OPTIMISATION OF ACCESS POINT DEPLOYMENT MODEL USING USER-TO-AP POSITION VECTORS IN A WI-FI IPS

By

Sahar Batool

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

MARCH 2018

ABSTRACT

Access Point deployment has been a course of major study over the years and numerous researchers have worked upon it. This thesis work focuses on an intelligent deployment of Wireless Access Points (WAP) in an Indoor Wi-Fi Localisation System. Novel in its regard, the work first uses Geometric Dilution of Precision and the Genetic Algorithm to deploy Access Points, followed by a selection of Access Points with the help of RSSI and user-to-AP Position vectors. A series of MATLAB simulations showed that the user-to-AP position vector approach showed an improvement in localisation in the deployment area under study. Further work includes a performance analysis in which simulation parameters are altered to check their effect on system accuracy. Overall, a position vector based selection of Access Points by the users have shown the best results in the CDF plots with better accuracy.

DEDICATION

This thesis is dedicated to

MY BELOVED PARENTS, SISTERS,

AND HONORABLE TEACHERS

for their love, endless support and encouragement

ACKNOWLEDGEMENTS

I am grateful to Almighty Allah, the Creator of this Universe for making me able to complete this research work.

A bundle of thanks to my parents and sisters for their love and support in making this work possible.

I would like to thank my thesis supervisor, Dr Mir Yasir Umair for providing me with intricate knowledge and sincere guidance throughout my research work.

Acknowledging the Department of Electrical Engineering at Military College of Signals, NUST for equipping the students with an environment fit for productive research and overall character building.

TABLE OF CONTENTS

LIST OF FIGURES

ACRONYMS

DOP - Dilution of Precision

GA - Genetic Algorithm

HDOP - Horizontal Dilution of Precision

ILBS - Indoor Location-Based Systems

IPS - Indoor Positioning System

KNN - K Nearest Neighbours

LOS - Line of Sight

NLOS - Non- Line of Sight

PDA - Personal Digital Assistant

PDOP - Position Dilution of Precision

RTLS - Real Time Location Systems

TDOP - Time Dilution of Precision

VDOP - Vertical Dilution of Precision

Chapter 1

INTRODUCTION

Accurate location estimation has been the concern of many technological solutions of today, and finding the correct location of a mobile or stationary user has become as inevitable as finding The Mask. Till date, a variety of location estimation techniques and applications have been developed, in which a record of a target's locations is kept to determine its next incremental location or in other words, the target is localised in a given environment.

For many years scientists and technologists have been researching upon finding technological solutions for the objects location problem on the lines of low cost and complexity. But as such a simple and easy to understand yet focal solution is yet to be developed. Researches on indoor localisation systems are based uopn the usage of short-range signals, like WiFi [4], [5], [27], Bluetooth [28], ultrasound [29], and infrared [30].

This thesis focuses on a novel idea of position vectors which essentially addresses the user localisation problem. In this solution the users can be viewed as cooperative targets because of a two-step localisation process. In the first step the object is localised using K-nearest neighbour algorithm and then in the next step the localised users' positions calibrated using reference points on the grid map of the coverage area are fixed as the first distance measurement points for the position vectors. The second set of points are the Access Point locations and the position vector formation is similar to Euclidean distance measurement.

Location systems work differently for outdoor and indoor environments, but the underlying principle is the same. Wi-Fi based localisation systems best work indoors, i.e. they are most suited for users targeted in indoor environments where GPS uniquely cannot provide their location precisely.

An optimization model is proposed in [1] for the deployment of additional APs in an existing Wi-Fi IPS (or Indoor Positioning System). Past works regarding location estimation are focused on finding suitable locations for satellites in GPS satellite location systems, in particular asserting that the volume of the polyhedron formed by the user-to-satellite unit vectors is related to the candidate positions of those satellites which when selected can result in reducing error in localisation. Other works show that the unit vectors formed from the user to the satellites are actually the coefficients of a GDOP matrix, to be constructed later on.

The variance in user position vectors for the performance assessment of Cricket Indoor Positioning System has been studied, but the position vectors hence determined have not been applied to improve the system efficiency. This research work combines a Wi-Fi Access Point deployment model with an RSSI and user-to-Access Point position vector based AP selection to reduce the error in indoor location system.

1.1 Motivation

Indoor location systems have gained popularity in recent years. These localisation systems have provided a new layer of automation called automatic object location detection. [9] There are many real-world applications depending on such automation. For example, one can consider finding the location of products kept in a warehouse, of medical personnel or equipment in a hospital, locations of firemen in a building, locating police dogs who are trained to find explosives in a building, and keeping tags on the maintenance tools and equipment scattered over industrial plants. During the last ten years, the primary progress in indoor location sensing systems has been made. [9]
It is a key challenge to design an accurate indoor positioning system which can be easily deployed on commercially available mobile devices without any hardware modification. [17] Indoor localisation using triangulation cannot predict the actual location of a user due to disrupted line-of-sight paths between the radio transmitters and the receiver. It is also susceptible to errors, making the technique unsuitable for congested

indoor environments like shopping malls, universities, offices, public hospitals and airports.

It is difficult to find a line-of-sight path between the transmitter and the receiver in indoor environment. Radio propagation in these environments often suffers from multipath and shadowing. The time and angle of an arrival signal can be affected by the multipath effect; thus the estimated location could be less accurate. There is an alternative approach in which the distance of the mobile user is estimated from a set of measuring units, using the attenuation of emitted signal strength. These signal strength attenuation-based methods actually calculate the signal path loss due to propagation in a cluttered indoor environment .

RSSI fingerprinting method does not require an LOS path since it records the received signal strengths, called the Received Signal Strength Indicator (RSSI), from the Wi-Fi Access Points (APs).The receive signal strength indicator (RSSI) is a parameter which traditionally has a value starting from 0 to $RSSI_{max}$. It is measured by the physical sublayer of the energy observed at the receiving antenna. [10] User-to-AP position vector-based optimisation when used in conjunction with RSSI based localisation in an indoor positioning environment can help to improve the system accuracy. A two-way approach as such can cater for the variations in the indoor environment e.g. obstacles of different types, moving objects and wall obstructions. The position vector-based selection of APs can be used instead of RSSI in case of availability of line-of-sight path between a user and an AP, justified by the RSSI levels exceeding a predefined threshold.

Although a standard for indoor positioning systems is not available as of today, the wireless local area network (WLAN) infrastructure already deployed in buildings is the most extensively used indoor positioning system technology. [16]

1.2 Goals and Objectives

Following are the objectives of this research:

1. Construction of Genetic Algorithm (GA) for deployment of additional APs.

2. Selection of APs based on RSSI and user-to-AP position vectors.

3. K-Nearest neighbour algorithm for user localisation using the APs selected based on first the RSSI and then on position vectors.

1.3 Relevance to National needs

Indoor Positioning using GDOP is suitable wherever an accurate location estimate is required whether it is an academic, industrial or military application. Using a novel approach can improve the efficiency of already deployed systems.

1.4 Advantages

a) Ability to provide coordinate information instantaneously.

b) Integration with machine guidance systems.

c) Low deployment and operational costs.

d) Low complexity algorithms.

1.5 Areas of Application

a) Passenger navigation systems in airports, bus, train, and subway stations and museums.

b) Health care systems requiring a location estimate for patient mobility.

c) Academic institutes and shopping malls.

d) As a key assistance in robotics, and protecting expensive assets and vehicles from theft, loss or unauthorised use.

e) In tracking and geo-fencing systems to ensure that machines are not taken outside of a project or mining lease boundary.

1.6 Location System Applications

Some of the real world applications of indoor location systems include the following:

a) Path finding

Path finding is the plotting of the shortest path of radio propagation between two points.

b) Tracking

A tracking system supplies a timely ordered sequence of location data for observing the motion of people or objects and initiates further processing.

c) Real-time locating systems (RTLS)

Real-time location systems are used to identify and track the location of objects or people in real time automatically, usually inside a building. Wireless RTLS tags are worn by people or attached to objects. In most of these systems, a set of fixed reference points receive wireless signals from tags and determine their location. These systems are able to track automobiles, locate merchandise in a warehouse, and find medical equipment in hospitals.

1.7 Properties of a good localisation system

A good localisation system: [11]

1. Is light enough to run locally on the phone.
2. Requires minimum effort in setting up (fingerprinting).
3. Is stable with the variation of time.
4. Is robust to changes in the environment.
5. Works equally well with any measurement device (any mobile phone).
6. Is accurate enough to predict the room the user is in.

1.8 Thesis outline

The Chapter 1, (current chapter) gives an introduction to the research conducted. The objectives of the research are discussed in this chapter along with applications, advantages and relevance to national needs.

The second chapter, Preliminaries provides the complete details of the localisation techniques and systems developed, and level of research already carried in the field.

Chapter 3 explains the Proposed technique and System Model of the simulations. The mathematical models and equations used in the simulation are also discussed.

Chapter 4 Results and Discussions presents the results of the simulations, and the performance analysis under different scenarios and conditions.

Chapter 5 is Conclusion and Future work.

Chapter 6 is References and Bibliography.

Chapter 2

PRELIMINARIES

2.1 Types of Measurement

While the prime focus of this research is fingerprinting based positioning, some of the other distance measurement (lateration) techniques are discussed in the following sub-sections to develop a clear knowledge of the existing techniques.

2.1.1 RSS

The Recieved Signal Srength (or RSS) measurement works on the principle that the strength of the received signal varies inversely with distance. A Wi-Fi based Indoor Positioning System is commonly based on this type of measurement. Mapping between measured RSS and the distance between the receiving and the transmitting node is measured with the help of a mathematical model using Euclidean Distance, which is a measure of the distance between an 'online'RSS value and the 'offline'training database RSS records.

2.1.2 TOA

The time at which radio signal reaches from one transmitter to a remote server or receiver is called Time of Arrival (or TOA). Measurements of signal propagation delay are used to obtain distance between a pair of nodes. Time of Arrival makes use of the absolute time that a signal arrives at a certain base station rather than the time difference between the departure of signal from one station and arrival at the other station. As signals travel with a known velocity, the distance can be directly calculated from the Time of Arrival using d= v*t. At least three base stations are required to constrain the precise position of the receiver to a single point. A number of radio-location systems use TOA, including GPS.

