positions. High cost for deployment i.e. specialized equipment at high density, is required for these approaches which are the major drawback. By now present radio signals in air are used in existing framework-based techniques for position estimation like fingerprints. In this case the Wi-Fi [8, 52], GSM [53], FM [54, 55], acoustic [56], and geomagnetic field [57] are the most commonly use wireless signals.

In Internet of Things (IoT) context, the most important task is to provide accurate location information [58]. Location based services are rapidly growing and huge market in the scenario of IOT. In recent years, indoor localization has received become provision of location-based services [59, 60, 61]. Nowadays, Wi-Fi is everywhere covering all space like hospitals, supermarkets, airports etc., it is very economical, simple and low equipment dependent, to employ Wi-Fi for indoor localization [62, 63, 64].

2.1.5.2. kNN

In few indoor environments, due to weak satellite signals, received signals strength indicator (RSSI) fingerprinting [65] based indoor positioning systems are deliberated using Wi-Fi infrastructure. Bahl et al. [66] offered a Wi-Fi fingerprinting based system RADAR which uses KNN algorithm to locate the wireless device. In RADAR, the matric is taken, was Euclidean distance. The localization accuracy of this method was dependent upon the selected nearest neighbors. Then cosine similarity was used as the metric of new proposed algorithm i.e. Cos-kNN [67]. In localization, the frameworks, using kNN gave the error from 2 to 5m (meters).

In online phase, the RSS value/online query is received from user at unknown position. Then this RSS value is matched with the all RSS values present in the database which is fabricated in offline stage. It is very essential step to select a pattern matching algorithm for indoor positioning. Among all matching algorithms, kNN is the most commonly used algorithm for positioning. Because it only uses the minimum distance to locate the user at known location with the help of k fingerprints [68]. Location is estimated using the RPs regarding to these k fingerprints. In pattern matching algorithm, the metrics are very important to find the distance among assessed RSSI value and one of databases. It is very important to obtain similarity among two entities.

Among different signal vectors, distance is computed by feature scaling model. This FS model is used in proposed feature scaling-based k-nearest neighbor (FS-kNN) algorithm [69] which is one of the approved techniques. In discrete way, by fixing the feature scaling weights in identified intervals, the distance measuring is improved from classical kNN using this method.

2.1.5.3. FS KNN

Assigning various weights to signal variances at various RSS levels is the key thought behind FS-kNN (or called powerful signal distance) while evaluating the likeness among dual RSS values. For similarity comparison, it needs to construct a RSS-value based FS model. In offline phase, for this model to work legitimately, to actuate the RSS-level-based scaling weights, it must initially be calibrated databases, on the basis of which simulated annealing (SA) is utilized which are particularly required in proposed model. Whereas By making use of this model, in online phase, among an instantaneous RSS vector (detailed by a MS) and every reference fingerprint in database, the computation of effective signal distance is done and afterward restoration of the first k RPs causing least active signal distances is completed, and the same distances are then utilized for the area estimation. At long last, in real office condition on college campus, FS-kNN is implemented, where sensible calculations are performed. An average position error up to 1.70m at most, the test consequences indicate that FS-kNN can accomplish, which is better than existing work [59].

As mention in [59] the scaling weight assigned to distinguished intervals are very different from each other. In that case if one of the recently described RSS vector falls on joint of two contiguous breaks then question arises: which weight would be selected from both intervals for the newly reported RSS value? This issue is termed as “ambiguous boundary” problem. And the distance will be called as ambiguous as it is calculated from ambiguous boundary. Therefore, due to this “ambiguous boundary” problem wrong localization can take place which will result as low accuracy.

2.1.5.4. Cos-kNN

In single space Euclidean distance is being replaced with cosine similarity. Cosine similarity function is given below where: A = (x_1 x_2 x_3 …. x_n) , B = (y_1 y_2 y_3…. y_n). These are two vectors for cosine similarity function, so this function is explained as:

$$\cos\theta = \frac{\mathbf{A}.\mathbf{B}}{|\mathbf{A}|.|\mathbf{B}|} = \frac{\sum_{i=1}^{n}(x_i.y_i)}{\sqrt{\sum_{i=1}^{n}x^2\sum_{i=1}^{n}y^2}} \qquad 2.4$$

For vectors A and B, θ is the angle. Cosine range is between 1 and -1. In Euclidean space, it calculates the angle between two vectors and after calculation the difference and assign the impotence to this difference rather than position is shown in figure 2.8. Let's suppose if position of one vector let's say A is fixed and other point B from the actual position is away from coordinates of the origin. The major difference between ED and cosine similarity is that, distance between two points or two vectors changed but the angle between them is not changed so cosine similarity is also not changed. Among the dimensions of vectors absolute difference between numerical calculations is referred to as ED. Whereas, cosine similarity is the measurement of difference between directions of two vectors rather than their actual numerical difference between RSSI values [62].

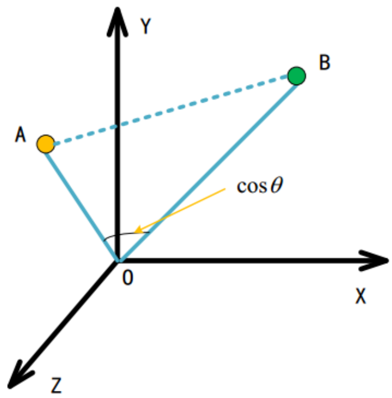


Fig.2.8: Cosine similarity between two RSS vectors in the signal space.

Change in RSSI values occur due to the diversity of mobile device or user and it is noticed as to subtract or to add the random noise in RSSI values. By changing a small value in RSSI calculation or by some minor change in RSSI value there exist a major change in Euclidean distance, but cosine similarity is almost same or there is very minor change in it. So, cosine similarity reduces the localization error to a large extent by matching the online RSSI calculation to offline database. So, localization system perform well, stable and accurate. In signal space angle is more important rather than absolute values of RSS vectors. So, to adopt the cosine similarity in signal space is more important in WLAN based indoor localization system.

2.1.6. Propagation Model

Estimating the signal propagation, some models [73-82] are described in this section. Each model is unique in its complexity level, but all are feasible to estimate the propagation in indoor environments.

2.1.6.1. ITU & Log-distance

In non-deterministic propagation model for indoor environments, ITU is recommended to estimate the total path loss within a structure i.e. room, hall or corridor. Having frequency array from 900 MHz to 5.2 GHz and up to 1 to 3 floors, the ITU model is easily applicable and can be expressed in Eq. 2.5 [73].

$$L = 20\log f + N\log d + p_f(q) + 28 \tag{2.5}$$

Where L is in dB called as path loss, frequency is denoted by f with unit MHz, space power loss factor is N, among TX and RX, q is number of floors and lastly the floor loss penetration factor is shown as $P_f(q)$, that remain 0 for the same floor measurements. with the exception of shadowing, ITU model seems to be like the log-distance path loss, by an additional term which is used to consider the shadowing effect, later tries to account for clutter among TX and RX.

To estimate the path attenuation (L) in dB, the model in Eq. 2.6, is generally utilized in the link budget enquiry.

$$L = L_o + 10n\log(d / d_o) + X_\sigma \tag{2.6}$$

1 m of a reference distance (d_o), the L_o is the free-space loss in (dB), the experimentally derived path loss exponent is shown as n and in meters, d is the space among the TX and RX in a straight line. With a mean of zero, X_σ is a Gaussian random variable (in dB) and standard deviation of σ dB.

2.1.6.2. Motley-Kennan Multi-wall Model

COST 231 and Motley-Kennan are generally comparative multiwall models. Notwithstanding the separation between handsets, they additionally require quantity of the panels among handsets. By allocating a firm attenuation factor to distinctly wall, the models represent extra offsetting caused building structure. This attenuation is caused by signal intruding walls to come to next handset. COST 231 additionally represents floor attenuation however to abstain from overvaluing, it moves off transition losses afterward a specific no. of floors is encroached [74]. Motley Keenan model and COST 231 (without the floor loss) are shown in Eq. 2.7.

Motely keenan model and COST 231 propagation models are similar relatively. Between transceivers, they not only find the distance but also need the total number of walls present between transceivers. The models calculate the extra attenuation caused by the structure of the building by allocating specific attenuation factor to each wall. In the non-LOS environment, signal intruding the wall for reaching the transceiver is caused attenuation.

After particular number of floors, a transition loss is impinged which is accounted by COST 231 model but it avoid overestimating [75]. Both models are shown mathematically in Eq. 2.7:

$$L = L_o + 20\log(d) + \sum k_j a_j \tag{2.7}$$

Where k_j is the no. of walls of type j between the TX and RX and for same type of walls, the attenuation factor is a_j.

2.1.6.3. Ray Tracing

In indoor environment, the estimation of propagation of radio signal is done by more accurate site-specific model, Ray Tracing. However, it is cultured as well as more demanding than previous mentioned models by having the capability of channel response prediction. At a given point, the channel impulse response h(t), in 3-dimensional ray tracing, is expressed in Eq. 2.8:

$$h(t) = \sum_{i=1}^{l}\left[\left(\prod_{u=1}^{M_{ri}} \Gamma_{iu}(\theta_{iu}) * \prod_{v=1}^{M_{pi}} P_{iv}(\theta_{iv})\right)\frac{1}{r_i}.e^{-j\omega\tau i}.\delta(t-\tau_i)\right] \tag{2.8}$$

Where number of rays are expressed as i which is reaching a specific point, number of reflected rays is shown as M_{ri} and for wall u, the reflection coefficient of ray i is expressed as Γiu, Likewise, for wall v, the transmission coefficient of ray i is shown as P_{iv}, and θi is the angle of incident of the ray i. With ri as a path length of ray i, with time delay of Ti, e^{-iwTi} represents the carrier [76].

2.2. Weaknesses of Wi-Fi for indoor positioning

Utilizing Wi-Fi signals for indoor localization is a standout amongst the most proper and ideal arrangements as a result of the presence of IEEE 802.11 b/g/n access points in buildings. Other than this, Wi-Fi additionally guarantees some negative impacts. Utilizing Wi-Fi for indoor localization is thusly not without downside.

2.2.1. Body Effect

European radio controls are institutionalized by the European Conference of Postal and Telecommunications Administrations (CEPT). The European Radio communications Office (ECO) creates controls for CEPT and obliges clients of the 2.4 GHz frequency band to work at low power. Subsequently, gadgets work at a most extreme power of 1 watt or 30 dBm [84]. As per [85], signal quality may drop 10-15 dBm because of the low penetration power. Note that this power misfortune represents 14-21% of the aggregate effective signal strength. Due to this power loss, the signal will highly impact the positioning system. In any case, there can also be other individuals and objects present and moving, in this way changing the power transmitted among sender and receiver. This is a noteworthy disadvantage for Wi-Fi based indoor localization systems, particularly in swarmed conditions.

2.2.2. RSS variations

Not just the body impacts the Wi-Fi signal, yet in addition differences in the environment e.g. physical entities add to the RSS variations in time. Because of walls and different configurations, which are in the environment of the proposed localization system, or by changes in the surroundings, multipath propagation typically conquers causing huge variances in the RSS. This impact, called fading, will be an disadvantage for the IPS when utilizing a cell phone to note the RSS [86]. This issue is most generally improved by calculating the normal RSS value over specific time duration as well as this must be borne at the top of the priority list in the proposed design for this thesis.

Another issue causing variations in RSS values is interference. The frequency band utilized by Wi-Fi radio signals is normally shared by different systems or gadgets e.g. microwave, Bluetooth gadgets. Interference may diminish RSS significantly when these gadgets are close-by and in process. RSS is in this manner enormously influenced by the body impact, fading, and interference. Strong RSS values would be to a great extent influenced by fading, while weak RSS values might be affected by at least one of the three elements.

2.2.3. Influence of the number of APs and RPs

The utilization of the quantity of access points isn't straightforwardly identified with the property of the Wi-Fi signal yet it is surely vital to say this as it can also be a drawback of utilizing Wi-Fi. The quantity of access points in total location, results in inaccurate fingerprint record, prompting poor performance. Putting extra APs to improve the database may cause extra expenses and placing the new APs can be tedious. [87] Observes at the connection between the quantity of APs and reference points (RPs) in the database that outcomes into optimal positioning results. An indoor space with measurements of 11 x 23 m, must contain around 5 APs and around 66 RPs for ideal confinement [88]. Truth be told, expanding the number of APs or RPs scarcely impacts the results [89].

2.2.4. Signal Aliasing

Another shortcoming that happens with Wi-Fi signals is a phenomenon, called signal aliasing. As indicated by [63], signal aliasing indicates: In signal space two points that are far separated might be close together. On account of the complex indoor propagation environment such aliasing can happen. The signal strength at a point near to an AP might be like that on the other point which is distant away essentially on account of a hindrance, (for example, a wall) attenuating the signal received at the previous point while the last point receives an unbarred signal. To take care of this issue, a first arrangement is the optimal deployment of APs in the building.

Suitable access point deployment in the building format is along these lines extremely fundamental in solving this issue. This phenomenon must also be considered in the proposed design. The positioning algorithm should implement a solution for this. The RADAR framework by [34] tackles this by continuous client tracing. If the system can estimate the position in a prior stage, it can choose between some calculated positions to estimate the actual location of the target.

2.3. Performance Metrics of IPS

IPSs utilize various localization techniques that differ incredibly as far as precision, cost, accuracy, innovation, scalability, robustness, and security. A few applications can involve low cost IPS whereas others might need high accuracy IPS, for example, medicinal tracing, industrial environmental tracking and indoor positioning system for blind. There are depicted diverse performance metrics of IPSs in this area.

2.3.1. Accuracy

As the similarity of agreement among a calculated value and an actual value of a deliberate [12] the term accuracy is characterized in the Joint Committee for Guides in Metrology (JCGM). Among measured location and the actual location, in this way, accuracy of IPS is average Euclidean distance. For some analysts in the field, the accuracy is as yet an extremely difficult region. a few compromises may be required among accuracy and other performance metrics [12], despite the fact that for most applications, the accuracy of an IPS is important driver.

2.3.2. Availability

With the needed accuracy and integrity, availability is the proportion of time that positioning service is accessible for use. As IPS Integrity is the certainty that’s why it can be set in the yield of the IPS. Just like communications congestion and scheduled factors, the availability could be constrained by irregular factors, for example, routine maintenance. Availability is viewed in three levels mostly i.e. less availability (< 95%), consistent availability (> 99%), and higher availability (> 99 . 9%) [62].

2.3.3. Coverage Area

The area which is covered by IPS is called coverage area. There is always a different range in which Every IPS works. The ones which shelter the extensive collection [62] are the most effective systems. Normally, there can be three levels of inclusion; local, scalable, and global for positioning systems. A well-defined, finite area is referred as local coverage that cannot extendible like a single room or building, whereas a scalable coverage is taken as a capacity of a system to expand the space by including equipment. Then again, a system which has worldwide area is referred as global coverage, for example, GPS. In present days, ranges from 5 meters to 50 meters of existing IPSs are going. In this manner, giving a framework which has coverage of in excess of 60 meters is challenging [62].

2.3.4. Scalability

To find the location of the objects around the world, a positioning system might have the capacity, inside a metropolitan region, all through campus, in a specific building, or inside a solitary room. Additionally, quantity of targets system can trace with a specific number of infrastructures or above given time might be restricted [37]. When scalability scales in one of the two measurements: geology and number of clients, then the scalability of an IPS implies system guarantees general positioning function. The quantity of clients' scale implies that no. of units found per geographic zone per time period rises.

2.2.5. Cost

At various measurements the cost of the IPS is estimated and these measurements are money, time, space, and energy. It is caused at various stages of framework: framework installation and maintenance, infrastructure components, and locating gadgets [20]. For framework establishment and maintenance, the cost incorporates the cost which is needed for establishment, and any costs compulsory to save the framework functional, whereas for purchasing components and preparing them, the cost for infrastructure components and locating gadgets might incorporate expense, space, and energy to use equipment. The IPSs which reuse the current infrastructures, for example, the network, are more cost effective. Just like passive RFID labels, some positioning devices, are totally energy submissive whereas others expend extra energy. This energy can be considered like a basic asset in IPSs to keep away from facility disturbance and deliver greater portability arrangements.

2.3.6. Privacy

For people utilizing IPSs, protection is essential, where a strong access command over how client's individual data is gathered and utilized is critical [9]. The Goal is to enhance client's privacy, with a specific end, confidential tools must be applied and conserved to shield information from interruption, robbery, and misappropriation. Tragically, in the majority of the assumed research in field of indoor positioning, the safety part of IPSs is not a noteworthy concern until now [36].

Chapter 3 - System Model

In this part of the thesis, complete overview of indoor positioning using multiwall model is given. In offline phase, construction of fingerprint-based database and in online phase, proposed algorithms for positioning, is discussed. Block diagrams are given and working of each step corresponding to chart is also discussed for deep understanding.

3.1. Introduction

RSSI localization techniques are based on measuring signal strength from a user to numerous access points located at different places and then to determine the distance among the user and the access points, this information is combined with a propagation model. To calculate the assessed user position according to predetermined position of APs, trilateration (sometimes called multilateration) techniques can be utilized.

Traditional fingerprinting is also RSSI-based, but from several access points in range, it simply relies on the recording of the signal strength, and this information is stored in a database with the known coordinates of the user device in an offline phase. However, this data can be deterministic or probabilistic. The current RSSI vector by an unknown location is then compared to those stored database as fingerprint, it happens during the online tracking phase, additionally the nearest match is served as the estimated user location.

A global overview of the proposed system architecture is shown in Figure 3.1. Obtaining the fingerprints is done through an Android application. Matching the target fingerprint with the database of collected fingerprints is done remotely on a webserver written in MATLAB. MATLAB is a functional language with a declarative style of function declaration. This allows the program implementation to be done in a short time. Tuples are used to represented data structures and lists are very efficient to manipulate data in MATLAB. Based on a best match algorithm, one tries to find the device’s location by sending a target fingerprint to the server. In the offline phase the database is built. The online phase is designed for localization.

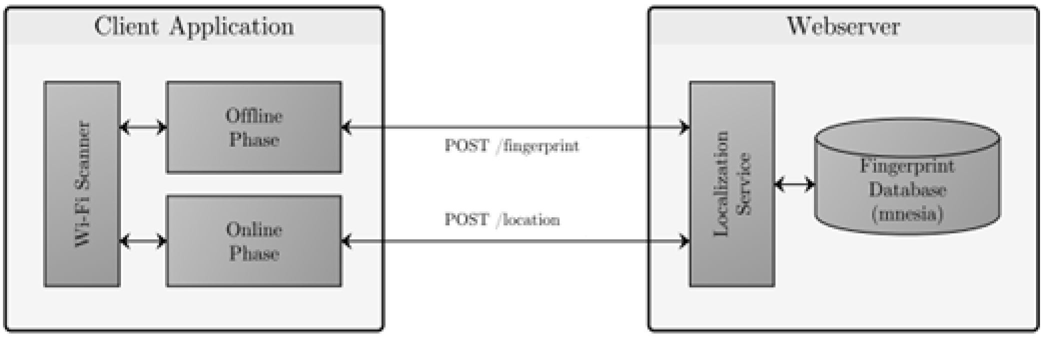


Figure 3.1: Overview of system Architecture

3.2. Proposed Algorithm:

As discussed earlier, indoor positioning is based on two phases, offline phase and online phase. In offline phase radio map is created based on RSSI values of all RPs from each AP and in online phase the online users are located by matching their online RSSI values with the RSSI values stored in the radio map. The matching of offline and online data is done with the help of matching algorithms. In this research the focus is mainly on two objectives. 1- Create strong databases by using appropriate propagation path loss model. Which is explained in next section and would be done in offline phase. 2- To develop the more accurate matching algorithm to locate the user. It would be done in online phase, by adding an extra feature to already existing model i.e. Feature-Scaling kNN. Deleting outlier is a statistical technique which is used to delete the individuals from a set C based on standard deviation. Therefore, as the FSkNN model finds the k nearest neighbors to select the estimated position, deleting outlier process will be held before, to remove the most non-related nearest neighbors from the selected k set of neighbors. By doing so, the most likelihood species having more chances to be estimated position will be left in the k set. This will help the algorithm to locate the user in more accurate way. Experimental results have showed that this proposed feature is being helpful to improve the accuracy of the existing model.

At the same time, existing algorithm (kNN, FS-kNN) will be processed with new databases and more accurate results would be extracted from these models.

At the end, all three algorithms (2 existing, 1 proposed) will be compared with each other to find the best possible algorithm. Figure 3.2 and 3.4 are the block diagram that show the main steps and their order of working in both phases i.e. offline and online.

3.2.1. Offline Phase

In indoor positioning systems, offline phase is the first step to start process. As mentioned earlier, in this phase databases are created of an area in which a target is supposed to be located. Offline phase further consists of number of steps. The detail of these steps is given below:

3.2.1.1. Initialization

Figure 3.2 depicts the backend workflow when the user makes a fingerprint in the offline phase. As this research work is software based so it does not provide hardware implementation and other hardware applications. Functionality detail of each step is given in next section.

3.2.1.2. Data Construction

The entire explained Program is implemented in MATLAB. The simulation is started by obtaining a 2D blueprint image of the structure, for multi-wall models. Then the image is calibrated as PPI (pixel per inch), hence in the picture, all the walls/panels represent the genuine building coordinates. Then adjustment of reference points is done, it is also known as mesh points (RXs). This can be performed by giving to the algorithm, the length of a certain wall (in meters), and then picking up two end pixels which relate to that specific wall. Four Access points are selected and served as transmitters (TXs). Firstly, Euclidean distance is calculated among TXs and RXs. Then Between the TXs and RXs, an algorithm calculates the LOS distance. To identify the number of walls and walls type among the TX and RX, an algorithm, named as Bresenham’s line, is used, on account of the multi-wall propagation model, to acquire the crossing point of a conventional line among RX and TX. In blueprint image, this algorithm works with the cooperation of image processing techniques as well. The propagation model used in this era is “Motley Keenan multiwall propagation model”. The specialty of this model is that no 3D model is needed and there is no prerequisite to characterize walls in a vector shape radio map, therefore this methodology is considered to encourage the basic outline. However, as a DXF (Drawing Exchange Format) file, the 3D ray tracking algorithm, needs the building model. Therefore, with the help of the set of four corners in the Cartesian axis and a common vector, every finite panel is defined, which specifies the location of the walls. Additionally, points of interest are added to simulation to anticipate the received power, for example the handset constraints (efficiency, VSWR and diffusing power), beam's angle of parting (from TX) and angle of appearance (to the RX).

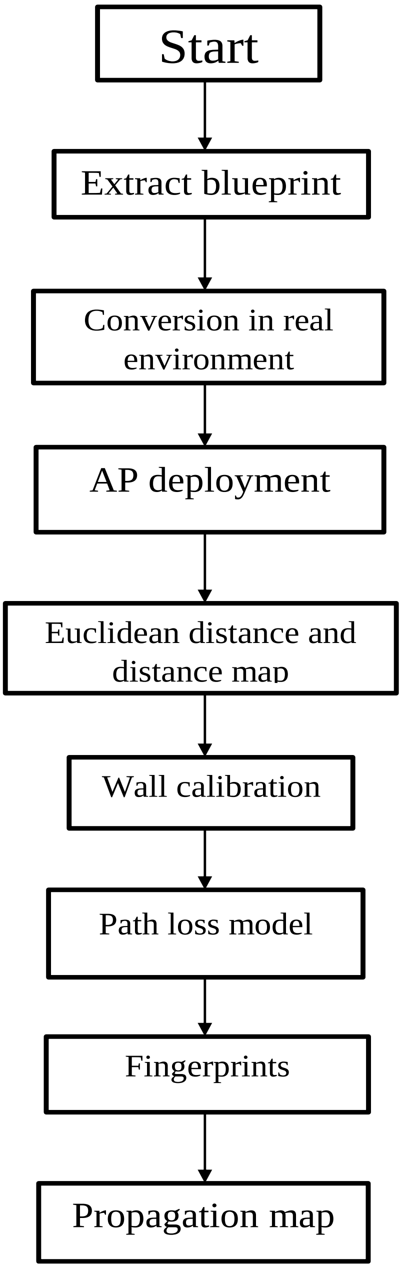


Figure.3.2: Block Diagram of offline phase.

Blueprint image is divided by an algorithm into a certain number of nodes and at each node, path loss is calculated by this algorithm with the help of proposed model. With 1-degree resolutions, Transmission coefficients are pre-calculated and then stored in a lookup table. The incident angle (θ_i) is restricted among 5° to 85° interval to avoid very severe LOS attenuations (when $\theta_i \approx 90°$) or complete transmission ($\theta_i \approx 0°$). Later a propagation estimation outline produced over the blueprint image by including other estimation data, (antennas gain and Tx power).

From the moment that the total scan time is reached, all the fingerprint data are retrieved. The value for every reference point in the dictionary is reduced to the average of the x maximum values observed during the scan, based on the algorithm using the average of a number (X) of selected maximum RSSI observations [36]. The received RSS values are sorted by descending order before taking the average of X RSS values from the reference points. Due to this the random poor sampled RSS values are removed before saving the fingerprint in the database. RSS vectors with the RSS values lower than X, are ignored while constructing the fingerprint, in the final fingerprint, the observed BSSID is not included. This factor helps to reduce the fingerprint size, saves storage in the database and also indirectly ensures faster response time from the server. The final fingerprint saved in the database consists of the location of the reference point and every obtained RSS value in dBm together from each access point.

3.2.1.3. Propagation Model

Wireless systems are a fundamental piece of our everyday life. They are coordinated into private, business, scholastic and even modern conditions because of their simple access, adaptability, cost productivity and portability. Consequently, wireless local area network (WLAN) which is effective and cost-productive is of great significance. Before deployment link budget analysis becomes complicated to guarantee that wireless services are accessible to portable clients inside a specific area. Because of their auxiliary unpredictability, installations, building materials and portability design, indoor environments always face more challenges. To relieve these tasks and encourage link budget design few diverse propagation models are subsequently used.

These models have their own pros and cons as they are intended for specific settings. For example, in any case better precision is achieved by deterministic models mostly, their implementation is complex though, a precise model of environment is required computationally. On opposite, there is not as mush computational demand of Semi-deterministic and semi-empirical models, they don't require a precise model of the scenario usually, because they are easy to execute and have satisfactory estimated accuracy [57 81]. The Motley-Keenan multi-wall model is one of the generally utilized models which has achieved a substantial consideration and has been enhanced [86 82, 87 83]. It is all because of its simplicity as well as its efficiency. In Eq. 3.1 the entire model is represented mathematically whereas L and L_o is the propagation loss and loss at reference distance of 1m respectively. Among TX and RxX, *d* is the distance whereas attenuation factor of the k^{th} wall is represented by a_k which is impeding the line of sight (LOS).

$$L = L_O + 20\log(d) + \sum a_k \qquad 3.1$$

The same model was extended in [88 84] to incorporate extra site-specific losses. To enhance its accuracy in [89 85] Lima et al additionally made a change in accordance with Motley-Keenan model. The thing which identify the thickness of the walls, by addition of an extra term to the model, this can be eliminated. The model is also adjusted in [90 86] to enhance the precision of the estimation by consolidating the breaking point phenomenon [91 87]. For indoor UWB propagation, the model is explored and altered in [92 88].

Authors deliberated two distinct situations of LOS and non-LOS and made a possibility i.e. by utilizing the LOS based path loss information to enhance the NLOS building loss derivation. And this idea is used in proposed model. In [93 89] the model is altered, among the accumulative diffusion loss and number of infiltrated panels to signify the nonlinear relationship. Authors particularly utilized that model like a choice to beam following because of its easiness. In [81] Authors utilized image processing methods to identify walls position and depth from a floor plan image rather than physically labeling the walls. In any case, impacts of angle of incidence and polarization are unnoticed.

Motley-Keenan model is therefore modified to accommodate the angle of incident, Eq. 3.1.1.

$$L = L_O + 20\log(d) + \sum \Gamma_k(\theta_i) \qquad 3.1.1$$

Instead of using a fixed attenuation factor, the transmission coefficient is used which is a function of incident angle and the polarization. This will adjust the signal attenuation level according to θ_i.

By considering the angle of incident and beam polarization, a modified model is developed which is known Motley-Keenan model. Without extra cost of implementation and complexity, accuracy of the model is increased.

a) Implementation of New Model

One of the major factor of this model is to calculate the angle of incident between rays of signal and walls when signal strike with the wall. In 2D and 3D this model is very straight forward because here walls are presented as vector equations. But it increases the cost of implementation and complexity. This model only need a blueprint image for its processing. By using mere an image, it is able to calculate the θ_i. Here Hough transmission is used to distinguish the walls from the image. Through Hough transmission space orientation or position of the walls θ_H can be known. Then by finding the position of the AP on image it calculates the angle of transmission (ϕ). So, after this angle for transmitted signal is found ($\theta_i = 90 - \theta_H + \phi$). When algorithm calculates the number of walls from blueprint image a gray scale image is presented which show the different walls having different intensity of pixels which are mostly starts from 254. It is done to separate different walls and then according to their intensity specific permittivity's ($\in_{r_k}$) and orientations (θ_{H_k}) is assigned. Then LOS is determined between AP and Rx by using a line algorithm Brenham. Intersected pixels of image with LOS are extracted. So, every intersection has some unique intensity levels, so it assigns different $\left(\theta_{H_k}, \in_{r_k}\right)$ values which calculates the θ_i and consequently Γ.

3.2.1.4. RSSI Propagation Map

A propagation/heatmap visually demonstrates the dispersion of a metric over areas inside the deployment region on a layout. The measurements to be seen in the heat maps join RSSI (the recipient signal strength indicator), the signal strength, the data rates, the throughput, the QoS Score and so on, which can be selected from the drop-down list in the heat map indicator. The heatmap contains a drop-down menu to allow the heatmaps to be seen by various factors.

It is possible to obtain metric either from single selected AP or all APs present on that floor. It is worth noting that readings of the metric can be different, which is mainly dependent upon the fact that whether the readings are taken from AP perspective or client perspective. Therefore, it is manually selected, to by AP or by user.

There is another fact about heat map is that if there is any change in in AP characteristics, user characteristics, their position, characteristics of walls, their positions and so on, will directly reflected by the heatmaps.

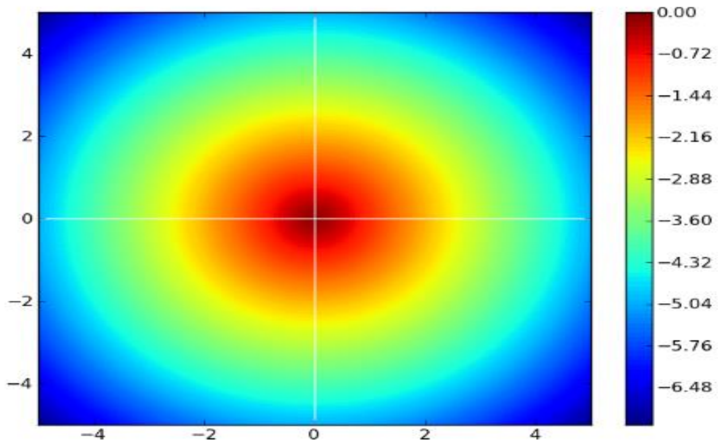


Figure 3.3: Heat map of RSSI vectors.

3.2.2 Online Phase

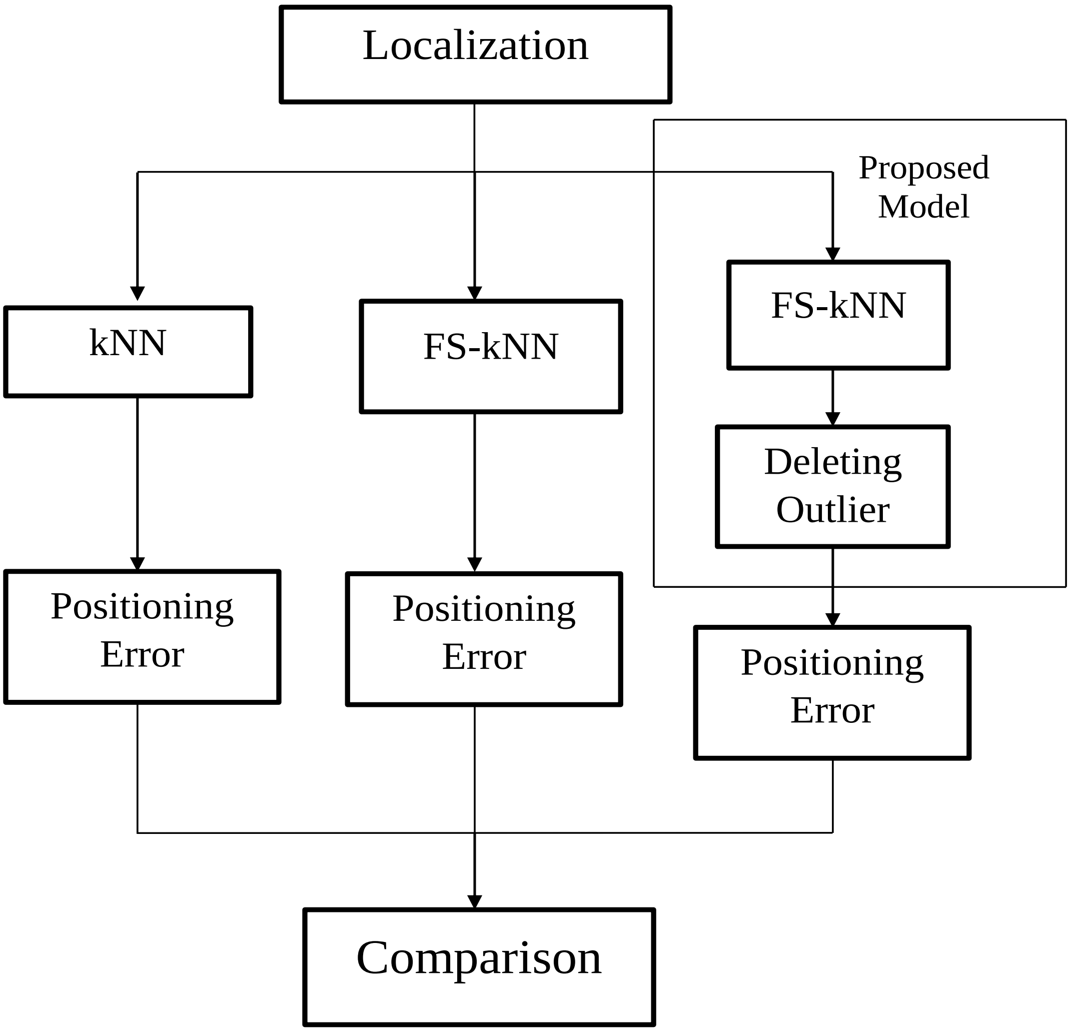


Figure 3.4: Block diagram of Online Phase.

3.2.3. Working Description

Much research also focuses on the localization algorithm itself used in the online phase. Most research implements a positioning algorithm that is either a deterministic or a probabilistic algorithm. These algorithms use real-time matching of RSS fingerprint data or the probability distribution of the signal strengths to obtain a user’s location respectively.

The most common techniques for deterministic positioning are data mining and machine learning algorithms such as KNN, Support Vector Machine (SVM) and artificial neural network. The KNN algorithm is an easy-to-use machine learning algorithm and is widely applied in indoor positioning systems because of its simplicity and high performance. The FS-KNN is a variant of KNN, this algorithm uses coefficient to calculate the position instead of just averaging the positions of the best matching fingerprints.

3.2.4. Localization Algorithms

As mentioned earlier, in this research work, matching algorithms are used to perform localization. Working detail of these algorithms is given below:

3.2.4.1. KNN

In the position decision phase, the kNN, pattern matching algorithms must be essential. k-nearest neighbor algorithm (kNN) is a generally utilized pattern matching algorithm that uses the recently announced RSSI vector gathered to discover the k fingerprints with least signal distances in the database using mobile device [59 90]. Thus, algorithm calculates the position based on the distance among the reference pattern and patterns present in the database. Various formulas can be used for distance calculation, e.g.: Manhattan distance or Euclidean Distance. It finds the K best matching fingerprints based on their mutual distance and calculates the position as the average of positions of same k patterns. After that, to find the expected position, the reference points relative to these k fingerprints are used. However, to find the distance among the assessed RSSI vector and one of the fingerprints, the vectors serve as the key component in the pattern matching algorithms. We can also say that, the resemblance between two objects to be matched is essential. The distance can be calculated by the formula given in Eq. 3.2.

$$d_m = \sqrt{\sum_{n=1}^{N}(RSSI_{m,n} - r_n(t)^2} \qquad 3.2$$

3.2.4.2. FS-kNN Localization

In this algorithm, RSS level-based scanning is introduced to measure a new scanning feature to measure the difference of effective signal during the corresponding synchronization among various signal vectors. d'm show an effective signal distance among online query and (fingerprint) attached to the mth RP, so that it can be calculated by:

$$d'_m = \sqrt{\sum_{l=1}^{L}(RSS_{m,l} - RSS_l)^2 * w(RSS_l)} \qquad 3.3$$

The value of effective RSS distance is calculated by newly estimated scaling weight function w(·). At the RSS level of RSSI, the one unit of RSS is changed according to the amount of distance. Because of introducing the RSS-level-based scaling feature, the relationship between the original signal distance and the geometric distance should be improved better, which can be achieved by effective signal distance which is supposed to be effective, as shown in Eq. 3.3.

In the FS model, this scaling weight fn. plays very important role. But it's difficult to provide a closed-form expression for w(·) in indoor radio environment which is quite complex. This model simply handle this situation by equally dividing the whole RSS plane into the $n(n \geq 1)$ breaks, after that it finds the one fix value for every interval as scaling weight with the help of specific calculations and tuning repeatedly. If there is n=1 in Eq. 3.4 then FS-kNN model degrades to simple kNN model. Thus RSS-value-to-scaling-weight plotting is defined as step function form, such as:

$$w(x) = \sum_{i=1}^{n} a_i X_i(x) \qquad 3.4$$

In above Eq. 3.4 real RSS value is denoted as x and *w(x)* as scaling weight at RSS vector (x) for actual signal difference. If *Ai* is the *i*th RSS break $1 \leq i \leq n$, for that *Ai* interval, *αi* represents the coefficient and *χi(x)* is sign function of the same break *Ai* then:

$$Xi(x) = \begin{cases} 1, x \in A_i \\ 0, x \notin A_i \end{cases}, \quad 1 \leq i \leq n \qquad 3.5$$

If a mobile station receives a value of RSS (x) falling in the interval *Ai* from the above Eq. 3.5. χi(x) equals to one while all other values i.e. χj(x) having position of $j \neq i$ equivalent to zero. After that the accomplished outcome w(x) from Eq. 3.5 (equivalent ai) is utilized to compute the active signal distance in Eq. 3.4 as a scaling weight for RSS value x. The method of tuning the values of scaling coefficient i.e. αi ∀i is deliberated later.

a) FS-kNN model FS-kNN model is mainly based on two stages: 1- offline stages 2- online stage. As discussed earlier, in offline phase, to calibrate a target environment a site survey is performed with the help of mobile station along with Wi-Fi interface. Particular number of RPs overlaid that area for collecting data at these RPs. Number of RSS values, at each RP, are collected repeatedly from different APs. A 2D matrix is built as database of the target environment.

This procedure discussed above is the same method used for traditional kNN algorithm. But FS-kNN also improves the kNN algorithm in two ways which are: 1- during offline period, FS-kNN not even builds database as well as tuning scaling weights for RSS-level-based FS model. 2- during online phase, in the calculation of active signal distances, it makes use of tuned scaling weights. Between online query reported by mobile user and database of each reference, the effective signal distances are calculated to find the kNN databases for positioning.

b) Simulated Annealing (SA) Simulated Annealing (SA) is a compelling and general type of optimization. Within the sight of vast quantities of local optima, it is valuable in finding worldwide optima. " Annealing " alludes to a similarity with thermodynamics, particularly with the way that metals cool and anneal. Rather than the energy of a substantial, simulated tempering uses the target capacity of an advancement issue.

Execution of SA is amazingly straightforward. The algorithm is essentially hill-climbing except as opposed to picking the best move, it picks an irregular move. On the off chance that the chosen move enhances the arrangement, at that point it is constantly acknowledged. Something else, the algorithm makes the move at any rate with some probability under 1. The probability diminishes exponentially with the " badness " of the move, which is the sum deltaE by which the arrangement is declined (i.e., vitality is expanded.)

Prob(accepting uphill move) ~ 1 - exp(deltaE / kT))

A parameter T is likewise used to identify this probability. It is closely resembling temperature in an annealing framework. At higher estimations of T, tough moves will probably happen. As T tends to zero, they turn out to be more unlikely, until the point when the algorithm carries on pretty much like slope climbing. In an ordinary SA optimization, T begins high and is step by step diminished by a "annealing schedule". The parameter k is some steady that relates temperature to energy (in nature it is Boltzmann's consistent.)