2.1.3 TDOA

Time Difference of Arrival (or TDOA) is the measurement of time difference of a radio signal between departure from one station and arrival at the remote station. TDOA does not rely on absolute distance (as in TOA) calculated between a pair of nodes. Two methods are used to calculate TDOA.

1) Multiple signals from synchronised anchor nodes are broadcast and the receiver measures the TDOA.

2) A reference signal is broadcast by a mobile transmitter or node and is received by a number of fixed anchor nodes.

Synchronisation of all the anchor nodes in the network is a prerequisite for each method. Two TDOA measurements and at least three anchors are needed to calculate the position of a mobile node.

2.1.4 PDOA

In Phase Difference of Arrival technique a couple of continuous wave signals at frequencies f_1 and f_2 are transmitted. Phase difference is measured at the receiving station which is proportional to the distance and inversely related to the difference of f_2 - f_1. The phase estimation error is very small due to the presence of extremely small signal bandwidth. Radar systems first made use of this technique.

2.1.5 Proximity

Proximity systems use beacons for positioning. These beacons are bluetooth units acting as transmitters which broadcast their identifying information to nearby smartphones. The simplest method which defines the proximity technique is the formation of a circle around the transmitting beacons with radius r_0. The main advantage of this method is that it does not require any additional hardware and also the time synchronisation among nodes is not needed.

2.1.6 RSS and Wi-Fi Positioning

One of the main advantages of Wi-Fi based indoor positioning is the cost effectiveness because compatible devices are localised without the installation of extra software and no LOS requirement. [34] Wi-Fi based Indoor Positioning Systems (IPS) have improved efficiency in quite a number of organisations and industries. An increasing number of companies have welcomed the new functionalities provided by Wi-Fi Indoor Positioning Systems. There is a prominent interest for Wi-Fi IPS and Real Time Locating Systems (RTLS) in the telecommunications industry because of their potential in creating new services from existing ones. The developers of IPS have great market potential for companies and this is predicted to be even larger in the coming years.

Indoor wireless positioning based on received radio frequency signal strength has become more popular for researchers in recent years. Wireless positioning is a rapidly growing technology used both in home and business networking sites. Nowadays, wireless networks are profoundly set up in educational institutes, hospitals, shopping malls, airports and so on. Wi-Fi location determination uses existing Wi-Fi equipment such as those installed in personal computers, mobile phones and PDAs.
Another Wi-Fi based technology is termed as the IoT or Internet of Things. Location estimate in IoT forms as the pioneer of research in this field.
Researchers have implemented various positioning algorithms to find the indoor position. Some popular algorithms among these are signal strength mean value algorithm, the K-nearest neighbors algorithm, and Bayesian positioning algorithm. K-nearest neighbour algorithm is a widely accepted algorithm and it has the least complexity and has been used for the purpose of this research.

2.1.7 Wi-Fi Multilateration

The basic idea of the multilateration is that the distances from an unknown station to several reference locations contribute to finding the presence of this node. [35] Multilateral algorithms [32] basically measure, compare and process the location data received from reference stations in the environment. [33]
In multilateration the position can be estimated by using the signal strengths received from several non-collinear (series of points that are not on the same line in a plane) Access Points. Using the Received Signal Strengths and a path loss model, the distance between the device and Access Points can be estimated. Then the relationship between this distance and signal strength gives an estimate of the position of the user.

2.1.8 Trilateration for Indoor Positioning

Trilateration for indoor positioning is relatively simple approach in which GPS receivers calculate the position of objects by using the mathematical process of trilateration. This method finds the position of an object or target node based on the following two measures:

1) The distances of the object from any three different known points.

2) The coordinates (position) of any three points.

In trilateration, the target nodes are unknown location nodes which have to be calculated by using known coordinates of any three points. By taking the coordinates of any thre Access Points we can easily find the position. In the conventional trilateration method, the intersection of three circumferences gives the position of the target node.

2.2 RSSI Fingerprinting Method

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.

In the offline phase the RSS fingerprints are sampled at each reference position from a number of access points (AP). APs are usually fixed transmitters such as Wi-Fi routers. A reference position 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 points are not sufficient to characterise 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 users mobile measures a vector of RSSI value at an unknown location and then compares the RSSI 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.

Hence, the method of finding loci of users through an RP database in which RSSI for each point is recorded and compared with the instantly received RSSI vector at the users is called fingerprinting for location estimation. Fingerprinting helps to significantly improve the accuracy and precision of conventional signal strength lateration techniques. [19]

2.3 Dilution of Precision (DOP)

Dilution of Precision, or DOP, is a concept taken from GNSS satellite navigation. It is a figure of merit which specifies the effect of geometry formed by navigation satellites on the precision of positional measurement. DOP is a number which signifies the effect of satellite geometry on the receiver's calculated position, inherently describing how good the satellite geometry is in estimating the position. [38] DOP only depends on the positions of Access Points relative to the location of the receiver. [40] Calculation of GDOP coefficient has been widely used to determine appropriate positions for satellites in Global Satellite Navigation System and can be applied to indoor positioning systems to find suitable locations for Wi-Fi Access Points. The coefficient of dilution of precision is a good candidate to specify the most adequate Access Points distribution in a Wi-Fi network. [40]

Using the relations for DOP in the indoor environments calls for an explanation of 'pseudorange'as discussed in Section 2.3.1.

2.3.1 Pseudorange measurements

A GPS receiver computes its three-dimensional coordinates and its clock offset from four or more simultaneous pseudorange measurements. These are measurements of the biased range (hence the term pseudorange) between the receiver's antenna and the antennas of each of the satellites being tracked. This is derived by cross-correlating the pseudorandom noise code received from a satellite with a replica generated in a receiver. The accuracy of the measured pseudoranges and the fidelity of the model used to process those measurements determine, in part, the overall accuracy of the receiver-derived coordinates.

DOP gives a measure of the extent to which the pseudorange deviates from the actual distance. Pseudorange is depicted in Figure 2.1. [43]

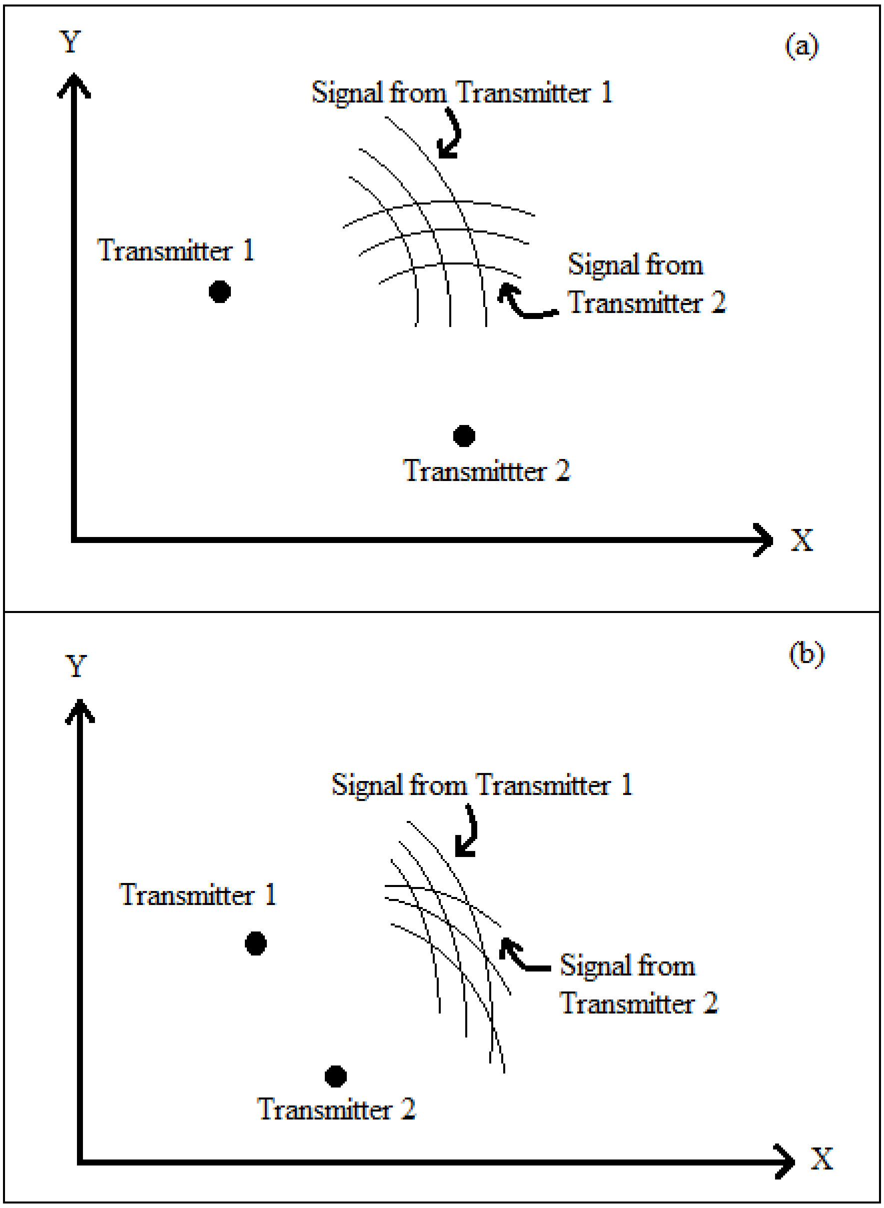

Figure 2.1: DOP: The uncertainty in the receiver's position [43]

Dilution of Precision (DOP) measures the relative degradation or reduction in the certainty of a navigation solution based on one-way range measurements from a set of transmitters. For GNSS, if four or more satellites are in view of a ground receiver, a navigation solution consisting of the position of the receiver and the offset between the receiver clock and the GPS clock is calculated.

DOP estimates the error in the geographical position of a receiver. The calculation of GDOP (Geometric Dilution of Presision) coefficient has been widely used to determine appropriate positions for satellites in Global Satellite Navigation System and can be applied to indoor positioning systems to find suitable locations for Wi-Fi Access Points. Candidate locations for APs are the solutions of a Genetic Algorithm (GA), as described later.