Simulated annealing is regularly utilized in discrete, yet huge, design spaces, for example, the arrangement of conceivable requests of urban communities in the Traveling Salesman issue and in VLSI directing. It has a wide scope of utilization that is yet being investigated.

This idea of moderate cooling executed in the simulated annealing algorithm is deciphered as a moderate lessening in the probability of tolerating more terrible arrangements as the arrangement space is investigated. Tolerating more regrettable arrangements is a crucial property of metaheuristics claiming it considers a broader exploration for the worldwide ideal arrangement. When all is said in done, the simulated annealing algorithms fill in as pursues. At each time step, the algorithm arbitrarily chooses an answer near the present one, gauges its quality, and after that chooses to move to it or to remain with the present arrangement dependent on both of two probabilities between which it picks based on the way that the new arrangement is preferred or more terrible over the present one. During the pursuit, the temperature is dynamically diminished from an underlying positive incentive to zero and influences the two probabilities: at each progression, the probability of moving to a superior new arrangement is either kept to 1 or is changed towards a positive esteem; rather, the probability of moving to a more terrible new arrangement is continuously changed towards zero.

Until now, properly tuning the value of n and ai(∀i) in Eq. 3.3–3.5 is the main problem. For achieving high localization accuracy, the no. of breaks n as well as break factors {α1, α2, . . . , αn} can be adjusted practically. The granularity of the ascending fn. w(·) is represented by RSS intervals number n. generally w(.) is expected in better way with the smaller value of granularity to show the relationship among rate based on RSS shifting and RSS change at various stages of RSS.

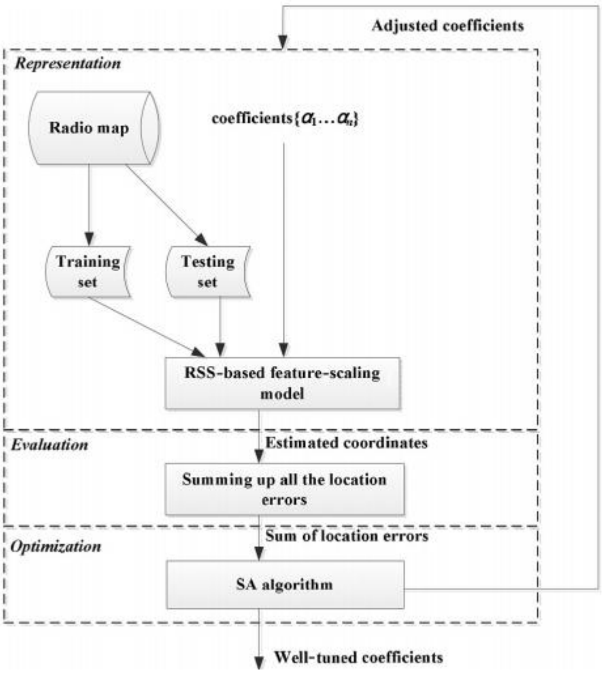


Figure 3.5: Flowchart for FS coefficient tuning in FS-*k*NN [90].

Number of intervals directly relate to the bigger searching space for the SA algorithm implementation as well as good positioning accuracy. Firstly, number of intervals n is supposed to have smaller value say 1 and then the main concentration is to tune coefficient for every interval. Typically, it is a parameter optimization issue, having number of solutions through other parameters. Selection of good parameters to tune the scaling weights, if there is no prior knowledge about important intervals, becomes difficult, if it is not impossible at all. In this technique, however stochastic optimization techniques are used to resolve such type of issues. Coefficients are tuned by SA algorithm. Tuning of coefficient of FS-kNN is described in flow chart given below: FS-kNN model consists of 3 basic steps: 1- Representation. 2- Evaluation. 3- Optimization. These components are carried out iteratively until or unless specific determined situation(s) are fulfilled for tuning. Every outcome of earlier step is input of next step. During optimization, randomly some coefficients are changed and served in the next iteration. Until the preset situation(s) are not fulfilled or specific other considerations are not achieved, the iterative calculation keeps on going to various iterations. The woring of each component is described below.

i) Representation: Two main components are used to represent the FS-kNN model: 1- The fingerprint-based radio map built in offline phase, 2- FS model based on received signal strength (RSS). For the FS-kNN model, the hold out technique is castoff to tune the weights properly. This method divides the whole database in to two sets i.e. an exercise set and a data set. The exercise set is used to locating the MS from unknown positions thus it associates to the RPs (database) with identified axis. Whereas for the assessment of performance of coefficient acquired in each iteration, testing data is used. When the new set of coefficients is obtained, FS-kNN model feeds the testing set to create estimated coordinates for every entity present in the set.

ii) Evaluation: As coefficient set is obtained, computation of sum of location error is needed to evaluate the resulting performance of these coefficients, which is mostly denoted as cost, and is given below:

$$Cost = \sum_{i=1}^{m_1} \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \qquad 3.6$$

where (xi, yi) is the genuine axis of the ith element in the trying set in Eq. 3.6, (xi, yi) is its assessed axis, and m1 is the entire no. of features in the trying set. The most important thing is that, in evaluation process, the RPs present in the trying set are taken as anonymous locations. Larger sum results the higher positioning accuracy. Thus, optimal solution is achieved by making cost =0 using coefficients.

iii) Optimization: For better accuracy, searching of new coefficient is done by SA, which utilizes the variable like temperature for cooling schedule. This temperature is going to lower and lower after iterative measurements. Randomly some coefficients are changed in every iteration. If the fresh calculated cost has smaller value, the fresh factors will automatically have converted into the present factors. And if it is not happened, new factors then still have some probability p to serve as present coefficients. And this probability is calculated as:

$$p = e^{-\frac{\Delta Cost}{temperature}}, \qquad \Delta Cost \geq 0 \qquad 3.7$$

As far the dissimilarity among the fresh cost and the optimum negligible cost is represented by Δ cost. If a predetermined maximum number of iteration is not achieved or some scheduled conditions are not met, the above process keeps on going. During the online phase, the time complexity of FS-kNN model is $O(M \times L)$ whereas L is the no. of APs and M is the number of RPs.

Whereas the time taken by offline phase can be found as $O(M \times L \times N \times K)$, K represent the no. of repetitions used in SA for regulation the factors of FS-kNN and N denotes the size of the testing set. This shows that FS-kNN consumes more time only in offline phase; however, time taken by localization algorithms is not affected in online phase [90].

3.2.5. Outlier

Outlier data can be defined in various ways. According to its one definition which is defined as, in either direction from distributions middle, values lie very far, and these far values are known as outlier.

There is another way to define outlier that is, outlier values are from away from bulk of the values and it consist of single and very low frequency values of variable.

Moreover, it is also defined as, a group of observations which consist of those values that are deviate from expected values. In another research it also said that outlier consist of values which are very different as compared to other values of the group to distinguish its self and to prove that these values are generated by some other mechanism.

In literature of data mining and statistics, outliers are taken as abnormalities, discordant, deviants or anomalies. In overall data sets of values, the subset which appear as abnormal or inconsistent is also known outlier according to [96].

There is a problem with identification of outlier values that its values are not easily distinguished from other values of data set because observed values may be the output of low probabilities values or of the perfectly extreme values of same distribution.

Outlier dataset is generated by some other mechanism, so its dataset is very different from rest of the observation. So by above mentioned definitions it is cleared that outlier values are abnormal while other data sets are referred to as normal or expected values.

But it is also said that outlier’s data is totally different from noise values because noise is a random calculation of error and in measured value it is variance. Before the detection of outlier noise is removed from the dataset as in outlier detection and data analysis, noise does not take any interest. If noise is taken into account in dataset then there are lot of false alarms and it also take lots of time and cost is also one of the major issue.

3.2.5.1. Causes

There are lots of anomalous causes for outliers. Transient malfunction is the major cause in taking measurements for physical apparatus. In data transmission and reception there is error in calculations.

There are lot of factors for creating outliers and it is mainly occur due to behavior change of environment, human disturbance, and apparatus error through natural deviations in populations. Due to other factors of population, samples may be contaminated with other elements.

Outlier may be caused due to the error in the supposed theory, which is major challenge for researchers now. Additionally, it is also said that, outliers that appear in pathological way from different datasets is indicating causative mechanism for the calculation is extremely different from extreme ends of data.

Assurance of quality for datasets is the main reason for measuring the outliers [96]. So, if there is need to improve quality and Impact of data analysis outliers must be removed and replaced from the other values of the data.

Sample mean and standard deviation which are some simple statistical calculations can be easily distinguished from outlier's calculation which have huge difference from middle of the distribution. In current literature, detection of outlier values is one of the major work in data analysis and it involves the identification of calculation or values from dataset which are far away from expected values.

Outlier values have major effect on analysis of data and sometimes it leads to fault calculations of results. So, it is very necessary to pay a proper attention on the detection and analysis of outlier's data sets especially it is more important when an important result have to be made from deviated and random data values. There are lot reasons to detect the outlier. It has effect on results and calculations, so its proper Investigation is very important before the analysis of data. It is very necessary to protect the outlier data for the assurance of quality of data. Many times, outlier calculated values are the result of fault calculation, inappropriate use of resources and apparatus. So, analysis of the data before final results is very necessary. When large number of calculation are observed it is advisable to detect the outlier values as these values may contain very important data although often these calculations are considered as noise and extra information. To detect the outlier in a dataset a test is performed which is called as Thompson Tau. This test provides the rejection zone which is statistically calculated after taking account the standard deviation and average of the data set calculations. So, for calculating the outlier in a dataset it provides an objective function [96].

3.2.5.2. Outlier Deleting

Here a set of k RPs as closest neighbor points is obtained after applying FS-kNN model. It can also be expressed in coordinate set as C = {RP1, RP2, ..., RPk}.

The matching process can be affected by a geographically dispersed set of RPs because of signal measurement ambiguity which results as poor estimation accuracy [91]. So, there might be a possibility of an outlier present in the set C. In statistical environment, an outlier may be defined as a point among all observation which is distant from other observations. In other words, from a clustering center of a set, a point which is far away is considered as an outlier. As the position is estimated by remaining k neighbor points directly, therefore final localization result takes a huge impact of the outlier. In this process a method is required to identify the outliers. For this purpose, an altered Thompson Tau test offered by John M.Cimbala [92] is used in the proposed model. That statistical decision can be made as:

$$\delta_x = \frac{|x_i - x'|}{std_x} \geq \tau, \tag{3.8}$$

$$\delta_y = \frac{|y_i - y'|}{std_y} \geq \tau, \tag{3.9}$$

If (x’, y’) and stdx, stdy are the average value of set C and standard deviation corresponding to the x and y respectively, and given circumstances are fallen then outlier is identified as xi and yi and would be removed from the C set. And feature scaling weights related to that outliers will be deleted as well. Where the rejection threshold is $\tau > 0$ and given as:

$$\tau = \frac{t_{\frac{a}{2}}(k-1)}{\sqrt{k}\sqrt{k-2+t^2_{\frac{a}{2}}}} \tag{3.10}$$

In the Student t distribution, α is the desired confidence, stated as a fraction, after selecting k nearest neighbors from FS-kNN model, k is the model size and

$$std_x = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(x_i - x')^2}, \tag{3.11}$$

$$std_y = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - y')^2}, \qquad 3.12$$

As the size of set C is with insufficient sample so t-distribution is used. The $t_{(\alpha/2)}$ can be obtained from the t-distribution table directly.

After deleting the outlier, a set C' consisted on novel clarified nearest neighbor RPs and their equivalent active signal distance vector set [dw1, dw2...dwl] ($l \leq k$) are extracted. And then from these extracted vectors, one of the vector is selected as user estimated position that is based on least active signal distance.

Chapter 4 - Results and Discussion

This chapter is mainly based on simulation results and discussion. Proposed model's working has been explained in the previous chapter. Here are the implementation results of the model in the simulations form. Additionally, comparison has been performed among different positioning algorithms. Case study has also done for performance analysis of the model by changing different parameters. The impact of these changes is discussed as well.

4.1. Simulation Model

Indoor environment of target area is the first floor of the building covered by certain number of WLAN APs. In fingerprinting technique, a radio map would be needed to build, so the two-dimensional floor has divided into square grids of 1m by 1m. In the offline phase, the readings from Wi-Fi Access Points (APs) have been taken at the centers of those grids which are also called as calibration points or Reference Points (RPs) and stored in the fingerprinting database according to the coordinates of the grid points, so it can be used for matching purposes later in the online phase. The coordinates of 2D floor plan are represented as (x, y).

4.2 Building Offline Fingerprint Database

Coordinates of the area have been taken from the picture, so each pixel value of the blue print needs to convert in real time distance (m). For this purpose,

$$\text{pathUnit} = \text{pathLength./pathPixels} \tag{4.1}$$

Eq. 4.1 is used. The pixel value (Pix) for division along x-axis and y-axis to form uniform grid is taken 13 for x-axis and 15 for y-axis. The readings are calculated at the center points of each grid, so the x, y coordinates of the Reference Points are uniformly distributed.

The signal strength of received signal can be used to estimate distance because the power (or RSS) of electromagnetic waves is inversely proportional to the distance as shown in the following expression:

$$P \propto 1/d \tag{4.2}$$

Here P is the RSS value at a distance d from the transmitter. Practically the RSS value along with distance depends on other environmental factors also. So, path loss coefficient or exponent is added in the equation to present a complete path loss model. The path loss coefficient represents the behavior of the environments, specified in the path loss model along with RSS readings.

4.2.1. Overview

It is cleared from the previous literature of this research that the block diagram is followed to achieve expected results. The whole procedure has two phases. Each phase has few steps. And the offline phase is based on the following steps:

- Extract image as blue print
- Meshing the floor plan
- AP deployment
- Implementation of Motely Keenan model
- Calculate RSSI vectors from all AP at each RP.
- Make radio map
- Verification of all APs' signal strength by heat map.

4.3 System Parameters

The main system parameters are given in the table 4.1.

Parameter	Value
No. of RPs	900
Grid size	1m*1m
No. of APs	4
Reference Distance	1 m
Transmission power	14 dBm
No. of samples collected at RP 'samp'	50
k in FS-kNN	5
k in FS-kNN with deleting outlier	5
No. of online queries	1000
Frequency	$8.652 \text{x} \ 10^{8}$ Hz

Table 4.1: System parameter.

4.4 Results

The results of each step in the offline phase are given and discussed in the following section.

4.4.1 Floor Plan

The floor plane is extracted as blue print image. Figure 4.1 represents the targeted area of indoor environment, having structure with black lines with width of 1 pixel. Image has dimensions 400*450.

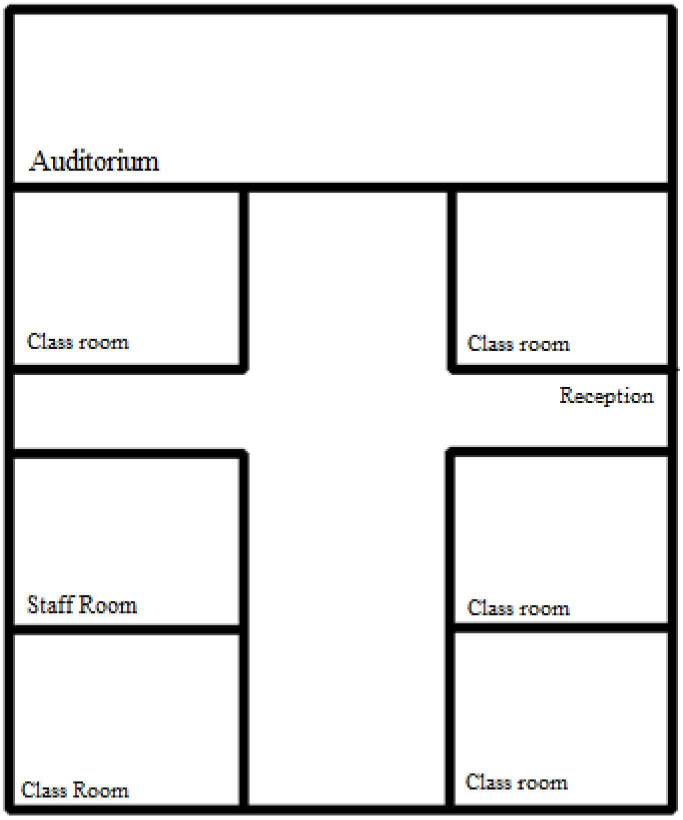


Figure 4.1: Floor plan of the targeted area.

4.4.2. Grids and Reference points

The fingerprinting-based localization system is highly dependent on the grids and RPs. The area of interest is divided into equally spaced grids, and RSS samples are gathered at the center point or corner point of the grid shown in figure 4.2. The granularity of the RPs highly affects localization accuracy. In literature [40-45] it is stated, for a better accuracy, more RPs are required so more refined results can be obtained. But there is a point that is, more the RPs, higher the accuracy until it reaches a threshold. The grid points or RPs should not be that mush close to each other as no significant change in RSS level is observed. Whole area is divided into equal grids and then reference points are set in each grid according to the required database.

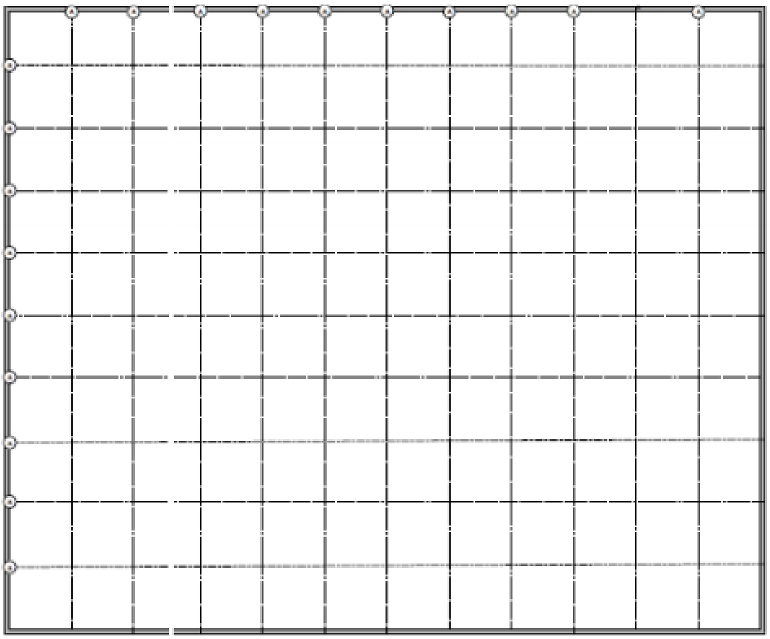


Figure 4.2: Grids having equal size in x-y coordinates.

These reference points (pixel values) are then converted into real time distance. In this technique, the 900 reference points are created having equal distance of 1m as shown in figure 4.3:

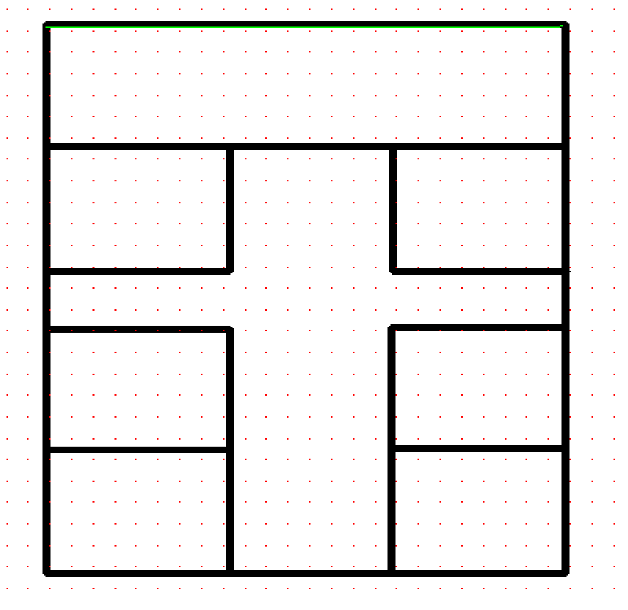


Figure 4.3: Reference point

To convert the pixel value into real distance, the two points of any wall are selected and then length between these points is entered. A formula is used to calculate the real time value in meter from pixel value. The same expression Eq. 4.1 discussed above is used here:

$$pathUnit = pathLength./pathPixels$$

From the figure, the two points have been selected from the upper most line and given that the length of 22 meter. That is denoted as pathLength in the above Eq.4.1. Whereas pathPixel is the pixel value of the image/blue print of the building.

Pathunit is multiplied with maximum number of pixel and the maximum distance value is obtained. Hence, real distance can be gotten of any value in the image (in pixel) by multiplying it with pathUnit. The value of the pathUnit is obtain as: 0.075.

So, the total area in meter is calculated as 30*33 m^2 of 400*450 pixels image.

4.4.3. AP deployment

Four Access points are manually deployed at different locations of the targeted area. It is very important to select area for the deployment from where AP can have maximum coverage. So, number of experiment have been performed to find optimal solution, that is all APs would cover the whole space. AP deployment is shown in the figure 4.4 below:

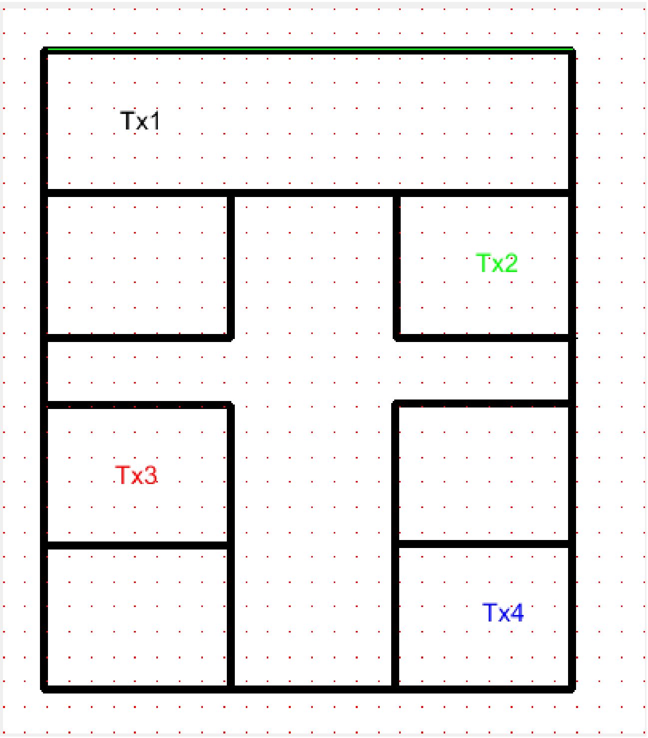


Figure 4.4: Deployment of four access points.

Locations of deployed APs in real distance are as:

AP1= [4.4670, 4.9334] (a)

AP2= [10.738, 19.333] (b)

AP3= [19.004, 4.2667] (c)

AP4= [24.605, 18.668] (d)

4.4.4. Euclidean Distance

After deploying APs, first Euclidean distance is calculated from each access point to all reference points and this distance is called as Euclidean distance and can be calculated as:

$$d = \sqrt{\left(x_2 - x_1\right)^2 \left(y_2 - y_1\right)^2} \qquad 4.3$$

A1 and A2 in Eq. 4.3 are distance in terms of pixels and B is the distance in terms of meters

In figures 4.5(a-d), the results as ‘distance map’ after calculating the distance have been represented from all APs. Distance is very less near to the AP. And as users go away from the AP, the distance increases. And it can be said that reference points with least distance values have strong signal strength or good coverage of AP.

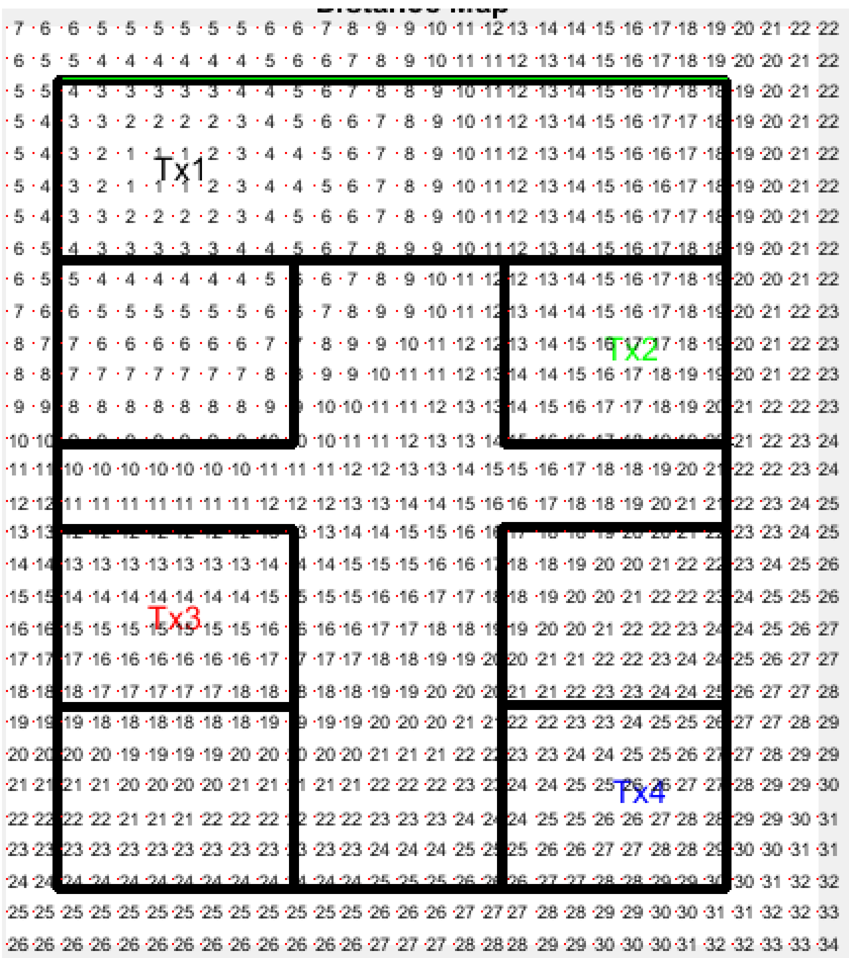


Figure 4.5a: Distance map of TX1.

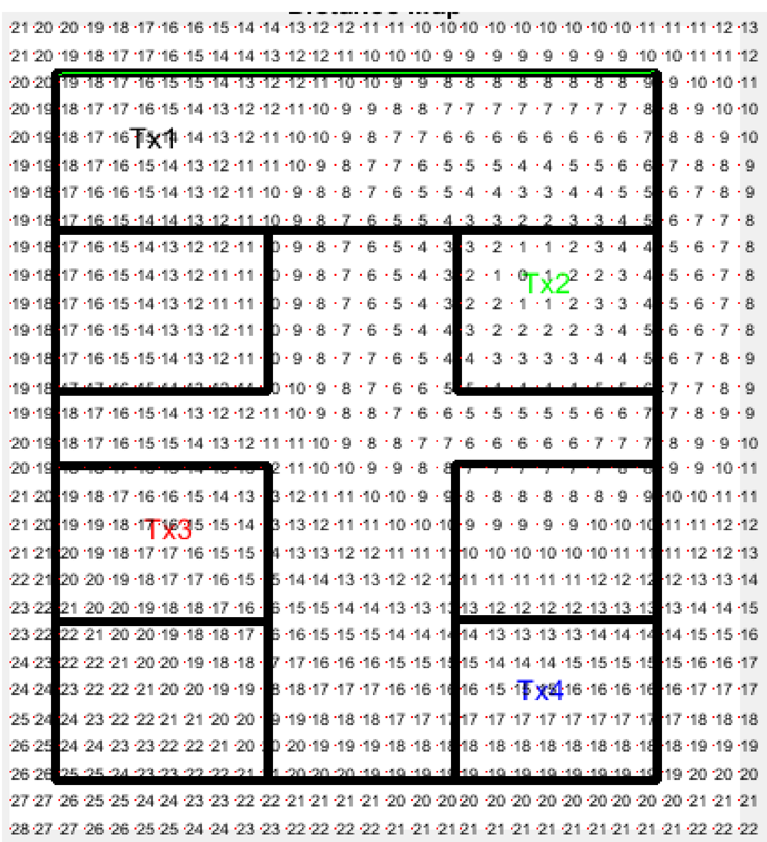


Figure 4.5b: Distance map of TX2.

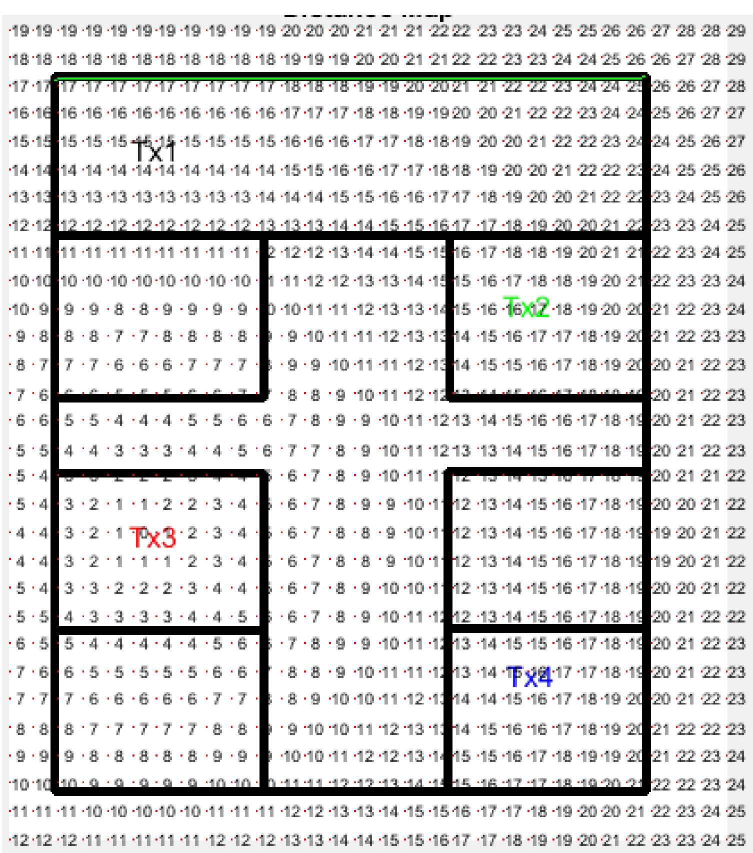


Figure 4.5c: Distance map of TX3.

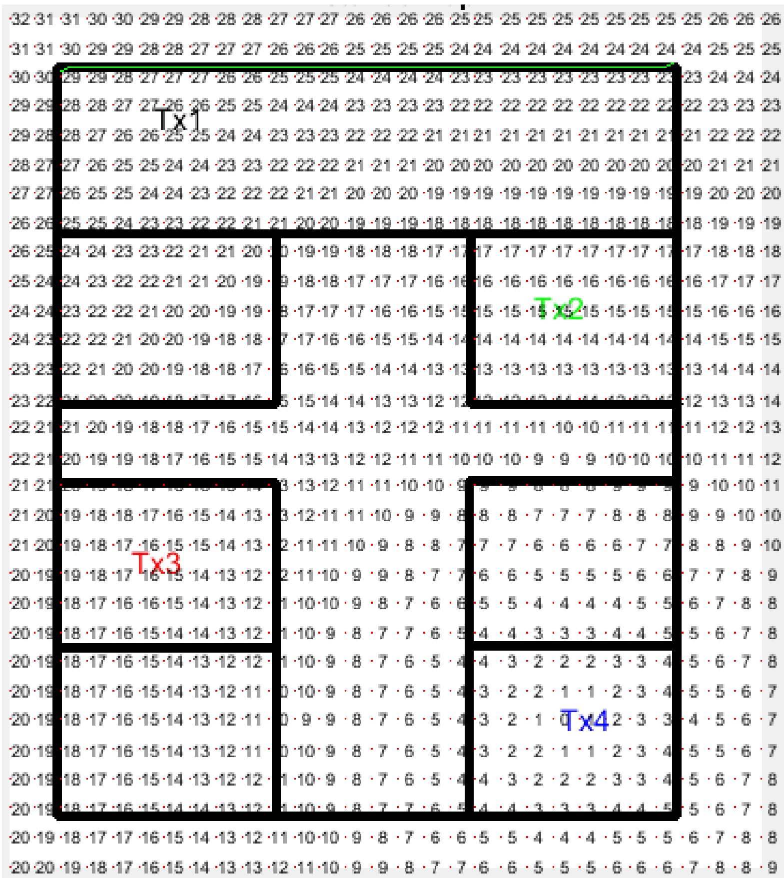


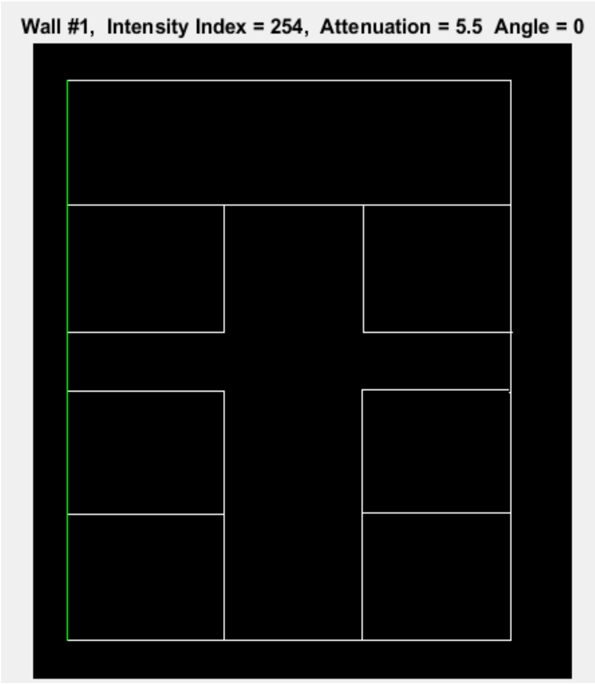
Figure 4.5d: distance map of TX4.

4.4.5. Implementation of Motely Keenan model

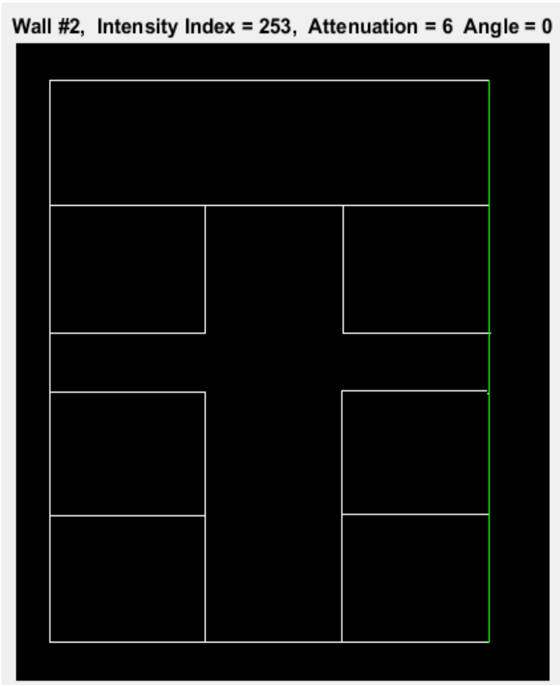
In this step, a path loss model which is named as Motely Keenan multiwall propagation model is implemented to calculate total path loss due to obstructions and hurdles. Number of walls are counted between transmitters and receivers to calculate total path loss. Wall type is examined as well to account the thickness of the hurdle. That's why each wall is calibrated separately by the model. Model has its mathematical expression as:

$$L = L_O + 20\log(d) + \sum a_k \qquad 4.4$$

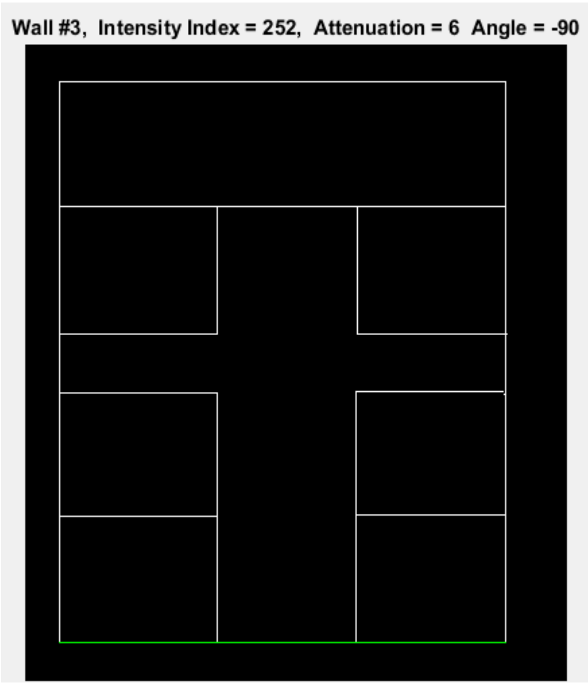
In the figures 4.6(a-o) given below, each wall which is calibrated is represented by green color. Based on intensity index of the wall, this model helps to find the attenuation factor of the wall with angle of incident.


Wall #1, Intensity Index = 254, Attenuation = 5.5 Angle = 0

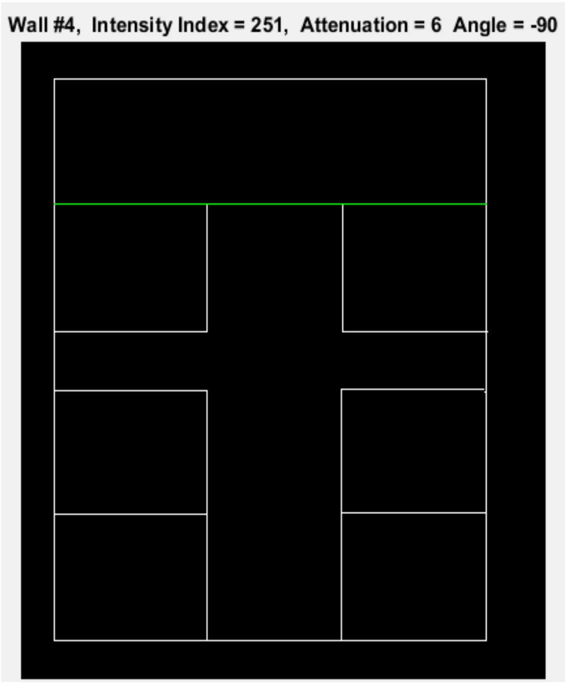
(a)


Wall #2, Intensity Index = 253, Attenuation = 6 Angle = 0

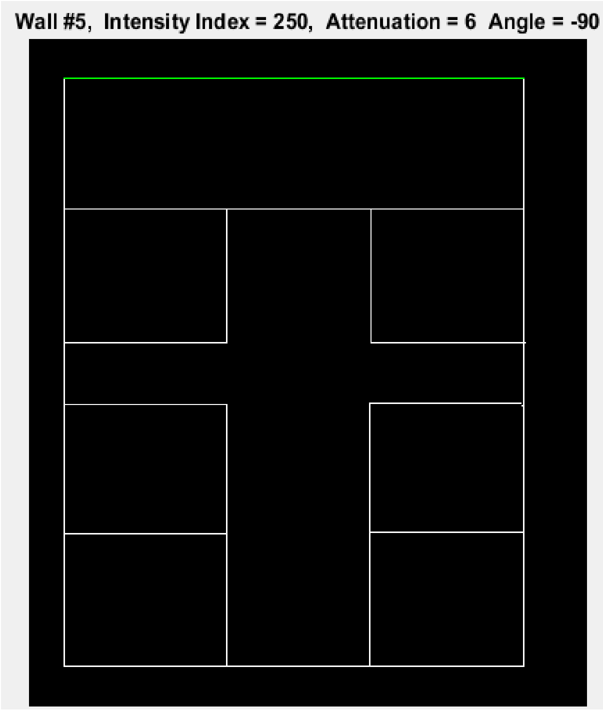
(b)


Wall #3, Intensity Index = 252, Attenuation = 6 Angle = -90

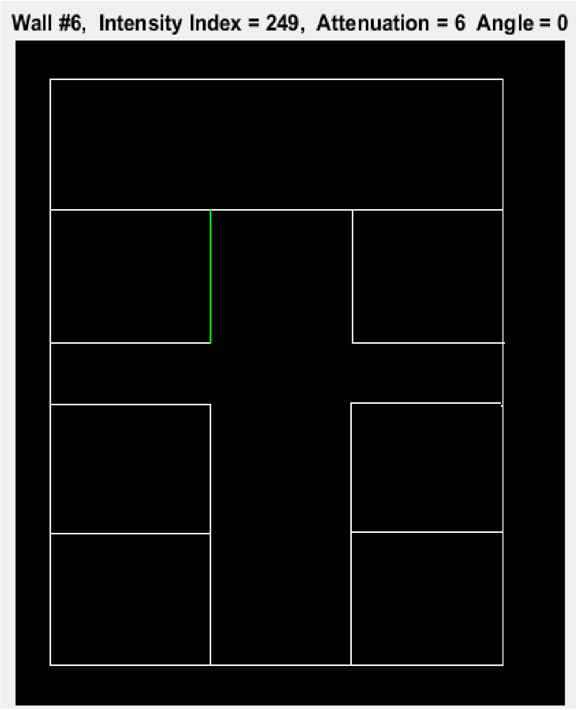
(c)