The position, horizontal, vertical, and time DOPs are defined as follows:

$$PDOP = \sqrt{(\sigma_E^2 + \sigma_N^2 + \sigma_U^2)}/\sigma \tag{2.1}$$

$$HDOP = \sqrt{(\sigma_E^2 + \sigma_N^2)}/\sigma \tag{2.2}$$

$$VDOP = \sigma_U/\sigma \tag{2.3}$$

$$TDOP = \sigma_T/\sigma \tag{2.4}$$

Where σ represents the variance in East, North and Up direction for the position dilution of precision; in the East and North for horizontal dilution of precision; in only the Up direction for vertical dilution of precision and the Time offset for time dilution of precision.

For 3D positioning the East, North and Up directions are considered, whereas for four positioning nodes, the time offset is also included; otherwise it is skipped.

The variants of Dilution of Precision are summarised as follows:

a) GDOP (Geometric Dilution of Precision) measures the dilution of precision for the entire navigation solution, concerning with the geometry of the localising entities. GDOP combines the dilution of precision of the position and optionally, clock-related components of the navigational solution.

b) PDOP (Position Dilution of Precision) measures only the dilution of precision associated with the positional portion of the navigational solution. It is further divided into four categories:

- HDOP (Horizontal Dilution of Precision) measures the dilution in precision for the latitude/longitude (horizontal) components of the Positional portion of the navigational solution.

- VDOP (Vertical Dilution of Precision) measures the dilution of precision for the altitude (vertical) component of the Positional portion of the navigational solution.

- EDOP (East Dilution of Precision) calculates the dilution of precision in the Eastern component of the navigational solution.

- NDOP (North Dilution of Precision) determines the dilution of Precision in the Northern component of the navigational solution.

c) TDOP (Time Dilution of Precision) measures the dilution of precision in the time portion of the navigational solution.

Table 2.2 shows DOP values and their significance with respect to system accuracy. A DOP value of around 1 is considered ideal and gives the most accurate location estimate.[3]

DOP	Rating
1	Ideal
2-3	Excellent
4-6	Good
7-8	Moderate
9-20	Fair
21-50	Poor

Table 2.2: DOP Rating [3]

2.4 Localisation

Even though GPS is a promising and reliable technology, the GPS signals cannot be detected due to environmental obstructions which renders GPS unsuitable for some terrestrial applications. Since the satellite signals are heavily attenuated passing through the atmosphere and their intensity is further reduced due to trees, heavy fog or man-made structures e.g. buildings, this system can not be especially useful for densely populated urban areas and also inside buildings and corporate setups. Even in indoor environment, the GPS does not work well because the walls, floors and other construction objects greatly attenuate the signals received from satellites. [20] Because indoor applications require a continuous, seamless and ubiquitous positioning, the GPS is not suitable due to great signal attenuation. [21]

To overcome this issue, researchers restored their focus to land-based positioning and tracking systems such as Wi-Fi IPS for a variety of situations which cannot make use of satellite signals.

The fundamental techniques of locating and tracking radio frequency devices differ greatly from those of data communications as in communications, the information is transferred from one station to another. The information carrier might be RF, sound or light and a single, reliable link is sufficient to transfer data between the stations. Locating a device whose location is completely unknown however requires a completely different approach than traditional data communication in which only data is transmitted.

As seen from various researches, localisation can be achieved in the following ways: [14]

1) Geometric methods e.g. trilateration, triangulation, hyperbolic methods.

2) Fingerprinting methods or signal mapping.

The geometric methods include those techniques which can locate or track devices based on properties of the signals being transmitted. Time of Arrival, Time Difference of Arrival and Angle of Arrival are examples of geometric location measurement techniques, while an alternate technique is the Received Signal Strength based fingerprinting method which is more robust and less susceptible to errors.

Figure 2.2 shows the architecture of a position estimation system [42].

The received RF signals being observed at various reference points are characterised

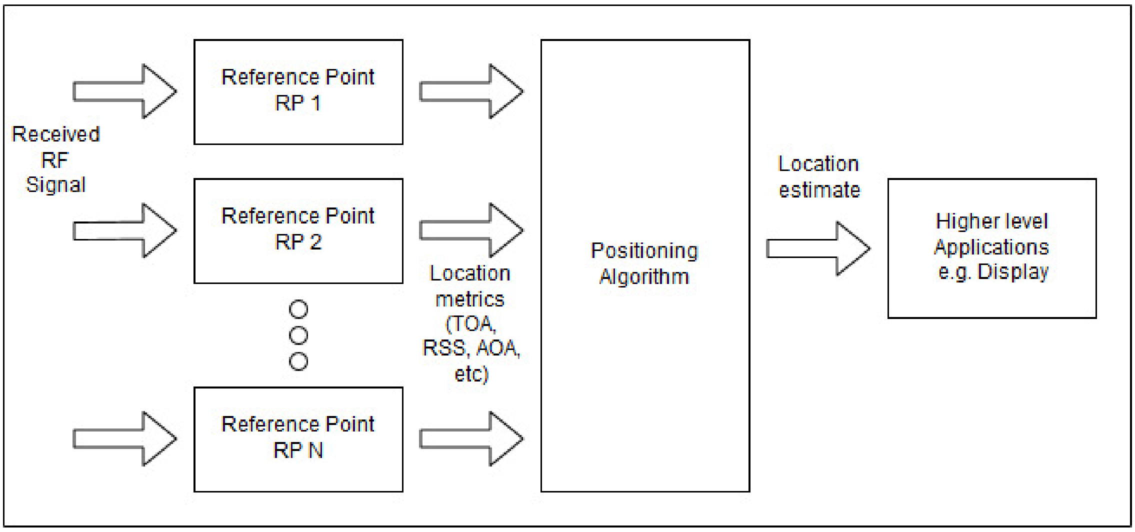

Figure 2.2: Positioning system architecture [42]

by location metrics e.g. TOA, RSS or AOA etc and a positioning algorithm is used to find the location estimate to be observed at a higher level application such as display.

2.5 Distance/position estimation metrics

A location system needs to obtain range estimates from fixed anchors or reference points in order to estimate the location of a node. Range estimates can be obtained using different metrics. RSS and TOA/TDOA are examples of such techniques.

2.5.1 RSS

The energy of an RF signal, after being transmitted, experiences propagation loss which is proportional to the distance that the signal travels. Because RSS level changes instantly, it is necessary to use the average value of RSSI in computations. [15]

A commonly used radio signal propagation model is given as

$$P_r(dB) = P_t(dB) - 10\alpha log_{10}d \tag{2.5}$$

Where P_r (dB) and P_t (dB) represent the received signal power and transmitted signal power in decibel. α is the indoor attenuation factor and is dependent on the propagation environment. For free space propagation, α = 2. There are wide range of values possible for α, e.g. for a hard partition office environment α is found to be 3.0 and for a lab environment with metal partitioning it is reported to be 3.3. [41]

2.5.2 AOA

Angle of Arrival (AOA) information from two different anchor nodes can be used to determine the position of a user by using triangulation. AOA estimation is also referred to as Direction of Arrival (DOA) estimation, direction finding or bearing estimation in many contexts and has been researched extensively. AOA is demonstrated in Figure 2.3.

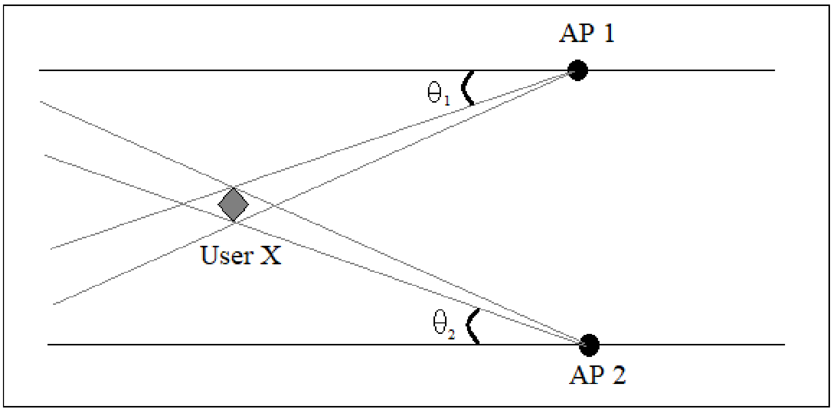

Figure 2.3: Angle of Arrival [14]

2.5.3 TOA

Time of Arrival (TOA) is another distance estimation method in which the distance is estimated based on the time the signal takes as it travels from a transmitter to the receiver. The speed of RF propagation is very well known in both free space and air, therefore it gives a direct estimate of the distance between the transmitter and the receiver when the time travel is measured. TOA is the only important parameter which needs to be estimated correctly in a multipath propagation environment. TOA needs an LOS path or the pirect path (DP), without which the TOA measurements can be inaccurate.

2.5.4 TDOA

TDOA, or hyperbolic positioning, is a positioning method in which the receiver calculates the differences in the TOAs from different reference points. This method automatically removes the clock biases between the transmitters and receivers, since the differences between the TOAs from two transmitters are only considered.

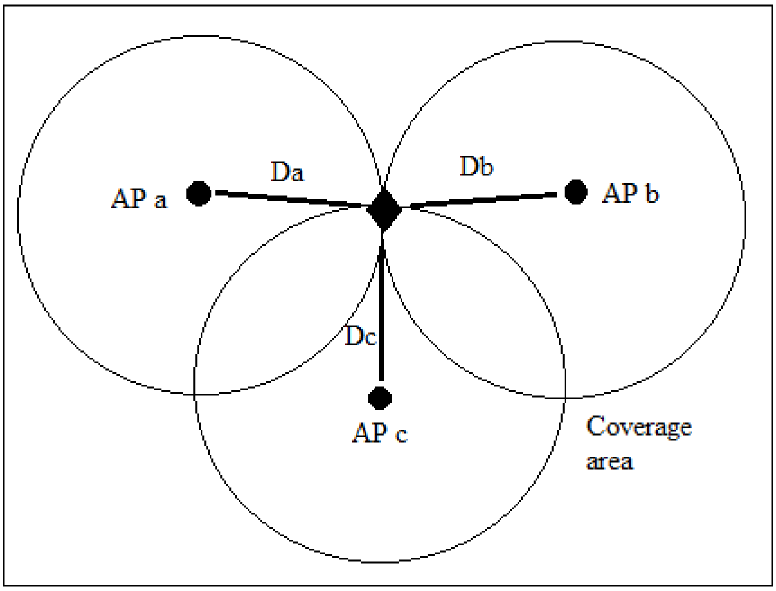

Figure 2.4: Time of Arrival

TDOA is based on Hyperbolic Lateration in which TDOA between two sensors is calculated and the value is used to construct a hyperbola. The focal points of the hyerbola are the respective locations of the sensors. The position estimation using TDOA is shown in Figure 2.5.