Wall #4, Intensity Index = 251, Attenuation = 6 Angle = -90

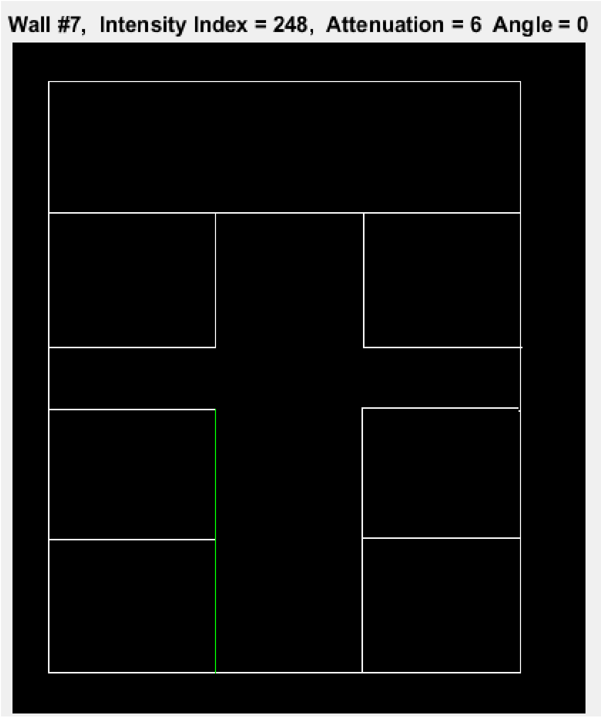
(d)


Wall #5, Intensity Index = 250, Attenuation = 6 Angle = -90

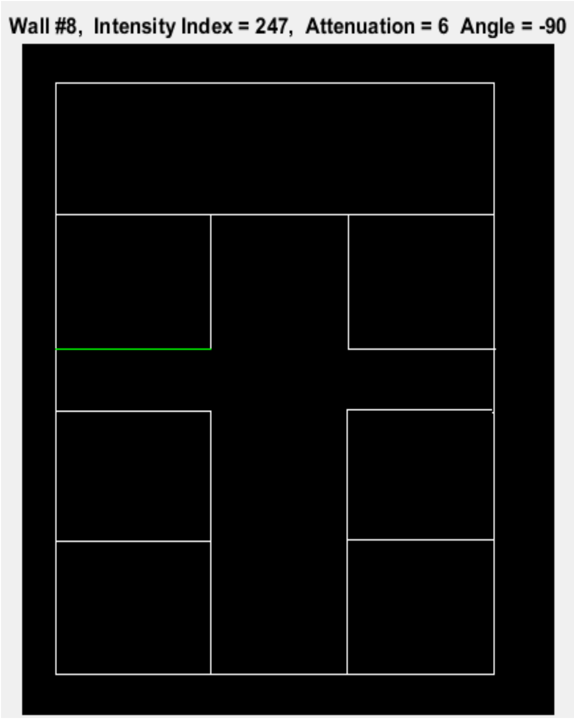
(e)


Wall #6, Intensity Index = 249, Attenuation = 6 Angle = 0

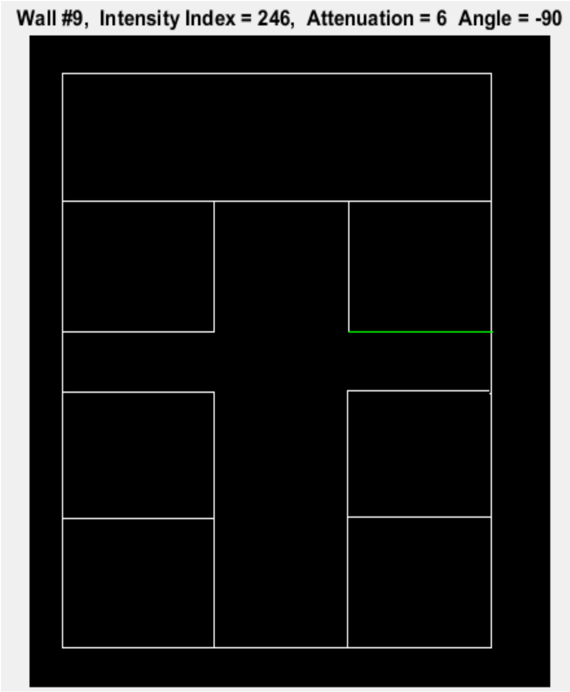
(f)


Wall #7, Intensity Index = 248, Attenuation = 6 Angle = 0

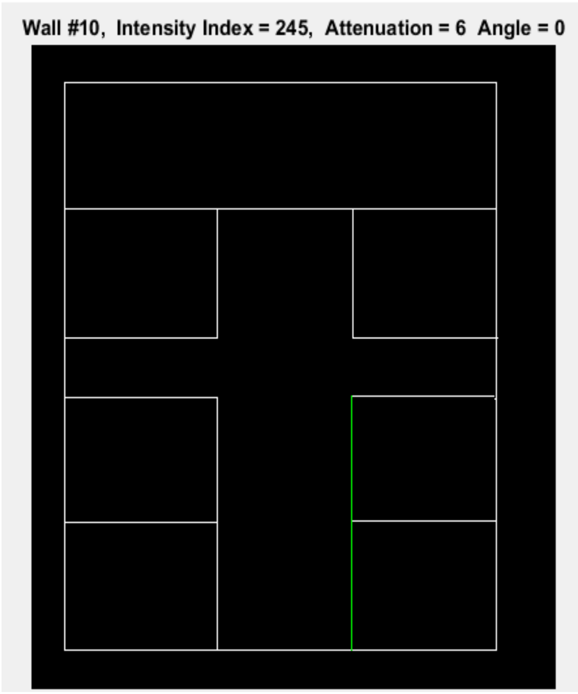
(g)


Wall #8, Intensity Index = 247, Attenuation = 6 Angle = -90

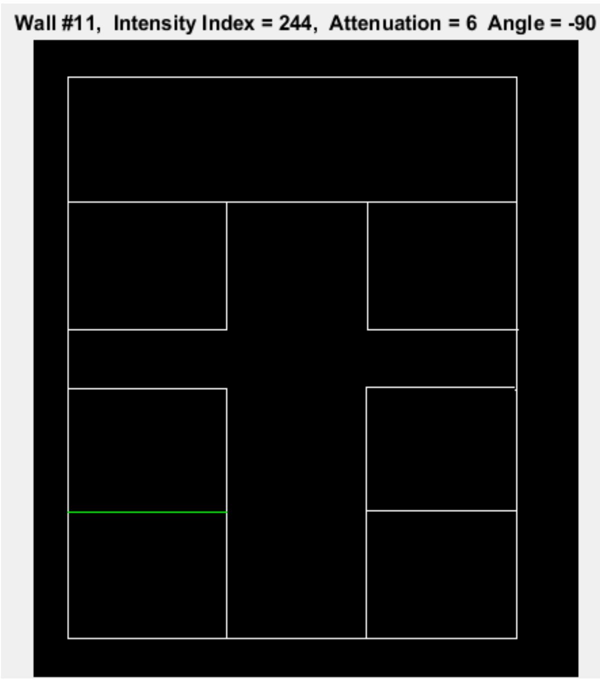
(h)


Wall #9, Intensity Index = 246, Attenuation = 6 Angle = -90

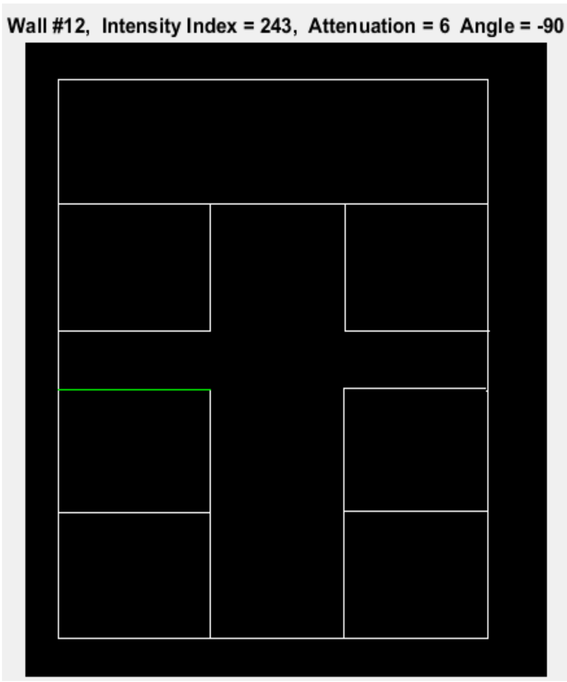
(i)


Wall #10, Intensity Index = 245, Attenuation = 6 Angle = 0

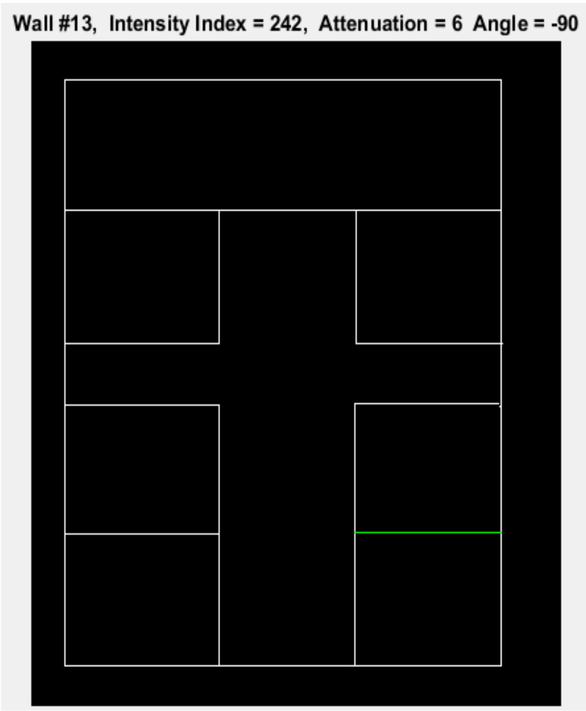
(j)


Wall #11, Intensity Index = 244, Attenuation = 6 Angle = -90

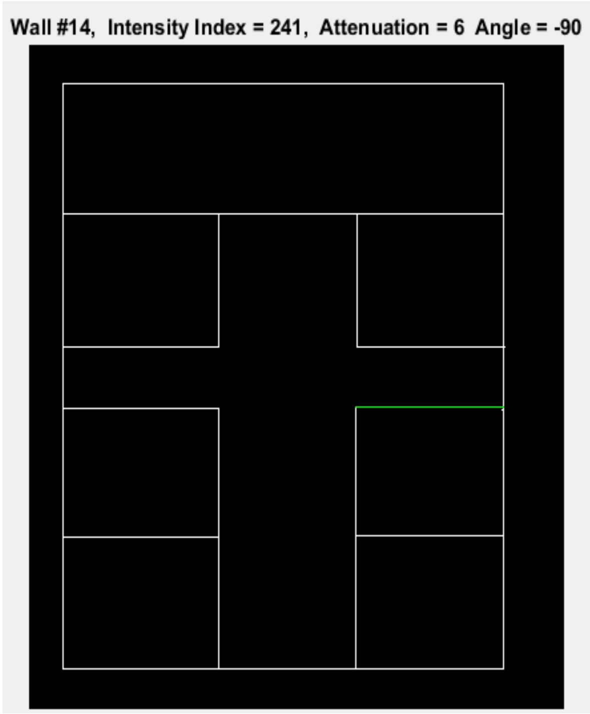
(k)


Wall #12, Intensity Index = 243, Attenuation = 6 Angle = -90

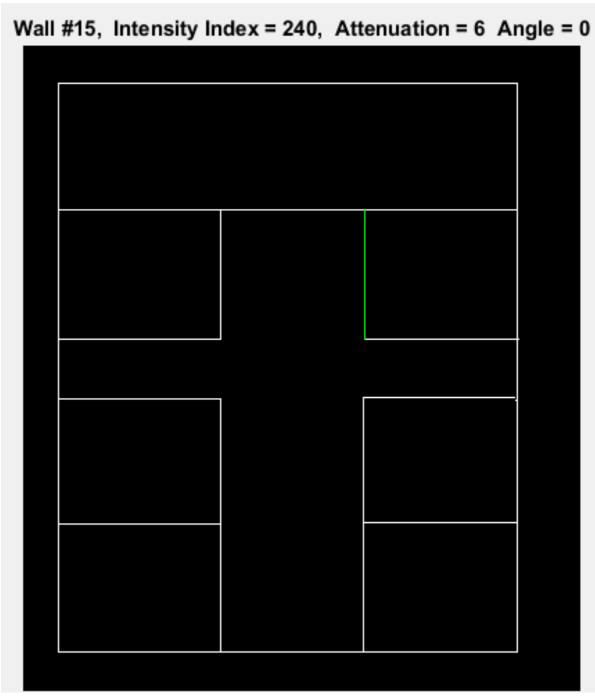
(l)



(m)



(n)



(o)

Figure 4.6 a-o: Calibration of each wall of the structure.

4.4.6. Radio map

After applying the propagation model, RSSI vectors are calculated at each reference point by using formula given above. The RSSI vector at some distance has unit of dBm.

Once these values are calculated for all reference points then these values are stored as fingerprints. Single samples of RSSI recorded from the nearby access point are not sufficient to characterize a fingerprint. Due to presence of noise in the environment it is necessary to obtain an average of the readings to successfully identify a fingerprint. Therefore, 40 values have taken at each RP to mitigate the noise effect by taking average of these values. Hence, a radio map is created in which each reference point has four RSSI values from four APs perspectival. So, a matrix of 900x4 is constructed. Fingerprints of few reference points are shown in figure 4.7 from the radio map of all reference points.

FingerP

900x4 double

	1	2	3	4
1	-53.4378	-75.9985	-80.1888	-90.3084
2	-52.0731	-75.3180	-79.7340	-90.0807
3	-51.2551	-75.1385	-73.2234	-89.8354
4	-50.6871	-74.9841	-72.7185	-89.6037
5	-50.4098	-74.8362	-72.1493	-89.3550
6	-50.5380	-74.7078	-71.5453	-89.1053
7	-50.9985	-80.6067	-70.9441	-82.8710
8	-51.7484	-80.5210	-70.2620	-82.6217
9	-52.5858	-74.4629	-69.5810	-76.3893
10	-53.5317	-74.4254	-68.8068	-82.1440
11	-60.4746	-74.4139	-67.9810	-87.9027
12	-61.3279	-74.4270	-55.6603	-87.6815
13	-62.1923	-74.4661	-60.2412	-87.4525
14	-63.0051	-74.5304	-59.2938	-81.2325
15	-63.7206	-80.6128	-58.4123	-81.0364
16	-64.4367	-80.7227	-46.0427	-80.8401
17	-65.0672	-80.8449	-50.8841	-74.6701
18	-65.6998	-74.9939	-50.4680	-74.5062
19	-60.7948	-75.1601	-50.4422	-74.3624
20	-72.8216	-75.3297	-50.7800	-74.2481

Figure 4.7: Radio map of first 20 RPs.

4.4.7. Propagation Map

This is the last step of offline phase in which RSSI value of access points is verified. By using this methodology, one can identify that whether the AP covers the whole space or not. In other words, it is type of observation to make sure that each reference point has a strong coverage of any AP on it. In figure (4.8a-d), the RSSI verification of deployed APs is done. The area from high signal strength to weak signal strength changes color from red to blue, respectively.

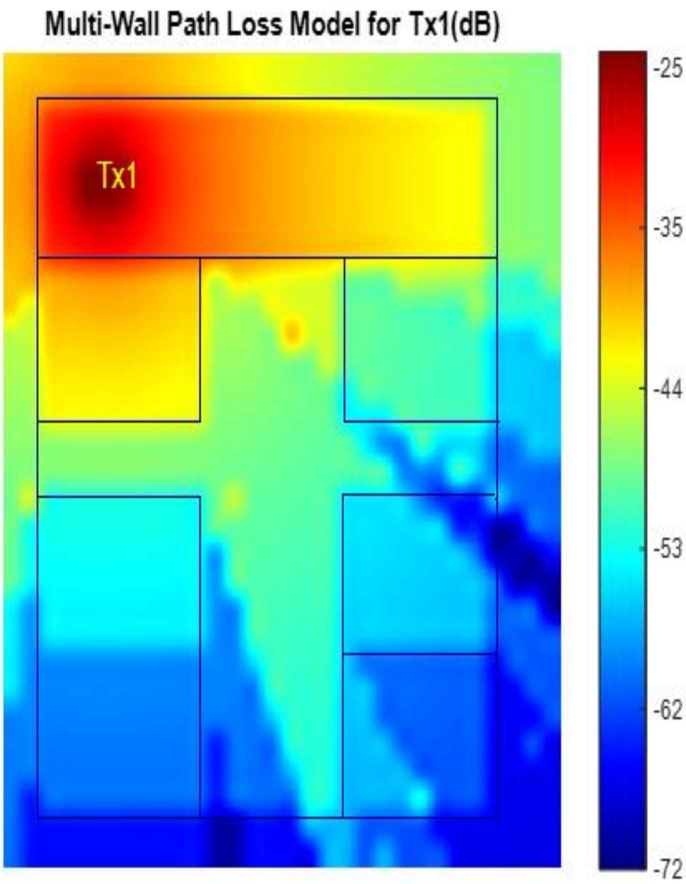


Figure 4.8a: Heat map of TX1

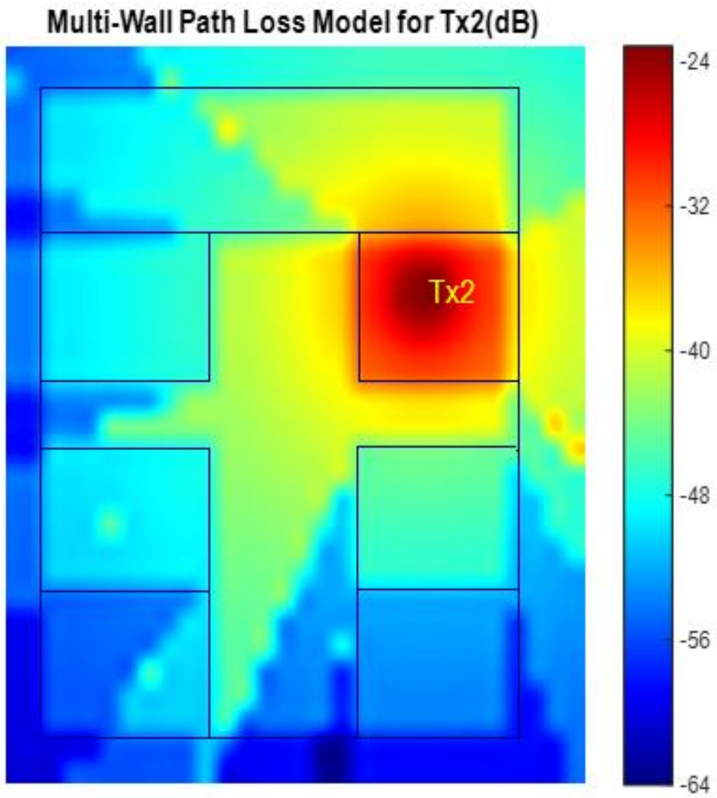


Figure 4.8b: Heat map of TX2

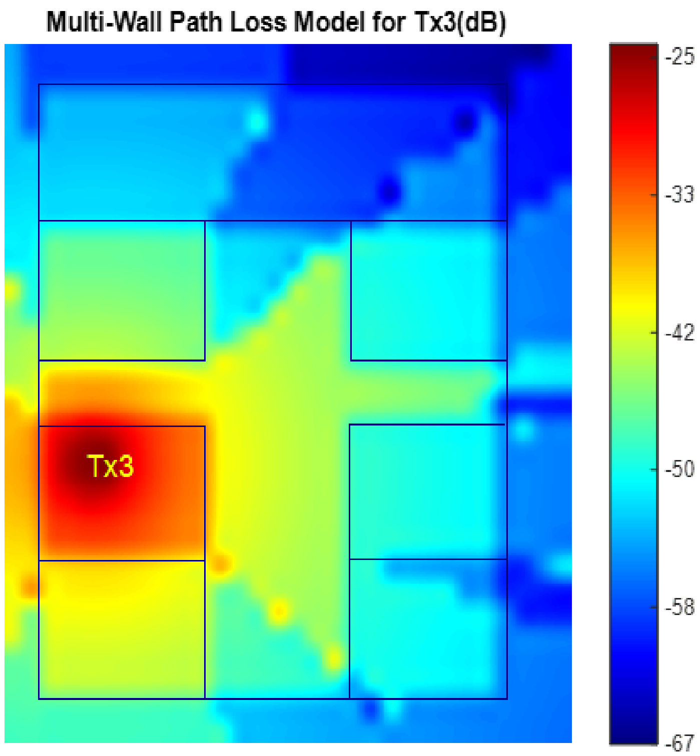


Figure 4.8c: Heat map of TX3

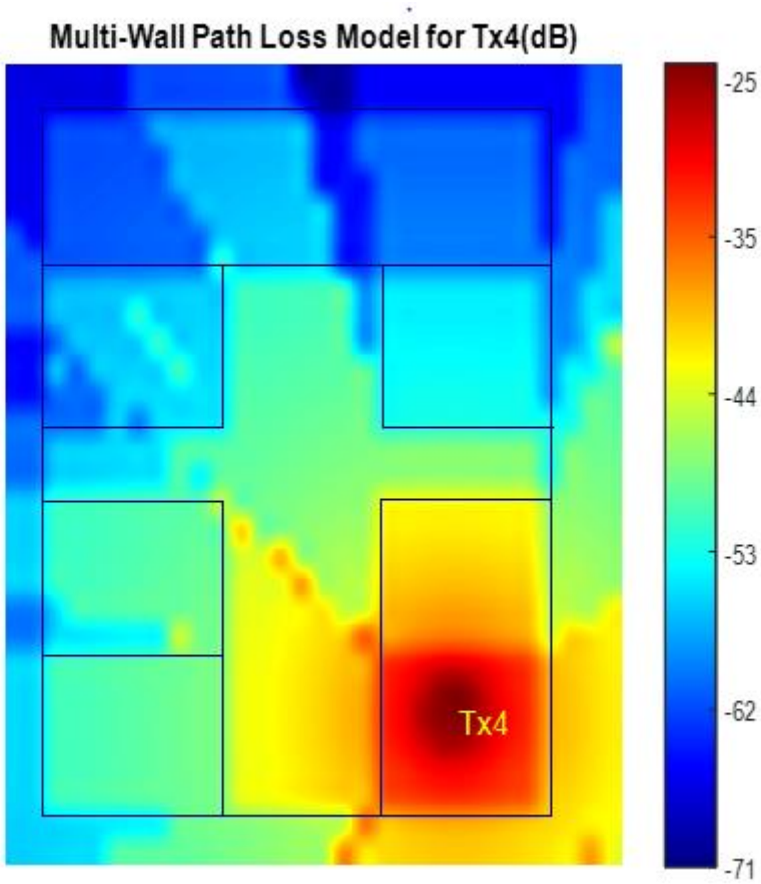


Figure 4.8d: Heat map of TX4

4.5. Online Phase

In this phase random users are taken at different locations from the targeted area and the distance is calculated in same way as in offline phase for RPs. Same propagation model is used to calculate the total path loss between users and APs. And the online RSSI values are calculated from all APs for each user. It is worth noting here that, this time the RSSI values calculated are at user side and these values are sent to server which will then use matching algorithm to compare fingerprints with online query and gives the best results as positions (in x-y coordinates) of all users. As it is mentioned earlier, Existing matching algorithms for positioning, which are: 1- k-Nearest Neighbor (kNN) algorithm, 2- Feature Scaling k- nearest Neighbor (FS k-NN) algorithm, 3- Feature Scaling k- Nearest Neighbor (FS k-NN) algorithm with deleting outlier. FS k-NN and FS k-NN with deleting outliers are the variants of k-NN that provide more accurate matching results than simple k-NN. In below section, positioning results are shown by using cumulative distributed function CDF. The CDF represents the overall percentage error of positioning of all users. In the end, the CDF results are compared of all algorithms as well. This comparison will clear the fact that which algorithm is best as matching algorithm to locate the users and will show the efficiency of the proposed model.

Online data (RSSI vector regarding users' position) for first 20 random user is shown in figure 4.9:

P1ower

1000x6 double

	1	2	3	4	5	6
1	2.7251	6.5136	-38.4314	-66.6998	-73.0658	-82.9651
2	2.7251	4.7190	-35.6569	-67.5645	-73.0210	-83.2886
3	5.5166	7.9094	-41.2796	-65.1086	-71.5343	-81.7829
4	18.8097	20.3384	-75.6755	-61.7645	-66.9509	-52.0683
5	15.3535	18.2115	-67.9113	-51.3804	-59.8619	-61.9877
6	9.2387	7.7764	-52.1960	-64.5375	-62.7206	-74.4426
7	11.4985	8.1752	-55.0505	-64.2375	-60.7098	-73.4216
8	12.3625	12.0302	-63.7604	-54.8486	-56.5073	-65.7191
9	6.9789	24.4592	-63.0787	-58.8669	-82.2556	-80.3266
10	24.5921	25.3233	-90.3466	-72.9371	-81.7759	-53.8346
11	24.8580	6.3142	-81.4272	-74.9978	-53.7398	-64.9508
12	18.0785	14.4230	-67.6215	-56.4420	-56.7446	-60.4075
13	23.3293	23.3293	-83.6213	-71.8171	-74.8008	-44.8829
14	11.0332	21.7341	-68.3288	-38.8582	-68.4378	-71.7341
15	21.2689	18.5438	-75.9112	-63.9601	-66.0755	-46.1010
16	7.1118	3.6556	-40.7366	-73.3127	-70.1974	-76.2506
17	8.8399	3.3897	-50.6413	-67.3157	-62.7942	-75.8240
18	11.5650	25.6556	-76.0106	-53.2550	-75.9519	-72.3378
19	6.7795	7.5106	-42.1339	-65.0713	-70.6122	-81.4204
20	21.7341	3.9215	-73.9831	-68.8544	-42.3032	-72.5930

Figure 4.9: Online data of first 20 random users.

4.5.1. k-NN Algorithm

This algorithm calculates the position based on the distance among the reference pattern and patterns present in the database. Various formulas can be used for distance calculation, e.g.: Manhattan distance or Euclidean Distance. It finds the k best matching fingerprints based on their mutual distance and calculates the position as the average of positions of same k patterns. After that, to find the expected position, the reference points relative to these k fingerprints are used. However, to find the distance among the assessed RSSI vector and one of the fingerprints, the vectors serve as the key component in the pattern matching algorithms. It can be said as well that the resemblance between two objects to be matched is essential. The distance can be calculated by the formula given in Eq. 4.5.

$$d_m = \sqrt{\sum_{n=1}^{N} (RSSI_{m,n} - r_n(t)^2} \qquad 4.5$$

In this work, value of k is settle as 3, as from the online data, the coordinate position of first random user is [2.73, 6.51] and then k-NN will select the 3 best possible solutions from the databases based on minimum distance. The coordinate values of these 3 RPs are taken from the results and then the algorithm picked the close enough to online query which is ‘D1 = [2.06, 6.05]’.

these results are also shown in the simulation form in figure below:

```
The Estimate Position | X = 2.06,Y = 6.05
The Actual   Position | X = 2.73,Y = 6.51
Percentage Error = 0.81
```

This is the result for one user. As the model is designed to locate 1000 user at a time so the total location error can be seen in the CDF plot of k-NN algorithm given below:

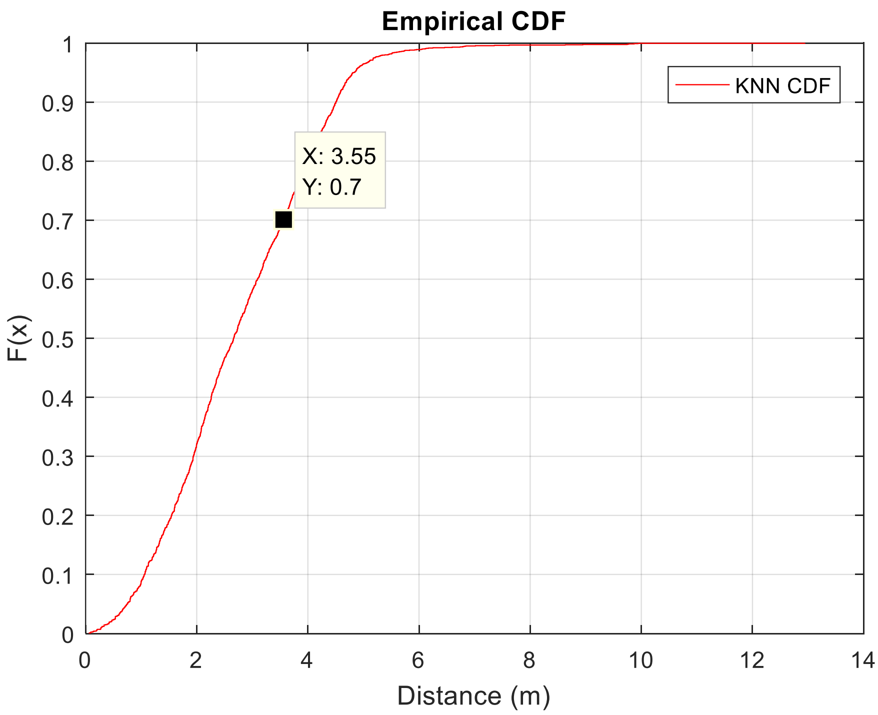


Figure 4.10: Total error CDF plot of kNN

From the plot shown in figure 4.10, we can say that 70% results are with distance error of 1.5 meter. That is better than the previous results using simple path loss model. So, one can say that proposed technique for offline phase has worked it to improve the positioning error even with most light matching Algorithm i.e. k-NN.

4.5.2. FS-kNN

FS-kNN also improves the kNN algorithm in two ways which are: 1- during offline period, FS-kNN not even builds database as well as tuning scaling weights for RSS-level-based FS model. 2- during online phase, in the calculation of active signal distances, it makes use of tuned scaling weights. Between online query reported by mobile user and database of each reference, the effective signal distances are calculated to find the k=5 databases for positioning.

In this algorithm, RSS level-based scanning is introduced to measure a new scanning feature to measure the difference of effective signal during the corresponding synchronization among various signal vectors. d'm show an effective signal distance among online query and (fingerprint) attached to the mth RP, so that it can be calculated by:

$$d'_m = \sqrt{\sum_{l=1}^{L} (RSS_{m,l} - RSS_l)^2 * w(RSS_l)} \qquad 4.6$$

The detail of the model is discussed in the previous chapter. So here, the basic points are will be discussed which directly related to the results.

FS-kNN model consists of 3 basic steps: 1- Representation. 2- Evaluation. 3- Optimization. These components are carried out iteratively until or unless specific determined situation(s) are fulfilled for tuning. Every outcome of earlier step is input of next step. During optimization, randomly some coefficients are changed and served in the next iteration. Until the preset situation(s) are not fulfilled or specific other considerations are not achieved, the iterative calculation keeps on going to various iterations.

4.5.2.1. Representation The training data and testing data is made in this section to analyze the performance of the model. The training set is used to locating the MS from unknown positions thus it associates to the RPs (database) with identified axis, whereas for the assessment of performance of coefficient acquired in each iteration, testing data is used. Figure 4.11 and figure 4.12 are the first 20 training and testing data respectively.

Training_data

450x6 double

	1	2	3	4	5	6
1	0.0665	0.0665	-53.4378	-75.9985	-80.1888	-90.3084
2	2.1269	0.0665	-51.2551	-75.1385	-73.2234	-89.8354
3	4.1873	0.0665	-50.4098	-74.8362	-72.1493	-89.3550
4	6.2477	0.0665	-50.9985	-80.6067	-70.9441	-82.8710
5	8.3082	0.0665	-52.5858	-74.4629	-69.5810	-76.3893
6	10.4350	0.0665	-60.4746	-74.4139	-67.9810	-87.9027
7	12.4955	0.0665	-62.1923	-74.4661	-60.2412	-87.4525
8	14.5559	0.0665	-63.7206	-80.6128	-58.4123	-81.0364
9	16.6163	0.0665	-65.0672	-80.8449	-50.8841	-74.6701
10	18.7432	0.0665	-60.7948	-75.1601	-50.4422	-74.3624
11	20.8036	0.0665	-73.3536	-75.5234	-51.4438	-80.1498
12	22.8640	0.0665	-74.3065	-75.9287	-53.1749	-74.0351
13	24.9245	0.0665	-81.1710	-82.3628	-60.9911	-74.0264
14	27.0514	0.0665	-81.9857	-82.8293	-62.7064	-74.1292
15	29.1118	0.0665	-82.7110	-83.2899	-64.1736	-74.8299
16	0.0665	0.9305	-52.5708	-75.6925	-80.0944	-90.6581
17	2.1269	0.9305	-49.9084	-74.8049	-79.1051	-89.6676
18	4.1873	0.9305	-48.7099	-74.4776	-77.9974	-89.1670
19	6.2477	0.9305	-49.5547	-80.2277	-70.7424	-88.6604
20	8.3082	0.9305	-51.6369	-74.0706	-69.3024	-82.1533

Figure 4.11: Training set for locating the MS

Testing_data

450x6 double

	1	2	3	4	5	6	7
1	1.0634	0.0665	-52.0731	-75.3180	-79.7340	-90.0807	
2	3.1239	0.0665	-50.6871	-74.9841	-72.7185	-89.6037	
3	5.2508	0.0665	-50.5380	-74.7078	-71.5453	-89.1053	
4	7.3112	0.0665	-51.7484	-80.5210	-70.2620	-82.6217	
5	9.3716	0.0665	-53.5317	-74.4254	-68.8068	-82.1440	
6	11.4320	0.0665	-61.3279	-74.4270	-55.6603	-87.6815	
7	13.5589	0.0665	-63.0051	-74.5304	-59.2938	-81.2325	
8	15.6193	0.0665	-64.4367	-80.7227	-46.0427	-80.8401	
9	17.6798	0.0665	-65.6998	-74.9939	-50.4680	-74.5062	
10	19.7402	0.0665	-72.8216	-75.3297	-50.7800	-74.2481	
11	21.8671	0.0665	-73.8574	-75.7282	-52.2979	-80.0777	
12	23.9275	0.0665	-80.7628	-82.1499	-54.1238	-74.0171	
13	25.9879	0.0665	-81.5874	-82.5944	-61.8741	-74.0639	
14	28.0483	0.0665	-82.3439	-83.0515	-63.4396	-74.2146	
15	30.1752	0.0665	-83.0637	-84.0286	-64.8612	-74.9682	
16	1.0634	0.9305	-51.4893	-75.4985	-79.6290	-89.9222	
17	3.1239	0.9305	-49.1140	-74.6379	-78.5855	-89.4264	
18	5.2508	0.9305	-48.8983	-74.3379	-71.3701	-88.9060	
19	7.3112	0.9305	-50.5680	-80.1342	-70.0250	-82.3983	
20	9.3716	0.9305	-58.7857	-74.0295	-68.4717	-69.8939	

Figure 4.12: Testing data for the assessment of performance of coefficient

4.5.2.2. Evaluation: As coefficient set is obtained, computation of sum of location error is needed to evaluate the resulting performance of these coefficients, which is mostly denoted as cost, and is given below:

$$Cost = \sum_{i=1}^{m_1} \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \tag{4.7}$$

4.5.2.3. Optimization: For better accuracy, searching of new coefficient is done by SA, which utilizes the variable like temperature for cooling schedule.

$$p = e^{-\frac{\Delta Cost}{temperature}}, \qquad \Delta Cost \geq 0 \tag{4.8}$$

In this algorithm, 4 invervals are made to assign 4 different cooefficents or weights. Whole space of fingerprints is divided into forur intervals and inintially, the value of all four cooeficient as 1. After running the SA algorithm, it sets the values of cooeficents as given in the resuts shown below:

```
coeff_new =

    0.8643    0.8656    0.8645    0.8632
```

Once these weights are found and assigned to the intervals, FS-kNN model find the effective signal distance and calculate the estimated position of the target. In the result, given below, it is showed the resultant estimated position of the first random user and its actual position and calculated percentage error.

```
start FSKNN
The Estimate Position | X = 2.10,Y = 6.12
The Actual   Position | X = 2.73,Y = 6.51
Percentage Error = 0.74
--
```

The total positioning error of all random users is shown in figure as a CDF plot.

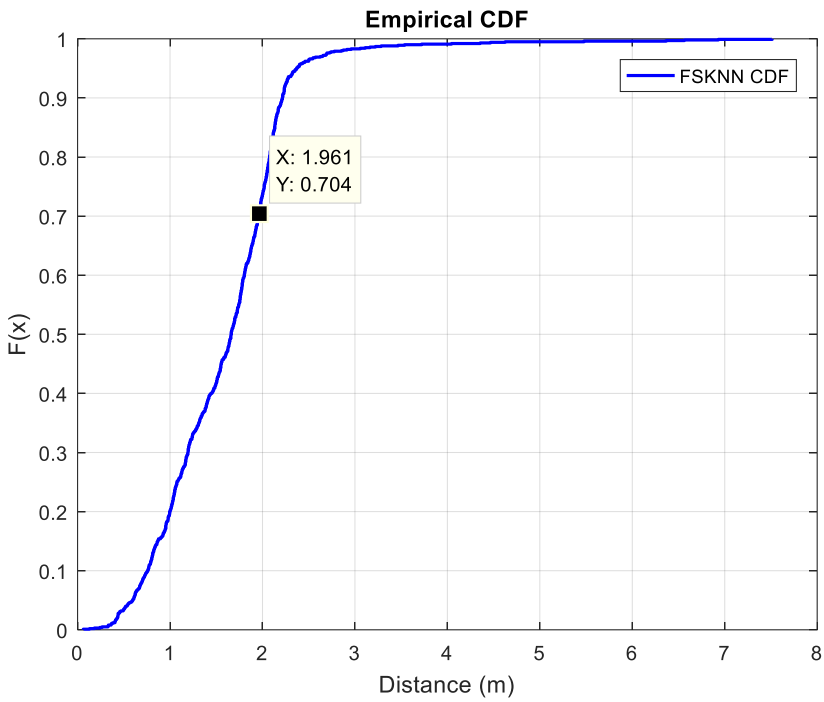


Figure 4.13: Total error CDF plot of FS-kNN

Which shows that 70% of users are located with positioning error with 1.9m

4.5.3. FS-kNN with deleting Outlier:

In this model starting procedure is the same as followed by the FS-kNN but at the point where k fingerprints are selected as nearest neighbors based on least effective distance. Once these fingerprints are extracted, the altered Thompson Tau test is performed to find and remove the outliers. That statistical decision can be made as:

$$\delta_x = \frac{|x_i - x'|}{std_x} \geq \tau, \qquad 4.9$$

$$\delta_y = \frac{|y_i - y'|}{std_y} \geq \tau, \tag{4.10}$$

Results after deleting the outlier from nearest neighbors is shown:

```
start FSKNNROL
The Estimate Position | X = 2.11,Y = 6.14
The Actual   Position | X = 2.73,Y = 6.51
Percentage Error = 0.72
```

As the outliers have been removed, the position is estimated from the remaining fingerprints based on least distance. The cumulative distribution function of positioning error by this model is given as:

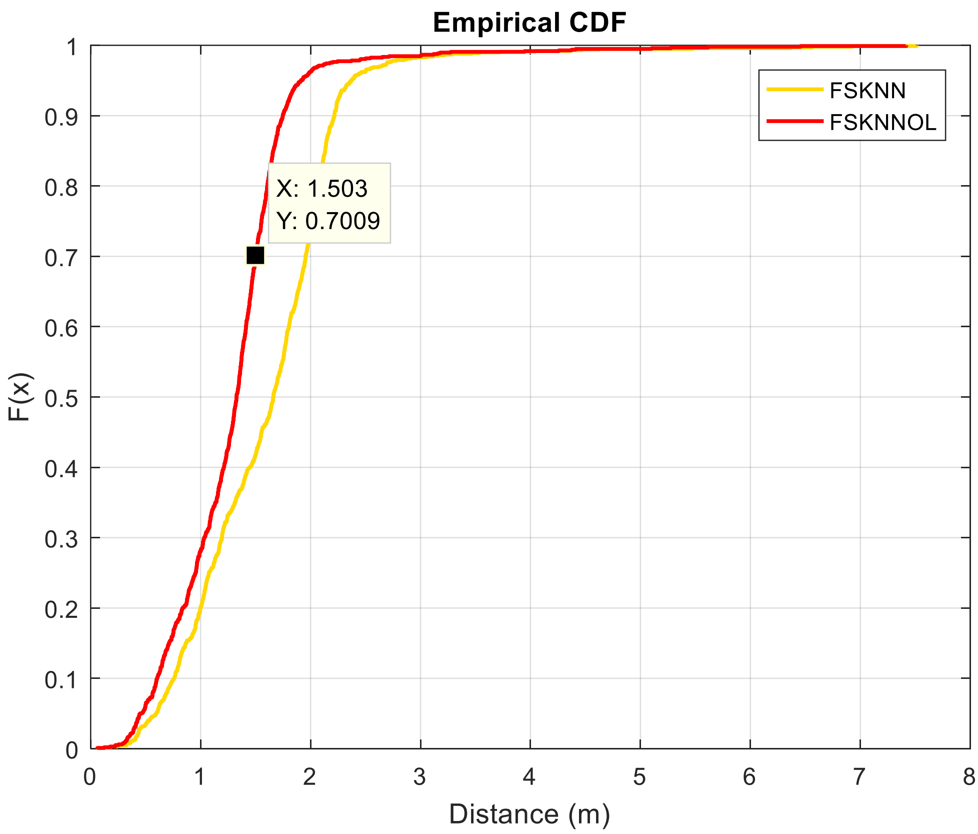


Figure 4.14: Total error CDF plot of FS-kNN with Deleting Outlier w.r.t. reference technique.