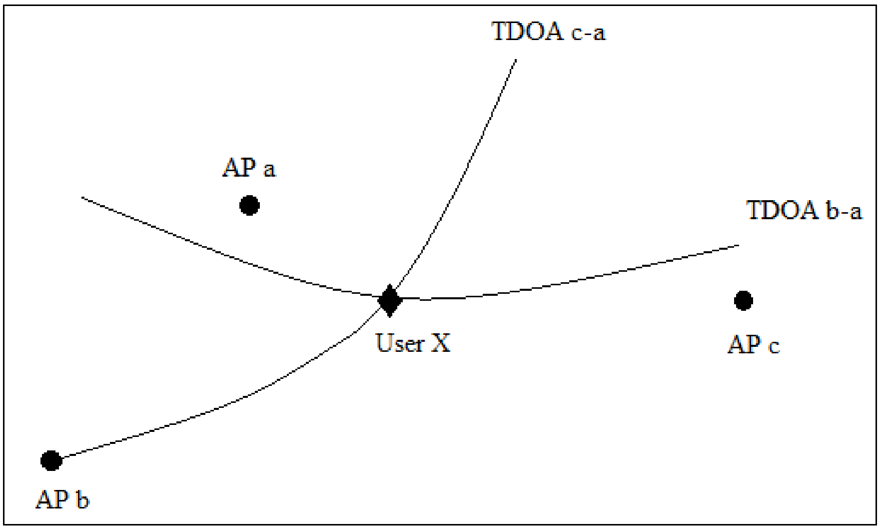

Figure 2.5: Time Difference of Arrival

2.5.5 Fingerprinting

Fingerprinting is a training-based method which is estensively used in positioning systems in WLANs because it is not affected by the time varying characteristic of radio propagation caused by multipath, shadowing, and interference. [17]

Geometric methods of positioning do not provide robustness in terms of performance as compared to fingerprinting method. This method is called so because the RSSI vectors received at the Reference Points (RPs) from various Access Points (APs) are all different from one another as in a fingerprint collection. Also this method is more accurate and convenient to implement as it does not require an LOS path between the RPs and the APs.

This method is essentially a two-step approach. The first step is the construction of an offline training database in which a set of grid points and the RSSI received at various RPs are used. The next step is the online phase in which actual localisation is done by comparing the RSSI vector received from APs at the user with the training RP database using some low-complexity algorithm like the K-nearest neighbour algorithm. The position of the user is then the position of the RP 'nearest'to it and the difference between the RP's location and the user's actual location contributes to the localisation error calculated by various samples since the RSSI values are also averaged.

Fingerprinting technique inherently captures all the environment related signal propagation effects e.g. multipath, shadowing and scattering and hence it can be used for applications in which geometric methods do not work up to the mark.

Figure 2.6 describes a location fingerprinting process, highlighting the offline and the online phases.

The Reference Points (RPs) are represented by RSS values from a number of Access Points, constructing a 'fingerprint'database for each of the received RSS vectors. These vectors are all different from one another attributed to the distance kept between reference locations and hence are named as fingerprints. When the offline database formation is complete, the location of a user can be determined in the online stage by comparing the RSS vector at the user with the vectors in the RP database using an algorithm like the k-nearest neighbour algorithm. The accuracy (the error which exists between the location estimated by the IPS and the actual location of the user) of such systems depends upon the underlying techniques used and improving accuracy is also

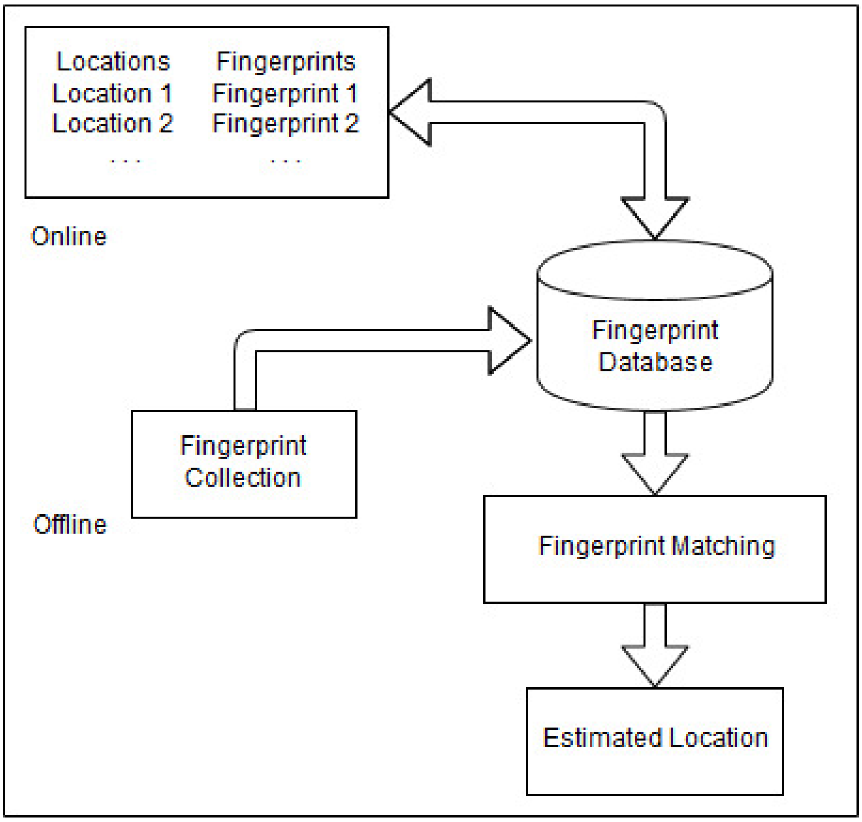

Figure 2.6: Location fingerprinting method [14]

the main concern of this research.

The time taken by a signal to obtain a position and the energy needed to successfully localise the user is directly dependent upon the complexity of the indoor positioning system. It is important to keep the complexity as low as possible as it also decreases the power consumption of the system. [16]

Chapter 3

SYSTEM MODEL AND PROPOSED SCHEMES

Indoor location-based services (ILBS) have gained interest in the recent years because of the social and commercial usage. [12] Indoor environment is more often complex, defined by objects and non-line-of-sight (NLOS) distances between the RF tranmitters and receivers. The presence of obstacles, signal fluctuation or noise and environmental changes clutter the indoor environment. In order to gain satifactory ILBS, a high localisation accuracy is required despite such a complex environment. [13]
This chapter provides definitions and detailed description of a Wi-Fi based Indoor Location technique. Starting from definitions, this chapter covers an Access Point deployment algorithm (the Genetic Algorithm or GA) and the user-to-Access Point position vector algorithm which classifies the Access Points on the basis of RSSI and position vector objective functions.

Wi-Fi based indoor location systems (Indoor positioning systems or IPS) contain a set of wired or wireless Access Points (APs). These are networking devices which allow wireless Wi-Fi devices to be connected to a wired basic network infrastructure. The wireless APs make up a Wireless Local Area Networks (WLAN). The access points act as central transmitters and receivers of wireless radio signals.

To achieve the full performance potential of the IPS, it is a best practice to ensure the proper placement of APs. The APs are mainly distributed throughout the interior spaces in many office WLANs, providing service to the surrounding workplace. While the AP locations are selected traditionally on the basis of coverage, WLAN bandwidth, channel reuse, security, aesthetics and feasibility, the indoor positioning system requires that a good location fidelity is achieved too.

The Access Points should not be clustered solely in the interior or center of floors, rather they should be arranged in perimeter locations to provide full coverage for the location based systems. [14] This complements the location fidelity and adds robustness to the positioning environment. The Access Points in some cases might also be providing data services along with a location based service, therefore it is a best practice to distribute the APs in for example, the four corners of a room or office building where positioning is required to be done. These perimeter APs play a key role in ensuring good location fidelity within the areas reached by their signals.

However, in some cases, customer may prefer to deploy APs in places other than perimeter locations which can make the location estimate less efficient in terms of accuracy. With the underlying principle of RSSI-based positioning, since the Reference Points (RPs) have to be calibrated in the same environment, perimeter locations for APs are the best. Center locations of an AP can either be very close to the RPs or overlap the RP altogether, and ambiguity will be created in the RSSI measurements for RP calibration in the offline phase. In the online phase, again a centered location of AP will overlap a user's location, excluding the RP location which might have been calculated at that point but is left out due to the placement of an AP in that area. Also in situations where asset tracking is required, placing the APs other than perimeter locations will hinder the movement of goods inside the building.

This research work focuses on the deployment of additional APs in an existing Wi-Fi based IPS, based on certain search perimeters like the DOP (Dilution of Precision). The second stage of the research calls for a selection of APs from among the initially placed and the deployed ones based again or some simulation parameters as discussed in this chapter and in Chapter 4.

3.1 Received Signal Strength Indicator

The Received Signal Strength Indicator (or RSSI) is a device-specific identifier usually expressed in dBm for Wi-Fi Access points or routers. It is used to establish link between the mobile user and the Access Points as seen in sending CTS (clear to send) or ACK (acknowledge).

A user can receive signals or RSS from more than one access points at a given time and the RSS is used to estimate the device's location. According to studies, a minimum of three access points are enough to locate an object in 2-dimensional space. Although

systems have been developed which make use of two APs for 2-dimensional space, but those systems are more complex. Moreover, due to low cost of access points and ease of availability, increasing the number of access points is not an issue.
An increase in number of APs can also lead to improved accuracy of the system, as shown later in Chapter 4. Suppose an MS (mobile station) receives RSSI from Access Points (or base stations) in an indoor positioning system (IPS). The set of RSSI is recorded in a table, known as the RSSI database. This database is saved for the purpose of comparison with the instantly received RSSI vectors in the online phase for localisation.

3.2 The database

These are basically reference points (RPs) at which RSSI is recorded in the offline phase. These are different from grid points and their cardinal number is arbitrary. In the online phase, a user is referenced to one of these points using K-nearest neighbour (KNN) algorithm.

3.3 K-NN algorithm

The K Nearest Neighbour (KNN) algorithm finds the K 'fingerprints'which have the minimum signal distance to the online reported RSSI vector from among various fingerprints in the database that the location server has in store. [7]
Consider a two dimensional plane in which three Access Points are present, and the coordinates of one user have to be found out. From a set of Reference Points, the RSSI is compared with the instantly received RSS at the user.
If K=3, the algorithm allows for three neighbouring points at which RSS recorded earlier is compared to the instantly received RSS vector at the MS (vector here refers to RSSI from all the three Access Points).
Averaged over 'K', the coordinates or location of the user is then the coordinates of the nearest RP whose recorded RSS matches with the user's RSS vector.

3.4 Flow for the progress of this research

The following flowhart emphasizes the step-by-step progress of this research.
Starting with GA which optimises GDOP to minimum with each iteration, candidate

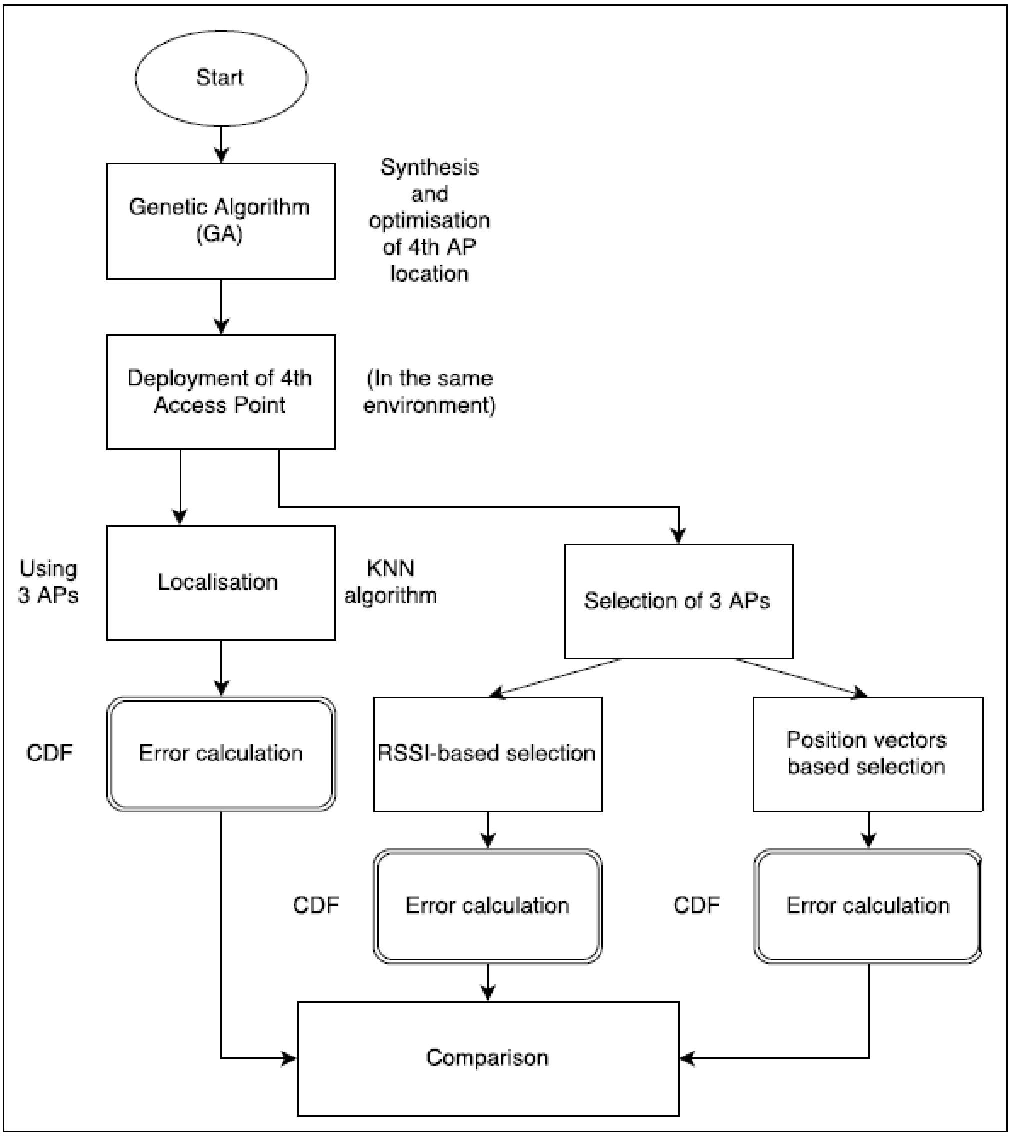

Figure 3.1: Flow for Access Point deployment and selection

locations for APs are found out, following by deployment of APs at the same. Then user localisation in the indoor environment is run in parallel with AP selection to conserve the spectrum bandwidth.

3.5 Access Point Deployment

One way to deploy Access Points is to analyse and compare all the possible AP setups, but that is a complex task. Also the propagation characteristics of the environment have to be taken into account. [18] Hence a deployment solution is acquired based on simulations because a real-world modelling of such system is complex. The usage of Genetic Algorithm (GA) in this research, therefore calculates a suitable location of AP while at the same time considering the multipath propagation loss and areas of optimum DOP.

3.6 The Genetic Algorithm

Genetic algorithm (GA) was proposed by John Holland and has been extensively applied in many fields. [39] Genetic Algorithm is one of the most important computation methods and has been used in searching, optimisation and approximation fields. [37] It is an algorithm which consists of a group of iterations to figure out solutions to the GDOP problem. Consisting of four stages, the Genetic Algorithm or GA aims to maximise a fitness function which results in the minimum EDOP. The Access point is then deployed on the location of minimum EDOP as shown in the next section.
The end result of GA is essentially the acquiring of x- and y- coordinates for the new deployable AP. GA consisted of four stages and the pseudocode for simulations is described below.

3.6.1 Elitism

Possible locations for APs are the specimens for Genetic Algorithm. In the first step, some specimen having the best value of fitness function are selected from a total population of candidate locations. The fitness function is a minimum requirement which should be met by a candidate solution, in this case, a fixed AP-to-AP distance which may not be decreased for good measurement. These specimens are isolated from the population and set aside so that the conservation of the presently best solution is ensured at each iteration.

3.6.2 Crossover

The selected best solutions from the Elitism stage are broken up based on a crossover probability and alternatively combined with each other, leading to new incurred solutions. This is called crossover of two parent genomes. The place of the crossover process is selected randomly and a new specimen originates. This new specimen is now the combination of information from the first specimen and information from the second specimen. The crossover process is shown in Figure 3.2.

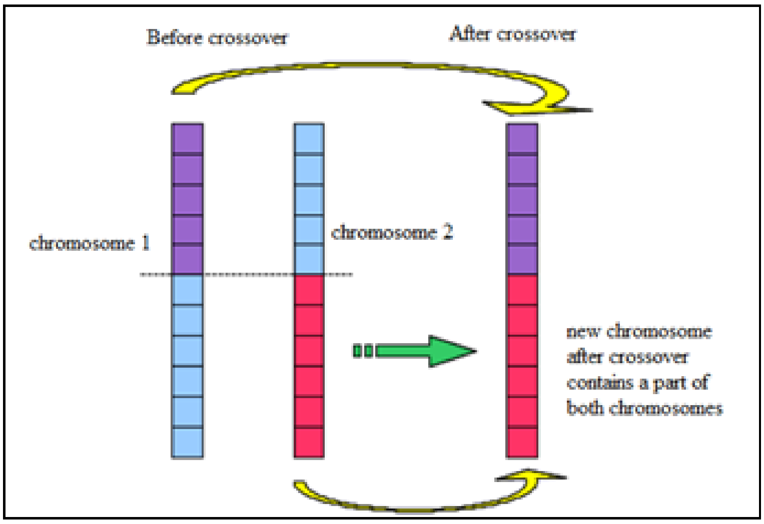

Figure 3.2: Process of crossover of two specimens and origination of a new specimen

3.6.3 Mutation

In order to prevent the group of crossed-over specimens (genomes) to be restrained just to a combination of already existing solutions, they are mutated at every step of the algorithm, based on a mutation probability. This is basically a random change of one gene in the crossed-over genomes.

3.6.4 Elimination

Running in parallel, the weak solutions are eliminated from the new updated set of solutions. In this phase the worst specimen is deleted from the population. After this phase the specimens which were preserved at the beginning in the elitism phase are returned back into the population.

Algorithm is terminated when all k iterations are over. It is necessary to select sufficient amount of iterations in order to obtain good results. For this, a selection of termination condition is necessary so that the value of fitness function will not change more than the defined certain value. The termination condition for simulations in this research is the end period of an evolution time of 300 seconds.

3.7 Pseudo code for Access Point Deployment using Genetic Algorithm

Deployment of Access Points as the first part of this research is carried out with the help of Genetic Algorithm. The GA selects position of new APs to be deployed based on a fitness function which in this case is the maximisation of distances between two closest APs. [3]

The following pseudo code defines a step by step progression of the GA algorithm.