Above figure 4.14 shows that new model has improved the positioning errors from 1.9m to 1.47 m. From the above results the combined CDF plot of all three algorithms has taken to observe the most accurate model among all.

4.5.4. Comparison

After analyzing the results from three different models, here those results are compared to verify which technique is better than the others. So, the compared CDF plot of three models shows that feature scaling model with deleting outlier, gives the best results and have the maximum possible accuracy among three of the models. Then comes the FS-kNN matching algorithm which is in between the proposed FS-kNN with deleting outlier and kNN algorithms according to accuracy. And it shows that the matching process often leads to a geographically dispersed set of RPs, resulting in unsatisfactory estimation accuracy, so it sometimes causes unacceptable positioning error. And in the last kNN gave the results which are poorest among all since it cannot recognize the best possible solution because it doesn't verify the RSSI vectors according to their distance. It takes any RSSI vector matching to the online query without noting the distance differences.

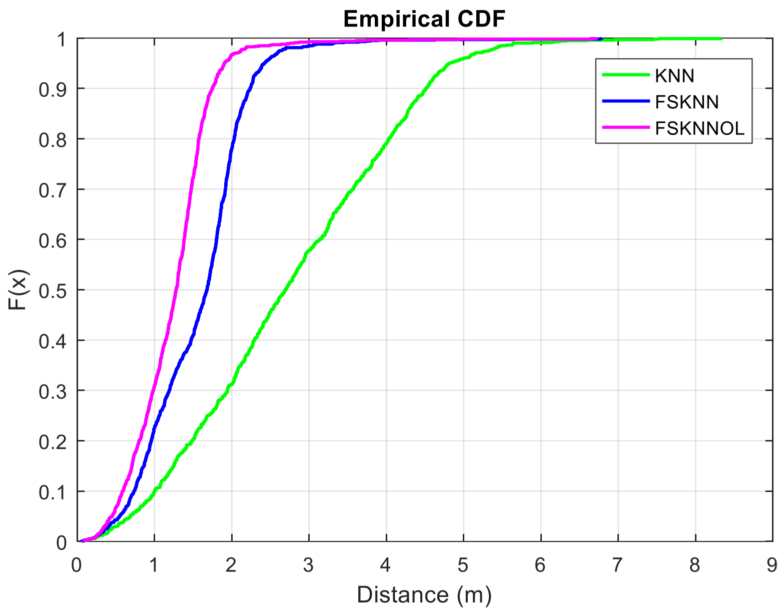


Figure 4.15: Comparison of CDF plots of 3 matching algorithms.

Algorithm	FS-kNN with deleting Outlier	FS-kNN	kNN
Mean Location error	1.503m	1.961	3.45

Table 4.2: Comparison of 3 matching algorithms.

4.5.4.1. Comparison with another algorithm

Here comparison with other variant of kNN i.e. Cosine similarity has been done to again analyze the performance of proposed model. Therefore, it can be said that proposed model is more accurate than the other existing models. Figure 4.16 shows the simulation results of the comparison.

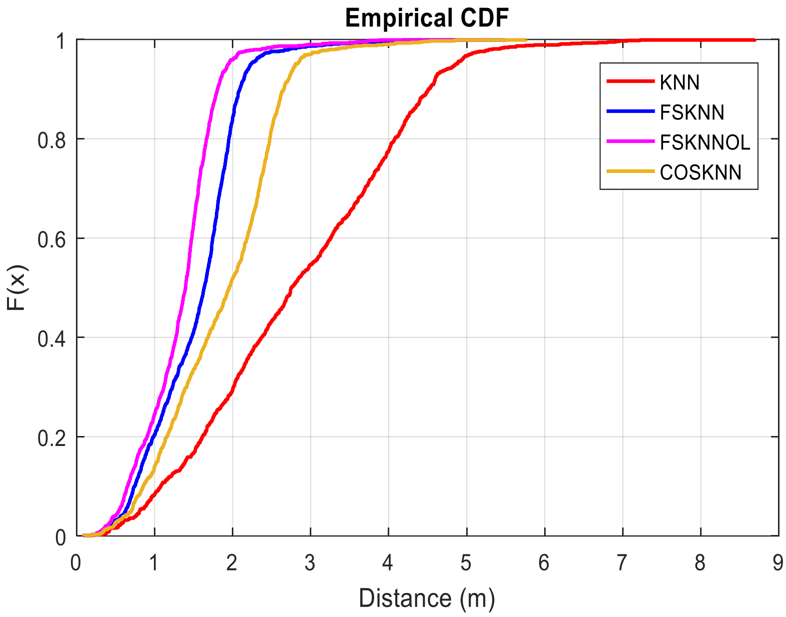


Figure 4.16: Comparison of CDF plots of 4 matching algorithms.

4.6. Performance analysis

In the following section, the consequences are examined by increasing or decreasing some basic parameters of the models, for example by changing different building structure with different blueprint size, by changing area, by changing no. of APs and by changing no. of RPs.

4.6.1. Different images with different sizes

In this case, model is proved as generic. Model can be implemented on blue prints of any building structure. A few changes in the code will make the model suitable for the new building structure. From the figure 4.17, it is shown that image is different in structure.

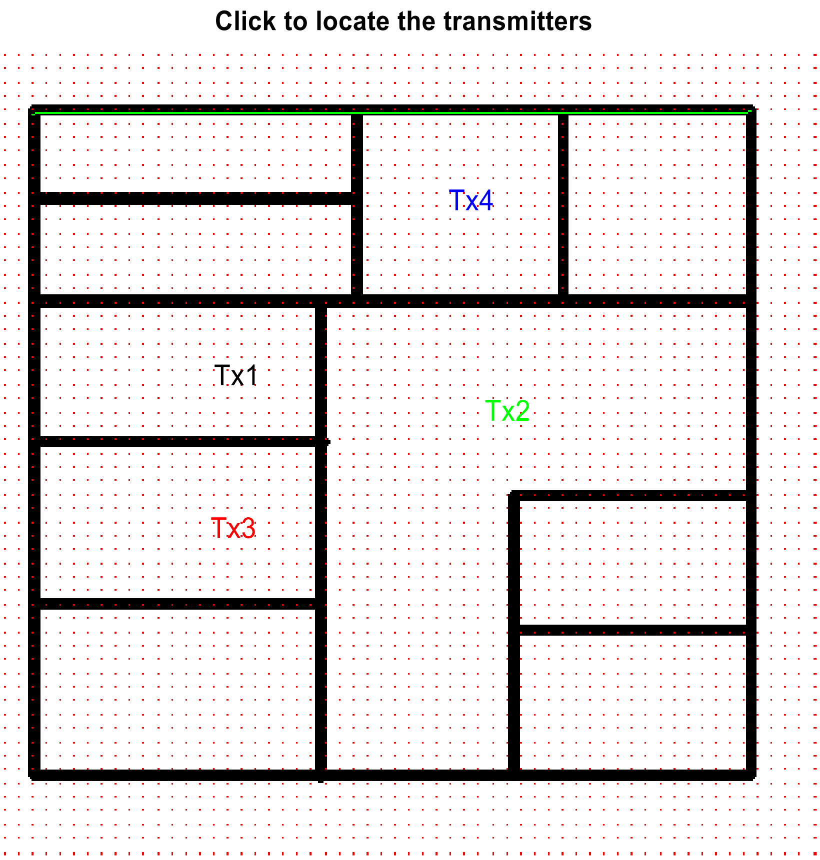


Figure 4.17: Case for changing the structure of the building.

The model not only works for different building structures but also it is generic for images with different sizes. If a blue print of a building has lower or higher no. of pixel from referenced image than small changes in the code regarding image size etc. will make the model appropriate for the new image. Figure 4.18 shows the results of 350-pixel image:

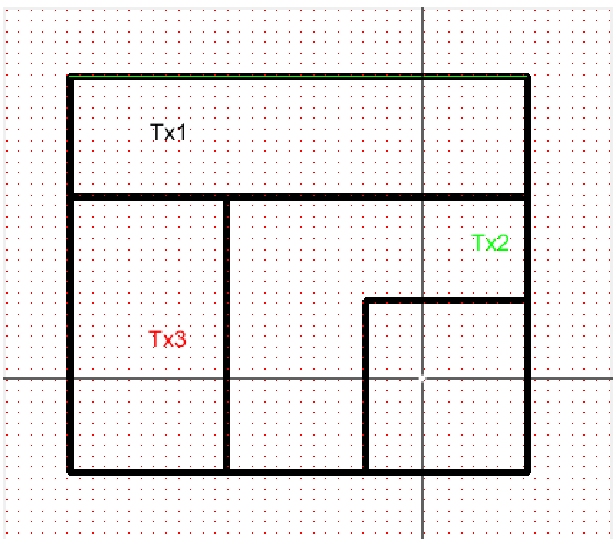


Figure 4.18: Case for decreasing the no. of pixels of the blue prints.

4.6.3. Number of APs

If I increase the number of access points in the same region then the location accuracy will be more enhanced, but the cost of the model will exceed to some extent as well. So, it is needed to present an optimal solution which can provide good accuracy with low budget. Figure 4.19 shows the consequences of the case.

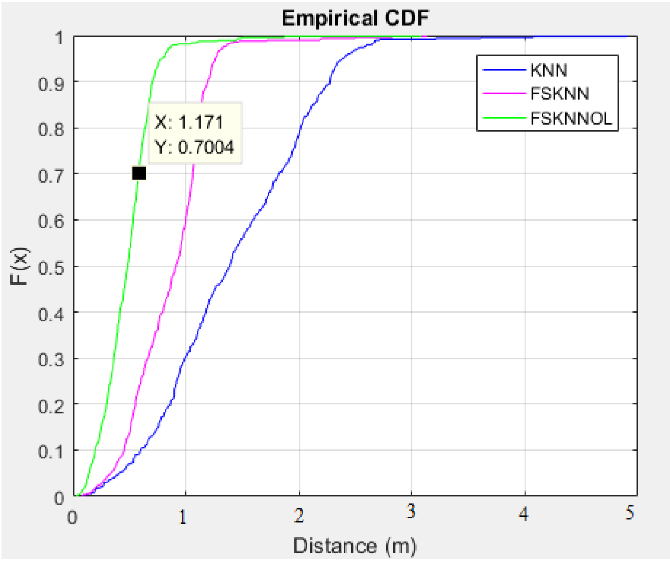
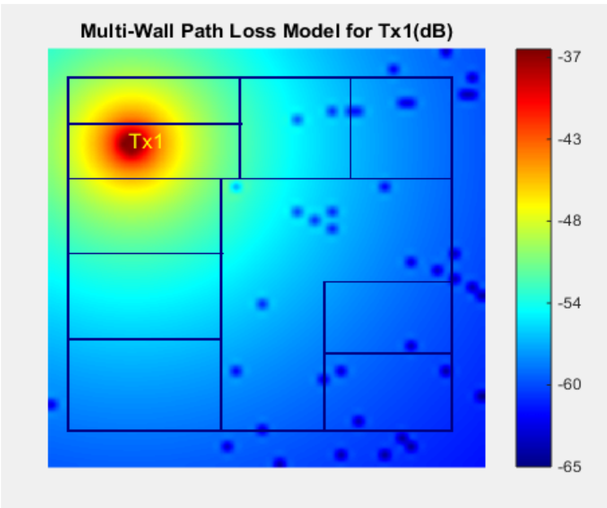


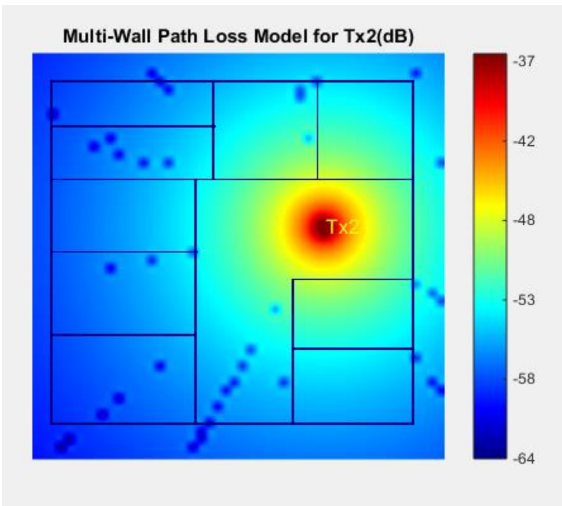
Figure 4.19: Case for increasing no. of APs.

4.6.5. Area

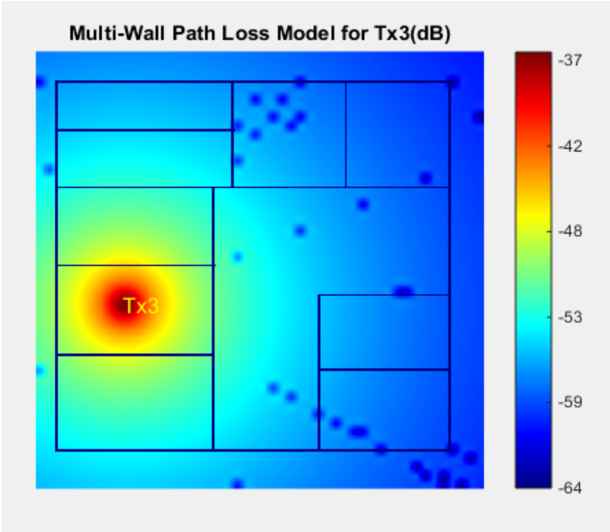
If area is increased, then it would be difficult to caver the whole space with strong AP RSSI values. Some of the area remain un covered because AP has low signal strength for the RPs which are away from the reference points. Therefor user standing away from the access points may not be facilitate with strong signal strength. In figure 4.20, it is cleared that by increasing the area, the signal strength is limited to small area and other is remained uncover.



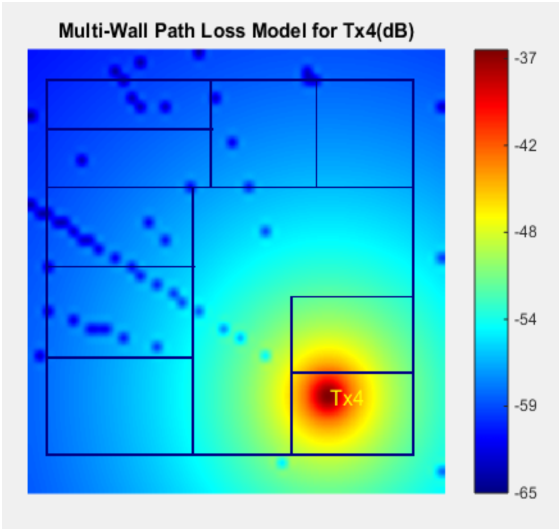
(a)



(b)



(c)



(d)

Figure 4.20: Case for increasing targeted area for same no. of APs.

4.6.6. No. of reference points 50/60/70

In this case, quantity of fingerprints is changed which are created in offline phase. It is done by changing the number of reference points from 50*50 to 60*60. By increasing the no. of RPs, the accuracy degrades to some extent because every user has more reference points around it and it is more confusing for the algorithms to identify that user is belonged which RP. Therefore, it is very important to choose number of reference points according to the total targeted area.

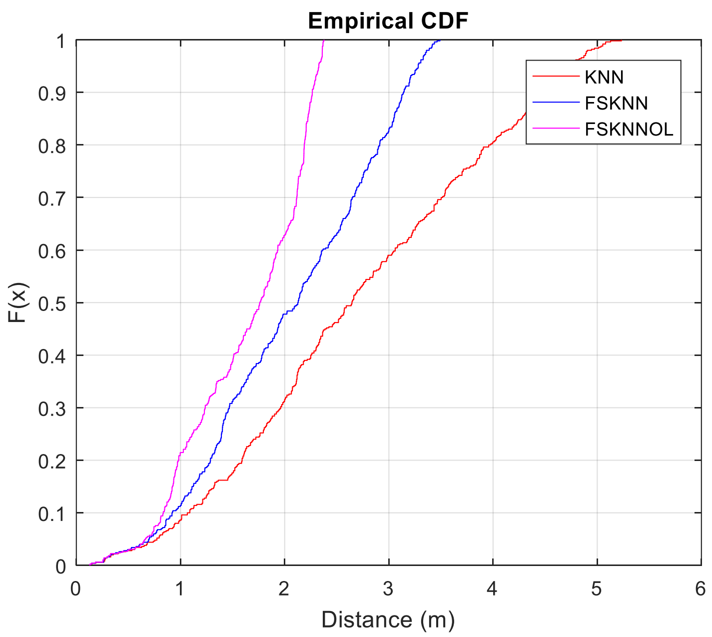


Figure 4.21: Case for increasing the no. of reference points.

Chapter 5 - Conclusion

In the previous chapter, simulation results of the model have been shown which was proposed in chapter 3. So, in this chapter results of proposed technique and possible future work are discussed.

Selection of blue print, floor plan meshing etc. has done for offline processing. Deployment of access points is very important in indoor localization, because with minimum number of APs, finding the appropriate positions of APs is necessary to achieve maximum accuracy. Therefore, number of simulations have been done to find the exact position of each AP for more efficiency. Therefore, one can apply this technique as a general method used for any building structure with manually access points deployment. Databases are created by using Motley Keenan model, it not only helps in finding the accurate path loss but also uses the simple 2D blueprint. This model gives the detailed information of each reference point separately and did not use a general term of wall factor for all RPs. The main advantage of using Motely Keenan propagation multiwall model is that it not only counts the total walls between transceivers but also examines the thickness of each wall with the help of angle of incident and polarization. Hough transmission is used to distinguish the walls from the image.

In wireless framework, verification of signal strength is very important to know that the area with strong strength and weak signal strength. In this case, propagation map helps to verify the signal strength of all APs in the targeted area. It was ensured APs must have covered the whole area under discussion.

For online results, some points are concluded about the matching algorithms that provided better results by using proposed method. Starting from the proposed matching algorithm which is FS-kNN with deleting outlier along with using new propagation model gives the positioning accuracy with 70% of results with 1.45 m of error. Whereas other two matching algorithms have also improved due to effective propagation model. FS-kNN with strong databases has improved and gives positioning errors up to 1.6m for 70% users from 1.75m. Moreover, kNN algorithm improved from 3.9m to 3.6m for 70% users. Hence, Positioning results affected by propagation model are also improved and accuracy of localization has increased up to 9% and 8% using the FS-kNN and kNN matching algorithm respectively. In other words, proposed technique has removed 20% of positioning errors from the referenced technique. This technique localizes 1000 random users in 60 seconds comparatively in larger area with less APs.

For the future, the proposed technique can be applied on most recent algorithms to improve the positioning accuracy. For indoor localization, various indoor positioning systems with different techniques would be fuse with each for all type of indoor environments. In the existing matching algorithms, spearman distance and Manhattan distance would be used as well, instead of the Euclidean distance for reducing positioning errors.