Step 1 *Generate Population*

*Specimens[x,y] = Graycode[size]*Total Population;*

Step 2 *Identify Best Specimens*

Define (Fitness Function) = A fixed AP-to-AP Distance;

G1 = Find GDOP for each Specimen(x,y);

Best Specimens = Specimens[x,y](minimum(G1));

Separate (Best Specimens);

Step 3 *Crossover*

Make two groups (Best Specimens);

Split (Best Specimen) at Crossover Probability;

Shuffle and Combine (Split Specimens);

Step 4 *Mutation*

Flip one bit (Combined Specimens) based on Mutation Probability;

Step 5 *Elimination*

Discard (Specimens) with AP-to-AP distance>Fitness Function;

Retain rest (from Step3, Step4);

G2 = Find GDOP (Retained Specimens);

GDOP = min (G2);

AP Location(x,y) = Specimen(x,y) for which GDOP(min) exists;

Deploy new AP at AP Location(x,y);

The algorithm first generates a total population of 256 specimens. The best solutions, called as 'elite' solutions, are gray coded to make up a set of 'chromosomes', with each chromosome containing two genes i.e. the x- and y- coordinates of new APs to be deployed. These solutions are preserved to be used at a later stage, before being subjected to crossover of any two solutions to form a new specimen with the crossover occurring at random bit locations. One bit at random is altered from each crossed over solution to avoid repetition of resultant solutions, the process being called as mutation. The crossover and mutation probabilities determine how many solutions remain unchanged in the next generation and how many of the new specimens are modified by 1-bit respectively.

The final set of solutions is then checked for solutions which deviate most from the fitness function and are discarded.

GA determines the coordinates of AP for calculation of GDOP coefficient. The algorithm uses a group of n-D solutions from which it selects the best ones (where n=2 for 2D and n=3 for 3D positioning). [3]

3.8 Access Point Deployment

The following discussion relates Euclidean Distance (ED), Geometric Dilution of Precision (GDOP) and Enhanced Dilution of Precision (EDOP) to the Access point deployment problem.

3.8.1 The ED

To ensure the positioning accuracy, the Euclidean distance (ED) should be maximized. ED is defined as, [1]

$$ED_i = \sum_{jN(i)} ||RSSI_i - RSSI_j||_2/N_i \tag{3.1}$$

$$Max : ED_{ave} = 1/n \sum_{i=1}^{n} ED_i \tag{3.2}$$

Where ED_i is the Euclidean distance between the i^{th} RP and the surrounding RPs, N(i) is the aggregate of surrounding RPs with cardinal numberN_i. ED_{ave} is the average ED of all the RPs and n is the number of RPs.

3.8.2 The GDOP

The locations of Access Points calculated through Genetic Algorithm are such that the GDOP coefficient is minimised. Computing the quotient between RSSI and GDOP to form an enhanced DOP, the EDOP, the resultant is used to find user locations.
Considering the Euclidean distance formula, [1]

$$r_{i,j} = ||(x_{r,i} - x_{a,j}) + (y_{r,i} - y_{a,j})|| \tag{3.3}$$

Where $x_{r,i}$ and $x_{a,j}$ refer to the x- coordinate of the i^{th} RP and the j^{th} AP; and $y_{r,i}$ and $y_{a,j}$ are their corresponding y-coordinates.
The coefficients of DOP matrix are calculated by:

$b_{i,x} = \frac{x_r - x_a}{{r_{i,j}}^2}$, $b_{i,y} = \frac{y_r - y_a}{{r_{i,j}}^2}$

Hence constructing the Matrix,

$$B = \begin{bmatrix} b_{1,x} & b_{1,y} \\ b_{2,x} & b_{2,y} \end{bmatrix}$$

The scalar quantity GDOP is the square root of the trace of B matrix, specifically:

$$GDOP = \sqrt{Tr(B^T B)^{-1}} \tag{3.4}$$

3.8.3 The EDOP

Introducing a new indicator, the Enhanced DOP, given as the quotient between ED and GDOP:

$$EDOP_i = ED_i / DOP_i \tag{3.5}$$

$$Max : EDOP_{ave} = 1/n \sum_{i=1}^{n} ED_i/DOP_i \tag{3.6}$$

$EDOP_i$ is the EDOP of the $i^{th} RP$ and $EDOP_{ave}$ is the average EDOP of all RPs. This EDOP is calculated as an objective function and fed into the GA optimisation algorithm, giving the candidate location for deployable Access Point(s). The number of Access Points to be deployed is according to the formulation of the planning problem. Due to low cost of deployment, additional AP placement in an already existing Indoor Positioning System will reduce positioning error and hence improve accuracy as shown in the results in Chapter 4.

3.9 Access Point Selection

In IEEE 802.11 Wi-Fi localisation networks, the performance given to the users greatly depends upon the users' ability to identify those Access Points which will offer the best location estimate. [36] A number of methods are used to determine the competency of Access Points, from which two techniques are discussed, the RSSI- based selection and the Position vector based AP selection.

3.9.1 User-to-AP position vectors

Figure 3.3 is a generalisation of position vectors as seen from a user to the Access Points.

The position vectors give position of a point relative to another point. If the location of a body changes, its position vectors formed relative to a point also change.
Considering that for the vector, its absolute value and a unit vector in its direction the following proposition is valid: [3]

$$eR = |R| \tag{3.7}$$

The vector R is then the Position Vector from user to a satellite, where R is its magnitude and e is the unit vector in its direction.
Taking unit vectors as the coefficients of the GDOP matrix, we can calculate position vectors by the magnitude of distance between the users and each of the APs. [2]
$| R |$ shows the position vector distance or specifically;

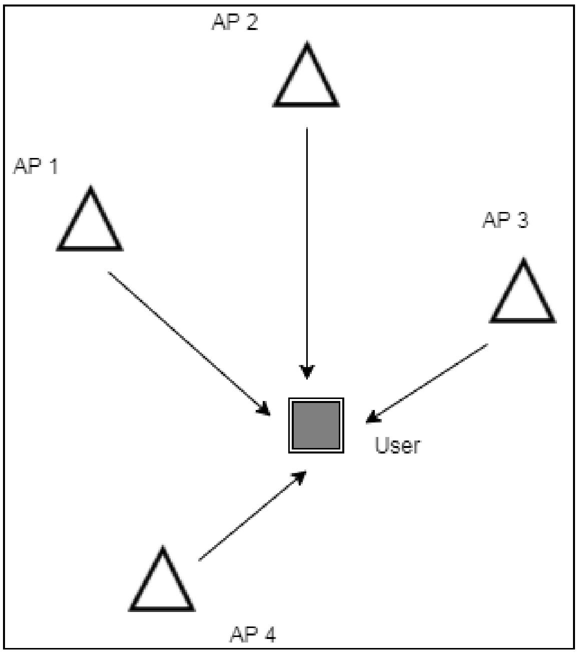

Figure 3.3: Position vectors from users to the Access points

Let P(0,0) and $P(x_1, y_1)$ be two points present in an x-y plane, then $\mid R \mid$ is calculated much like the Euclidean distance metric,

$$P(0,0) - P(x_1, y_1) = \sqrt{(x_1 - 0)^2 + (y_1 - 0)^2} = \sqrt{(x_1^2 + y_1^2)} \tag{3.8}$$

If P(0,0) $\neq$ Origin, and user a is present at location $P(x_a, y_a)$, and Access Point locations are $P(x_i, y_i)$, where i = number of Access Points, then position vectors seen from location $P(x_a, y_a)$ relative to the Access Points will be:

$$P_i = \sqrt{(x_i - x_a)^2 + (y_i - y_a)^2} \tag{3.9}$$

For a fixed set of Access Point locations, the probability of a user to be nearer to any three out of four APs is greater than that of being at equal distances from each of the APs. Therefore, with RSSI and position vector distances as metrics, an algorithm can reside at the user's end which selects those APs which are nearest to it for the k-NN localisation. There are two cases to be considered in this regard; first is the maximisation of the ED and hence RSSI objective function, defined above, and the second method consists of minimisation of the position vector distances returned by position vectors seen from the users to the APs.

As proposed by [2], there exists a relation between GDOP and the collective geometry formed by positions of signal transmitters and each of the receivers, setting the framework for a localisation algorithm which uses position vectors in addition to RSSI to determine a set of Access Points which will localise each user with the least error.

3.10 Pseudo code for AP selection

The algorithm developed for Access Point selection is the User-to-AP position vector algorithm, devised as follows:

Step 1 *Sorting*

(a) Sort descending[RSSI]

(b) Calculate position vector distances(Pvec)

(c) Sort ascending[Pvec]

Step 2 *Selection*

Select APs[greatest RSSI, smallest Pvec]

Step 3 *Return selected AP(s) index*

Step 4 *Select AP coordinate for the index value returned in Step 3*

The RSSI vectors are first sorted in descending order and the position vector distances are calculated using users' coordinates and the coordinates of the Access points. These position vectors, after being sorted in ascending order are used in the AP selection algorithm in which the APs which provided the highest RSSI and lowest position vector distances are selected to be used for a second-step localisation for better accuracy.

3.10.1 Explanation

For the indoor positioning system in which some Access Points are already present and their locations are fixed, while some have been deployed using the GA algorithm and EDOP coefficient, suppose a user is present at location $P(x_a, y_a)$, where x_a and y_a are the x- and y- coordinates to be determined for the user. After using k-NN algorithm, the location of user is found out and stored in a table.

Now in the position vector formulation step, the users coordinate information along with RSSI is also present and those Access Points are selected which are nearest to the user in physical space (x- and y-coordinates) and to the Reference Points in signal space (signal strengths).

Two methods of Access Point selection are proposed in the next sections.

3.11 RSSI based Selection

Objective function for this scheme is,

$$Max : RSSI = 1/n \sum_{i=1}^{n} ||RSSI_i - RSSI_j||_2 / N(i) \tag{3.10}$$

Where i = number of RPs with cardinal number N and j = number of APs.

After the RSSI at each user has been sorted in descending order as shown in pseudo code, the first n APs are selected for which RSSI is maximum.

3.12 Position Vector based Selection

A novel scheme of position vectors based AP selection when used in conjunction with RSSI based selection has the advantage that the localisation accuracy is improved. Also there is no need to introduce new infrastructure rather the existing IPS can be updated easily.

The objective function for this case is,

$$Min : P_{dist.} = ||P_x - P_a||_2 \tag{3.11}$$

Where x = number of Access Points and a = number of users.

Using user-to-AP position vector algorithm, those Access points will be used to localise users for which the position vector distances are minimum. The cumulative distribution functions for the RSSI based selected APs, the position vector-based selected APs and without the selection of APs (initial case) are discussed in Chapter 4.

Chapter 4

RESULTS AND ANALYSIS

This chapter contains the results generated using MATLAB simulations for a Wi-Fi based Indoor Positioning environment covering the floor of a building with dimensions of 16m*32m. Let this floor be composed of two small offices and two big classrooms. Dividing this floor into a uniform grid, a 1m space in each direction is set. Three existing APs are denoted by the triangles as shown in Figure 4.1.

Consider placing additional APs in an optimal way. Since the data collection process using real deployments is prohibitively time-consuming, it is replaced with a simulation methodology for simplicity.

4.1 Overview

The process of mapping offline RSSI measurements to real-time RSSI observations with the help of a Reference Point database and Access Points is called location fingerprinting. The offline phase comprises of updating a database of known reference locations in the grid map of coverage area. In the online phase, a user receives signals from Access Points and by using an algorithm e.g. the k-nearest neighbour algorithm, these real-time RSSI measurements are compared with the database of RP locations. The unknown location which is to be determined is then the location of the reference point closest to the user in signal space.

The database is made up of a large number of reference locations. In this study a total of 154 reference locations have been entered into a 154x2 data matrix.

4.2 Access Point deployment

Access point are placed in locations where they are likely to provide appropriate coverage. [24] For realizing the simulations, a rectangular area of the floor of a university building is considered, made up of two big classrooms and two small offices. Three Access Points are already placed at fixed locations as shown in Figure 4.1.

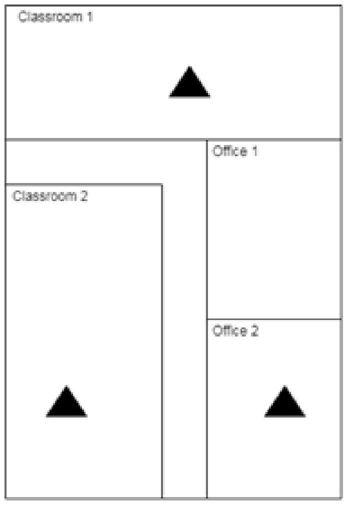

Figure 4.1: Map of experiment region

4.2.1 Multipath model

The foremost step of an Indoor Positioning system is to choose an appropriate Multipath model based on the design parameters. Empirical models are used in network design, while deterministic models are used for high precision applications. [8]

To devise a multipath model for indoor positioning systems, there are two types of obstacles to the signal path, the main or static obstacles and the mobile or moving obstacles. However, the accuracy of localisation algorithm depends more on the type of obstacles rather than their mobility. A Rayleigh model is used where multipath signal is present whereby it is impossible to achieve a line-of-sight path for a signal between the transmitter and the receiver.

A Rician model is applicable where an LOS is also present along with multipath signals received at the the MS in which the probability of an LOS present is determined using the received signal strength at the nodes.

Multipath dispersion due to the influence of many signal reflections and diffusion is one of the major issues in WLAN indoor environments. Walls, floors and roofs highly attenuate the signal and induce great variations in the mean received power. [22]
The partitions and the walls as depicted in the diagram are categorised in a path loss model by the Wall Attenuation Factor (WAF).

Using the following empirical path loss model for signal strength: [4]

$$P(r)[dBm] = P(r_o)[dBm] - 10\alpha log(r/r_o) - l.WAF \tag{4.1}$$

where $P(r)$ is the signal strength at distance r and $P(r_o)$ is the reference strength at distance r_o. α is the attenuation factor of the indoor environment. l is the number of walls between the transmitter and the receiver and WAF is the wall attenuation factor. To ensure the results reasonably accurate, the propagation factors should be approached with real signal strength measurement. [1] Simulating for an indoor Wi-Fi test bed of 154 reference points and 9 localised users, with 3 existing APs, the simulation parameters are defined in Table 4.1.

Dimensions of floor (meters)	14 x 25
Number of Reference Points (RPs)	154
Number of Existing Access Points (APs)	3
Transmit Power of Access Points (dBm)	24
Indoor Attenuation Factor (n)	3.1
Number of partitions in x - dimension	2
Number of partitions in y - dimension	2
Grid size (meters)	1

Table 4.1:Simulation parameters

Genetic Algorithm (GA) is used to optimise the EDOP defined in Chapter 3. The parameters for GA were as follows:

Evolution time = T= 300 s

Crossover probability = P_c= 0.8

Mutation probability = P_m= 0.05

Population size = PT = 256;

EDOP will be optimised in the evolution time of 300 seconds. Pairs of genomes from the elite solutions will be crossed over with crossover probability of 0.8. After the crossover is complete, one bit is flipped at random from the crossed over genomes so that the resultant ones are all different from one another. This is known as mutation and will occur at mutation probability of 0.05.

Table 4.2 shows the simulation results of variation in EDOP with respect to evolution time T.

T (sec)	0-20	21-245	246	247	248	249-300
EDOP	0.55	0.54	0.535	0.53	0.525	**0.5218**

Table 4.2: Evolution of EDOP with time T

The Genetic Algorithm is used to optimise the EDOP to a lower value and the coordinates of candidate AP location are updated with each iteration of the GA. The reduction in EDOP is not linear but constant for some ranges of T. This is a random process and evolves as the EDOP is calculated for the indoor environment considered.

From Table 4.2, at the end of evolution time T the EDOP coefficient is reduced to the value of 0.5218.

Figure 4.2 shows location of the new AP calculated through GDOP. Existing APs are depicted in blue while the new location for the deployed AP is depicted in red.

4.3 Localisation and Access Point Selection

The localisation step involves the following entities:

1) Three initial APs at locations (2,5), (14,5) and (8,25)

2) One additional AP at location of minimum EDOP

3) A set of 9 uniform and randomly distributed users

The localisation takes 3850 samples for error calculation in the CDF plots. The coverage area can have a number of users depending upon the time. For leisure hours users can be more for a university building while the number of users will be less during lecture hours. Nevertheless, an estimate has been made on a fixed number of users in the given area. In Figure 4.3, a set of 9 uniform and randomly distributed users are depicted in an area of 14x25 m^2.

Position vectors formed from each of the user to the Access Points are shown in Fig-

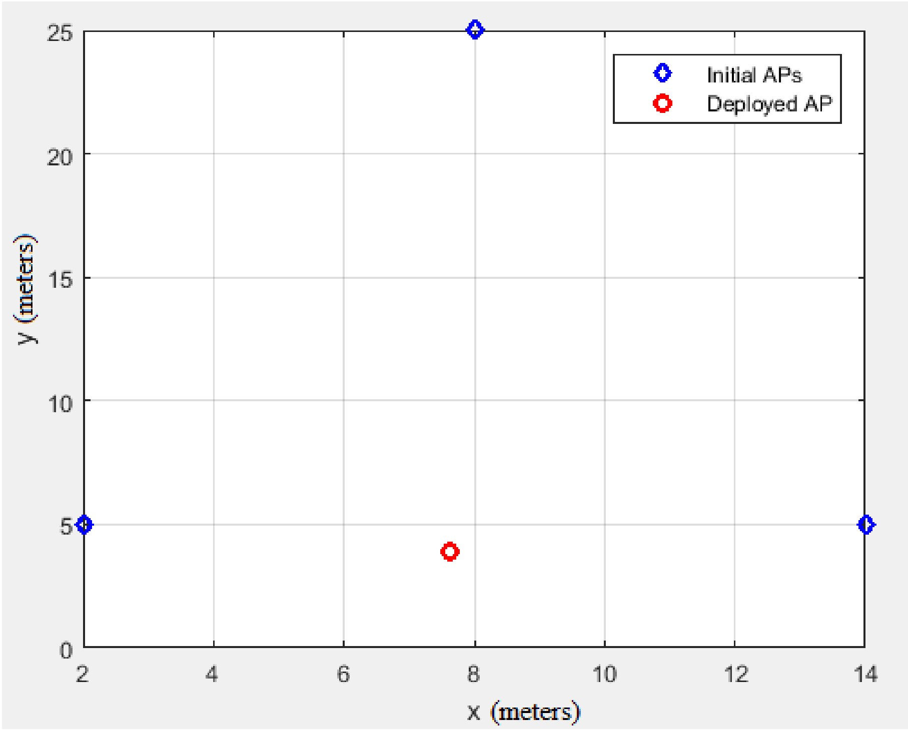

Figure 4.2: Initial Access Points (blue) and the Deployed Access Point (red)

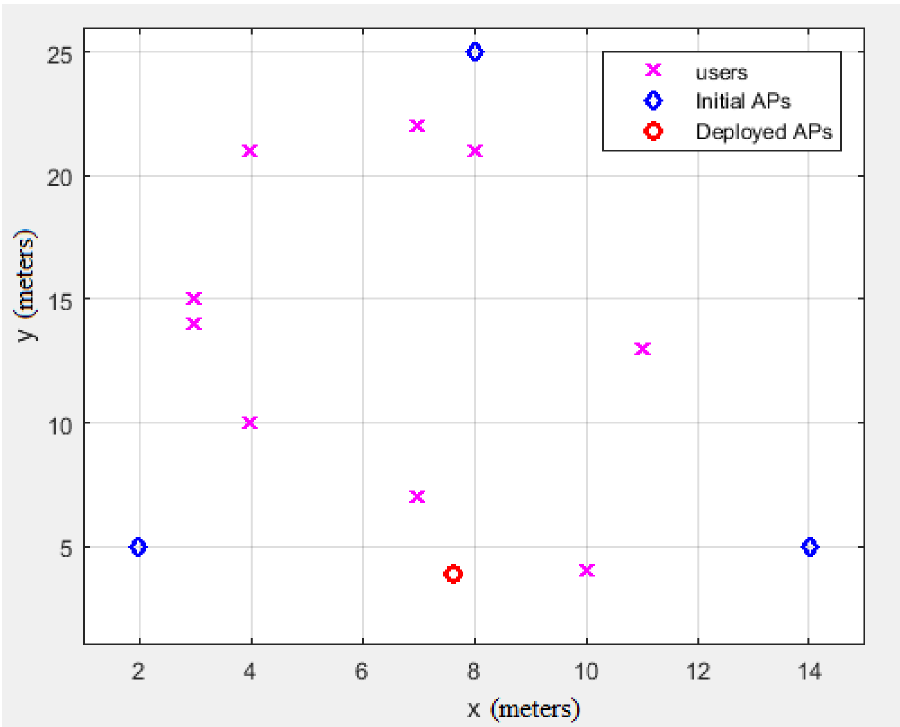

Figure 4.3: Set of uniform and randomly distributed users

ure 4.4. Minimisation of distances determined through position vectors is the framework for the second phase of Access Point selection; the first being the maximisation of RSSI at the users.