Chapter 6 - References

[1] Ville Honkavirta, Tommi Perala, Simo Ali Loytty, Robert Piche. "A comparative survey of WLAN location fingerprinting methods", 2009 6th Workshop on Positioning, Navigation and Communication, 2009

[2] Z. Li, C. Liu, J. Gao, and X. Li, “An Improved WiFi / PDR Integrated System Using an Adaptive and Robust Filter for Indoor Localization,” ISPRS International Journal of Geo-Information, vol. 5, no. 12, 2016.

[3] Q. Chang, S. Van de Velde, W. Wang, Q. Li, H. Hou, and H. Steendam, “WiFi Fingerprint Positioning Updated by Pedestrian Dead Reckoning for Mobile Phone Indoor Localization,” Lecture Notes in Electrical Engineering, vol. 3, no. 342, pp. 729–739, 2015.

[4] R. Want, A. Hopper, V. Falcão, and J. Gibbons, “The active badge location system,” ACM Transactions on Information Systems, vol. 10, no. 1, pp. 91–102, 1992.

[5] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, “Indoor location sensing using geo-magnetism,” in Proceedings of the 9th international conference on Mobile systems, applications, and services – MobiSys ’11, 2011, p. 141. [Online]. Available: http://portal.acm.org/citation.cfm?doid= 1999995.2000010

[6] H. Xie, T. Gu, X. Tao, H. Ye, and J. Lv, “MaLoc: A practical magnetic fingerprinting approach to indoor localization using smartphones,” in UbiComp ’14 Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2014, pp. 243–253.

[7] K. Liu, X. Liu, and X. Li, “Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors,” IEEE Transactions on Mobile Computing, vol. 15, no. 5, pp. 1144–1156, 2016.

[8] a. Mandal, C. Lopes, T. Givargis, A. Haghighat, R. Jurdak, and P. Baldi, “Beep: 3D indoor positioning using audible sound,” in Second IEEE Consumer Communications and Networking Conference, 2005. CCNC. 2005, 2005, pp. 348– 353.

[9] V. Radu and M. K. Marina, “HiMLoc: Indoor smartphone localization via activity aware pedestrian dead reckoning with selective crowdsourced WiFi fingerprinting,” in 2013 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2013, no. October, 2013, pp. 28–31.

[10] Bao, Haitao, and Wai-Choong Wong. "An indoor dead-reckoning algorithm with map matching." *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*. IEEE, 2013.

[11] R. Ivanov. Indoor navigation system for visually impaired. In Proceedings of the 11th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing on International Conference on Computer Systems and Technologies, pages 143–149. ACM, 2010.

[12] R. Mautz. Indoor positioning technologies. PhD thesis, Habil. ETH Zurich, 2012, 2012.¨

[13] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in Proceedings. 1987 IEEE International Conference on Robotics and Automation, vol. 4, 1990, pp. 850–850. [Online]. Available: http://ieeexplore.ieee.org/document/1087846/
[14] J. J. L. Durrant and H. F. Whyte, "Mobile robot localization by tracking geometric beacons.pdf," in IEEE/RSJ International Conference on Intelligent Robot Systems, 1989, pp. 376–382.
[15] N. Karlsson, E. Di Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, and M. E. Munich, "The vSLAM algorithm for robust localization and mapping," in Proceedings - IEEE International Conference on Robotics and Automation, vol. 2005, no. April, 2005, pp. 24–29.
[16] U. Bandara, M. Hasegawa, M. Inoue, H. Morikawa, and T. Aoyama, "Design and implementation of a Bluetooth signal strength based location sensing system," in Proceedings. 2004 IEEE Radio and Wireless Conference (IEEE Cat. No.04TH8746), no. October, 2004, pp. 3–4.
[17] M.-s. Choi and B. Jang, "An Accurate Fingerprinting based Indoor Positioning Algorithm," International Journal of Applied Engineering Research, vol. 12, no. 1, pp. 86–90, 2017.
[18] S. Yang, P. Dessai, M. Verma, and M. Gerla, "FreeLoc : Calibration-Free Crowdsourced Indoor Localization," in 2013 Proceedings IEEE INFOCOM, 2013, pp. 2481–2489.
[19] C. Wu, S. Member, and Z. Yang, "WILL : Wireless Indoor Localization without Site Survey," in IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 4, 2013, pp. 839–848.
[20] H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37(6):1067–1080, 2007.
[21] M. S. Svalastog. Indoor positioning-technologies, services and architectures. 2007.
[22] J. Korhonen, T. Ojala, M. Klemola, and P. Väänänen, "mTag-Architecture for Discovering Location Specific Mobile Web Services Using RFID and Its Evaluation with Two Case Studies," in Advanced Int'l Conference on Telecommunications and Int'l Conference on Internet and Web Applications and Services (AICT-ICIW'06), 2006, pp. 191–200. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=1602324
[23] Z. Song, G. Jiang, and C. Huang. A survey on indoor positioning technologies. In Theoretical and Mathematical Foundations of Computer Science, pages 198–206. Springer, 2011.
[24] N. A. Alsindi, B. Alavi, and K. Pahlavan, "Measurement and modeling of ultrawideband TOA-based ranging in indoor multipath environments," IEEE Transactions on Vehicular Technology, vol. 58, no. 3, pp. 1046–1058, 2009.
[29] M. S. Svalastog. Indoor positioning-technologies, services and architectures. 2007.
[25] DraMCo, "Nauwkeurigheid van dynamische RSS-based indoor plaatsbepalingssystemen," pp. 1–11, 2008. [Online]. Available: https: //iiw.kuleuven.be/onderzoek/dramco/research/ladi-witepa/Technieken.pdf

[26] M. B. Nabil Ali Alrajeh and B. Shams2, “Localization Techniques in Wireless Sensor Networks,” *International Journal of Distributed Sensor Networks*, vol. 2013, pp. 1–9, 2013.

[27] M. Sauter, “3.7.1 Mobility Management in the Cell-DCH State,” in From GSM to LTE: An Introduction to Mobile Networks and Mobile Broadband, 2010, p. 160.

[28] Cisco, “Location Tracking Approaches,” in *Wi-Fi Location-Based Services 4.1 Design Guide*, 2014.

[29] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, “Wireless RSSI fingerprinting localization,” *Signal Processing*, vol. 2017, no. 131, pp. 235–244, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2016.07.005

[30] K. Kaemarungsi and P. Krishnamurthy, “Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting,” in *Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous’04)*, 2004. [Online]. Available: https: //pdfs.semanticscholar.org/c87f/ca49f8e3dacbcd5ab11261508aa79a03ca8a.pdf

[31] Y. Chapre, A. Ignjatovic, A. Seneviratne, and S. Jha, “CSI-MIMO : Indoor WiFi fingerprinting system,” in *39th Annual IEEE Conference on Local Computer Networks*, no. January 2016, 2014, pp. 202–209.

[32] H. Wang, A. Elgohary, and R. R. Choudhury, “No Need to War-Drive : Unsupervised Indoor Localization,” in *Proceedings of the 10th international conference on Mobile systems, applications, and services (MobiSys ’12)*, 2012, pp. 197–210.

[33] P. Bahl, V. Padmanabhan, and A. Balachandran, “RADAR: An in-building RF based user location and tracking system,” in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. 2, no. c, 2000, pp. 775–784. [Online]. Available: http://research.microsoft.com/en-us/groups/sn-res/infocom2000.pdf

[34] K. D’hoe, G. Ottoy, P. Keersebilck, J. P. Goemaere, and L. De Strycker, “Indoor room location estimation,” *Advances in Electrical and Computer Engineering*, vol. 8, no. 2, pp. 78–81, 2008.

[35] Y. Gu, A. Lo, and I. Niemegeers. A survey of indoor positioning systems for wireless personal networks. *Communications Surveys & Tutorials, IEEE*, 11(1):13–32, 2009.

[36] J. Hightower and G. Borriello. Location sensing techniques. Technical report, IEEE Computer, 2001.

[37] M. Shchekotov, “Indoor Localization Method Based on Wi-Fi Trilateration Technique,” in *Proceeding of the 16Th Conference of Fruct Association*, 2014, pp. 177–179. [Online]. Available: https://fruct.org/publications/abstract16/ files/Shc1.pdf

[38] Z. Farid, R. Nordin, and M. Ismail, “Recent Advances in Wireless Indoor Localization Techniques and System,” *Journal of Computer Networks and Communications*, vol. 2013, p. 12, 2013.

[39] A. Taparugssanagorn, S. Siwamogsatham, C. Pomalaza-R, x00E, and ez, "A Hexagonal Coverage LED-ID Indoor Positioning Based on TDOA with Extended Kalman Filter," in Computer Software and Applications Conference (COMPSAC), 2013 IEEE 37th Annual, 2013, pp. 742-747.

[40] D. Trong-Hop, H. Junho, and Y. Myungsik, "TDoA based indoor visible light positioning systems," in 2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN), 2013, pp. 456-458.

[41] P. Lou, H. Zhang, X. Zhang, M. Yao, and Z. Xu, "Fundamental analysis for indoor visible light positioning system," in 2012 1st IEEE International Conference on Communications in China Workshops (ICCC), 2012.

[42] Z. Yuxin, Y. Feng, F. Gunnarsson, M. Amirijoo, E. Ozkan, and F. Gustafsson, "Particle filtering for positioning based on proximity reports," in Information Fusion (Fusion), 2015 18th International Conference on, 2015, pp. 1046-1052.

[43] Y. Feng, Z. Yuxin, and F. Gunnarsson, "Proximity report triggering threshold optimization for network-based indoor positioning," in Information Fusion (Fusion), 2015 18th International Conference on, 2015, pp. 1061-1069.

[44] Kul, Gökhan, Tansel Özyer, and Bülent Tavli. "IEEE 802.11 WLAN based real time indoor
positioning: Literature survey and experimental investigations." *Procedia Computer Science*34 (2014): 157-164.

[45] Khodayari, Shahrzad, Mina Maleki, and Elham Hamedi. "A RSS-based fingerprinting method for positioning based on historical data." *Performance Evaluation of Computer and Telecommunication Systems (SPECTS), 2010 International Symposium on*. IEEE, 2010.

[46] X. C. a. Y. D. Peerapong Torteeka, "Hybrid Technique for Indoor Positioning System based on Wi-Fi Received Signal Strength Indication," in *2014 International Conference on Indoor Positioning and Indoor Navigation*, 27th-30th October, 2014.

[47] N. G. D. A. Ali Khalajmehrabadi, "Modern WLAN Fingerprinting Indoor Positioning Methods and Deployment Challenges," *IEEE COMMUNICATIONS SURVEYS & TUTORIALS*, 2016.

[48] G. Kail, P. Maechler, N. Preyss, and A. Burg, "Robust asynchronous indoor localization using LED lighting," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 1866-1870

[49] Youssef, Moustafa, and Ashok Agrawala. "The Horus WLAN Location Determination System."*Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services - MobiSys '05* (2005)

[91] E. C. C. E. w. t. E. C. o. P. (CEPT) and T. Administrations, "The European table of frequency allocations and applications in the frequency range 8.3 kHz to 3000 GHz (ECA Table)," 2016. [Online]. Available: http://www.erodocdb.dk/docs/doc98/official/pdf/ERCRep025.pdf

[52] W. Wong, J. Ng, and W. Yeung, "Wireless LAN Positioning with Mobile Devices in a Library Environment," in *25th IEEE International Conference on Distributed Computing Systems Workshops*, 2005, pp. 633–636. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1437236

[53] W. Xue, W. Qiu, X. Hua, and K. Yu, "Improved Wi-Fi RSSI Measurement for Indoor Localization," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 1–1, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7835628/

[54] V. Moghtadaiee and A. G. Dempster, “Design protocol and performance analysis of indoor fingerprinting positioning systems,” *Physical Communication*, vol. 13, no. PA, pp. 17–30, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.phycom.2014.02.004

[55] P. Bahl, V. Padmanabhan, and A. Balachandran, “A software system for locating mobile users: Design, evaluation, and lessons,” 2000. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.43.3784{&}rep=rep1{&}type=pdf

[56] A. Alhamoud, M. Kreger, H. Afifi, C. Gottron, D. Burgstahler, F. Englert, D. Böhnstedt and R. Steinmetz, "Empirical investigation of the effect of the door's state on received signal strength in indoor environments at 2.4 GHz," in *Proc. IEEE 39th Conf. Local Computer Networks Workshops (LCN Workshops),* 2014. 4- A. Motley, J. Keenan, "Radio Coverage in Buildings," *British Telecom Tech. J,* vol. 8, no. 1, pp. 19-24, 1990.

[57] L. Li, Y. Ibdah, Y. Ding, H. Eghbali, S. H. Muhaidat and X. Ma, "Indoor Multi-wall Path Loss Model at 1.93 GHz," in *Proc. MILCOM 2013 – 2013 IEEE Military Communications Conf*, 2013.

[58] Z. Wu, K. Fu, E. Jedari, S. R. Shuvra, R. Rashidzadeh and M. Saif, “A Fast and Resource Efficient Method for Indoor Positioning Using Received Signal Strength,” in IEEE Transactions on Vehicular Technology, vol. 65, no. 12, pp. 9747-9758, Dec. 2016.

[59] Li, Dong, Baoxian Zhang, and Cheng Li. "A Feature-Scaling-Based $ k $-Nearest Neighbor Algorithm for Indoor Positioning Systems." *IEEE Internet of Things Journal* 3.4 (2016): 590-597.

[60] K. Al Nuaimi and H. Kamel. A survey of indoor positioning systems and algorithms. In *Innovations in Information Technology (IIT), 2011 International Conference on*, pages 185–190. IEEE, 2011

[61] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha, “FM-based indoor localization,” in Proc. 10th Int. Conf. Mobile Syst. Appl. Serv. (MobiSys’12), Jun. 2012, pp. 169–182.

[62] S. Han, C. Zhao, W. Meng and C. Li, “Cosine similarity based fingerprinting algorithm in WLAN indoor positioning against device diversity,” 2015 IEEE International Conference on Communications (ICC), London, 2015, pp. 2710-2714.

[63] Y. A. Sekercioggl, J. Violi, L. Priestnall, and J. Armstrong, “Accurate node localization with directional pulsed infrared light for indoor ad hoc network applications,” in Proc. 22nd Int. Conf. Telecommun. (ICT’15), Apr. 2015, pp. 384–390.

[64] V. Filonenko, C. Cullen, and J. Carswell, “Investigating ultrasonic positioning on mobile phones,” in Proc. Int. Conf. Indoor Positioning Indoor Navig. (IPIN’10), Sep. 2010, pp. 1–8.

[65] L. Jing, Z. X. Cheng, Y. H. Zhou, and J. B. Wang, “From U-tile to indoortracer: An indoor location sensing platform based on passive RFID,” in Proc. 4th Int. Conf. Cyber Phys. Social Comput. Internet Things (CPSCom’11), Oct. 2011, pp. 89–93.

[66] H. Zou, H. Jiang, X. X. Lu, and L. H. Xie, “An online sequential extreme learning machine approach to WiFi based indoor positioning,” in Proc. IEEE World Forum Internet Things (WF-IoT’14), Mar. 2014, pp. 111–116.
[67] H. Buyruk et al., “RF fingerprinting based GSM indoor localization,” in Proc. 21st Signal Process. Commun. Appl. Conf. (SIU’13), Apr. 2013, pp. 1–4.
[68] V. Moghtadaiee, A. G. Dempster, and S. Lim, “Indoor localization using FM radio signals: A fingerprinting approach,” in Proc. Int. Conf. Indoor Positioning Indoor Navig. (IPIN’11), Sep. 2011, pp. 1–7.
[69] S. P. Tarzia, P. A. Dinda, R. P. Dick, and G. Memik, “Indoor localization without infrastructure using the acoustic background spectrum,” in Proc. 9th Int. Conf. Mobile Syst. Appl. Serv. (MobiSys’11), Jun. 2011, pp. 155–168.
[70] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, “Indoor location sensing using geo-magnetism,” in Proc. 9th Int. Conf. Mobile Syst. Appl. Serv. (MobiSys’11), Jun. 2011, pp. 141–154.
[71] L. Kanaris, A. Kokkinis, G. Fortino, A. Liotta, S. Stavrou, “Sample Size Determination Algorithm for fingerprint-based indoor localization systems.” Computer Networks, pp. 169-177, 2016.
[72] S. He and S. H. G. Chan, “Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons,” in IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 466-490, Firstquarter 2016.
[73] "Propagation data and prediction methods for the planning of indoor radio communication systems and the radio local area networks in the frequency range 900 mhz to 100 ghz," Recommendation ITU-R P.1238- 8, 2015.
[74] COST Action 231: Digital Mobile Radio Towards Future Generation Systems: Final Report, 1999
[75] A. G. M. Lima and L. F. Menezes, "Motley-Keenan model adjusted to the thickness of the wall," in SBMO/IEEE MTT-S International Conference on Microwave and Optoelectronics, 2005. , 2005.
[76] A. Motley, J. Keenan, "Radio Coverage in Buildings," *British Telecom Tech. J,* vol. 8, no. 1, pp. 19-24, 1990.
[77] A. J. Motley and J. M. P. Keenan, "Personal communication radio coverage in buildings at 900 MHz and 1700 MHz," *Electronics Letters,* vol. 24, pp. 763-764, Jun 1988.
[78] P. Pechač and M. Klepal, "Empirical models for indoor propagation in CTU Prague buildings," *Radioengineering,* vol. 9, pp. 31-36, 2000.
[79] A. G. M. Lima and L. F. Menezes, "Motley-Keenan model adjusted to the thickness of the wall," in *Proc. SBMO/IEEE MTT-S Int. Conf. Microwave and Optoelectronics,* 2005.
[80] A. I. Sulaiman and M. A. Hussein, "A modified multi-wall wave propagation model for concrete based building structure," in *Proc. Int. Conf. Computer and Communication Engineering (ICCCE),* 2012.
[81] K.-W. Cheung, J. H. M. Sau and R. D. Murch, "A new empirical model for indoor propagation prediction," *IEEE Transactions on Vehicular Technology,* vol. 47, pp. 996-1001, Aug 1998.

[82] A. Durantini and D. Cassioli, "A multi-wall path loss model for indoor UWB propagation," in *Proc. IEEE 61st Vehicular Technology Conf*, 2005.
[83] M. Lott and I. Forkel, "A multi-wall-and-floor model for indoor radio propagation," in *Proc. (Cat. No.01CH37202) IEEE VTS 53rd Vehicular Technology Conf. Spring 2001*, 2001.
[84] L. Li, Y. Ibdah, Y. Ding, H. Eghbali, S. H. Muhaidat and X. Ma, "Indoor Multi-wall Path Loss Model at 1.93 GHz," in *Proc. MILCOM 2013 – 2013 IEEE Military Communications Conf*, 2013.
[85] H. Zou, Y. Zhou, H. Jiang, B. Huang, L. Xie and C. Spanos, "Adaptive Localization in Dynamic Indoor Environments by Transfer Kernel Learning," 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, 2017, pp. 1-6.
[86] S. Bartoletti, A. Giorgetti, M. Z. Win and A. Conti, "Blind Selection of Representative Observations for Sensor Radar Networks," in IEEE Transactions on Vehicular Technology, vol. 64, no. 4, pp. 1388-1400, April 2015.
[87] S. Bartoletti, A. Conti and M. Z. Win, "Device-Free Counting via Wideband Signals," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 5, pp. 1163-1174, May 2017.
[88] D. Pastina, F. Colone, T. Martelli and P. Falcone, "Parasitic Exploitation of Wi-Fi Signals for Indoor Radar Surveillance," in IEEE Transactions on Vehicular Technology, vol. 64, no. 4, pp. 1401-1415, April 2015.
[89] K. Chetty, G. Smith, H. Guo and K. Woodbridge, "Target detection in high clutter using passive bistatic WiFi radar," 2009 IEEE Radar Conference, Pasadena, CA, 2009, pp. 1-5.
[90] S. Bartoletti, A. Conti, A. Giorgetti and M. Z. Win, "Sensor Radar Networks for Indoor Tracking," in IEEE Wireless Communications Letters, vol. 3, no. 2, pp. 157-160, April 2014.
[91] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, 2000, vol.2, pp. 775-784.
[92] Z. Wu, K. Fu, E. Jedari, S. R. Shuvra, R. Rashidzadeh and M. Saif, "A Fast and Resource Efficient Method for Indoor Positioning Using Received Signal Strength," in IEEE Transactions on Vehicular Technology, vol. 65, no. 12, pp. 9747-9758, Dec. 2016.