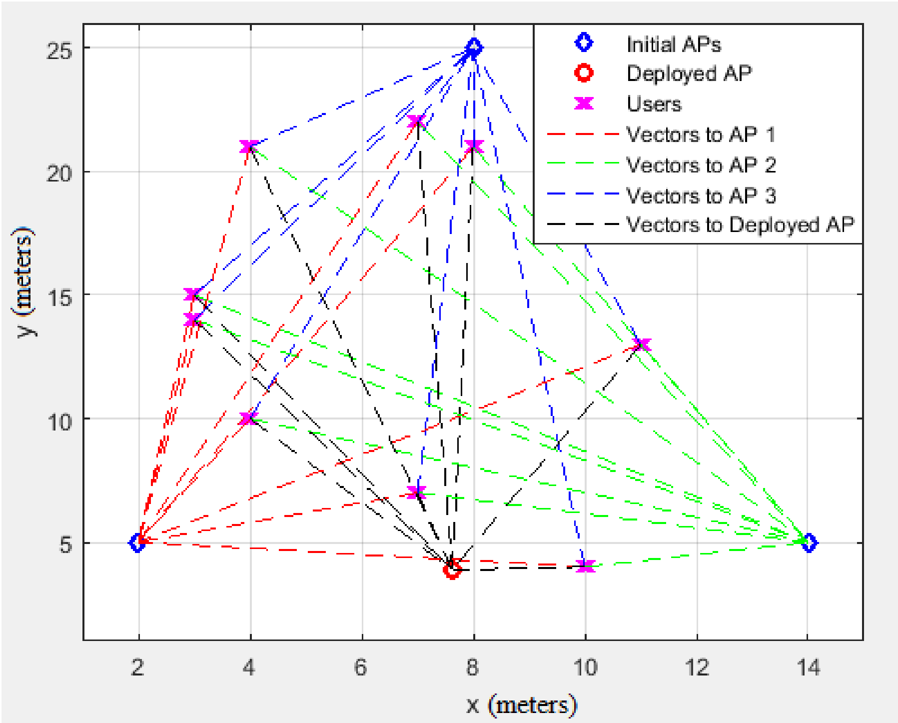

Figure 4.4: Position vectors from users to each of the Access Points

4.3.1 AP Selection

The next step after position vector formulation is the Access Point selection. In this step the users 'select' the APs which provide the best location estimate. Since we have used RSSI fingerprinting for localisation, the AP selection is first based on RSSI readings and then on the technique of user-to-AP position vectors. The purpose is to reduce the localisation error further which can not be achieved while solely using the RSSI measurements. To summarise, the Access Point selection is carried out based on the following two objective functions:

1) Maximisation of RSSI at the users

2) Minimisation of distance determined through position vectors (the user-to-AP position vectors)

In light of the algorithm for AP selection described in Section 3.3.2, the CDF localisation on three cases was observed.

Case I: Initial APs (3)

Case II: RSSI based selected APs (3 out of 4)

Case III: Position vectors based selected APs (3 out of 4)

Figure 4.5 shows cumulative distribution plots for localisation error with a total of 3 localising APs. The blue line represents error calculated using the existing APs for a set of 9 users, the red line shows error CDF for localisation of the same users with the help of RSSI-based selected APs, while the magenta line represents their localisation error calculated through APs selected by the user-to-AP position vectors.

Evidently, the position-vector based selection has produced the least error of 3.3 meters at the 90^{th} percentile; the RSSI-based selection yielded an error of 3.6 meters while the initial APs localised with an error of 4.3 meters.

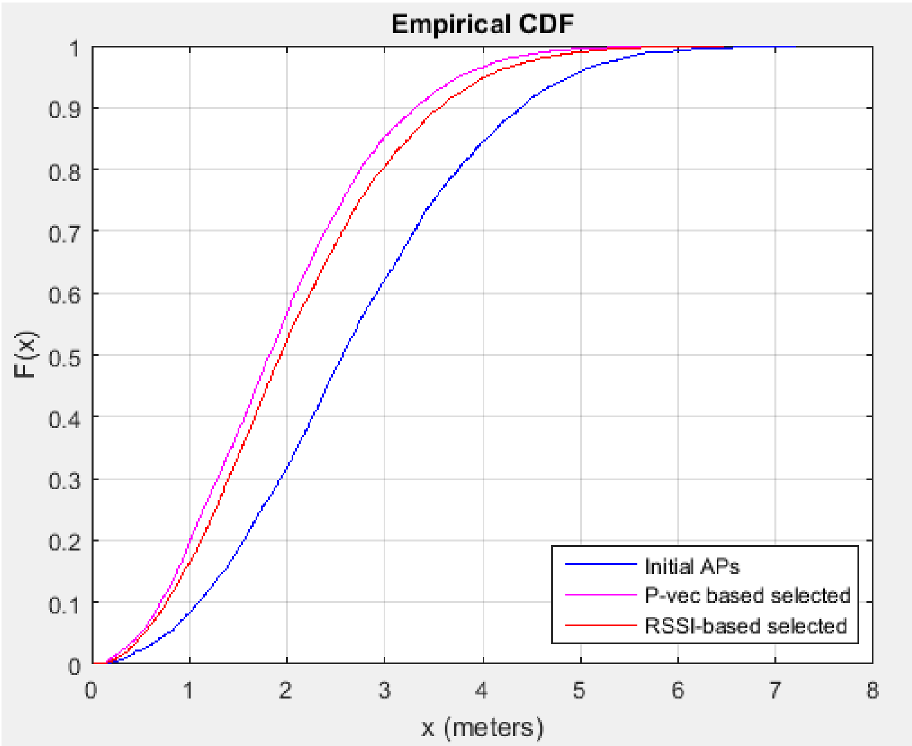

Figure 4.5: Comparison of localisation error for existing and deployed APs

4.4 Performance analysis

Altering various simulation parameters to check their effect on localisation accuracy is the main concern of this section. The parameters changed were:

1) L, the gray code bit length

2) K, the nearest neighbour parameter, and

3) AP locations.

4.4.1 Changing values of L

Parameter optimisation technique like the Genetic Algorithm, requires some method of representing the parameters. One approach used in genetic algorithms is to encode an integer parameter directly in its base-2 representation, using explicit bits, and then apply a standard binary-to-integer mapping to decode the parameter value. [5] Gray coding has been used for GA in this research with the initial bit length equal to L = 8. Table 4.3 shows the effect of changing L = size[graycode]= 7, 8, 9, 10 , with keeping K = 3 constant:

L	P-vectors based	RSSI-based	Initial APs	GDOP
7	3.414	3.735	4.27	0.5218
8	3.317	3.543	4.386	0.5262
9	4.426	3.324	3.777	0.5218
10	3.312	3.209	3.493	0.522

Table 4.3: Effect of changing gray code bit length 'L' on the localsiation error and GDOP

From the table we can see that the localisation error decreases with increasing L for the variables under study i.e Pvec, RSSI and initial AP locations. There was no significant change in GDOP; the first two significant terms to the right of the decimal are the same.

4.4.2 Changing number of neighbours (K)

The value of K should be chosen carefully since the number of nearest neighbors strongly affects the performance of the system. [31] If K = 1 the classification will be done in a way whch is very sensitive to the local characteristics of the data. Conversely, if a large value of K is chosen the algorithm averages over a large number of data points and hence averages out the variability due to noise associated with the individual data points. A large value like K=10 would simply predict the most frequent set in the training data set in all cases. This choice trades off the variability associated with a low value of K against the over-smoothing associated with a high value of K. [6]

The effect of changing the K parameter from 1 to 10 while keeping L = 8 constant on the localisation error and GDOP was as shown in Table 4.4.

K	P-vectors based	RSSI-based	Initial APs	GDOP
1	3.244	4.269	3.154	0.5218
2	3.287	4.542	3.86	0.5218
3	3.85	4.436	4.214	0.522
4	3.296	4.369	3.306	0.524
5	4.242	3.952	3.777	0.522
6	3.875	4.306	3.882	0.5219
7	3.542	3.593	3.984	0.5218
8	3.608	4.255	3.535	0.5219
9	3.744	3.348	3.238	0.5218
10	3.266	3.266	3.793	0.5232

Table 4.4: Effect of changing number of Nearest Neighbours 'K' on the localsiation error and GDOP

The results show that GDOP gains the constant value of 0.5218 for K = 1, 2, 7 and 9. The localisation error for position vector based AP selection is the lowest for majority of the values of K.

4.4.3 Changing AP locations

Changing the location of an Access Point would be simply to unplug it from the wall socket it is connected and move to another place. Unfortunately, analysing all the possible AP location setups cannot be accomplished due to the large number of location setups not feasible in practical deployment. [18]

Wireless Access Points provide better accessibility, yet for this research wired Access Points are considered. The locations of Access Points were changed and the effect on CDF error was checked as shown in Table 4.5.

Total APs	Deployed	Initial	AP locations
4	1	3	Original
4	1	3	New
5	2	3	New

Table 4.5: Changing locations of Access Points

Initial locations = [2 14 8 x_1, 5 5 25 y_1]

New locations = [2 14 8 x_2, 13 13 25 y_2]

For a total of 5 APs their locations are specified as follows:

[3 7 11 z_1 z_2, 20 1 23 t_1 t_2]

Figures 4.6 - 4.11 show the deployment areas, and localisation error CDF for the following three cases:

Case I: Total 4 APs, original locations

Case II: Total 4 APs, new locations

Case III: Total 5 APs, new locations

Consider Figure 4.6, in which three initial Access Points can be seen at locations (2,5), (8,25), and (14,5). One additional AP has been placed at (7.75,4.125) using genetic algorithm optimisation of GDOP. With a total of four APs among which 3 are

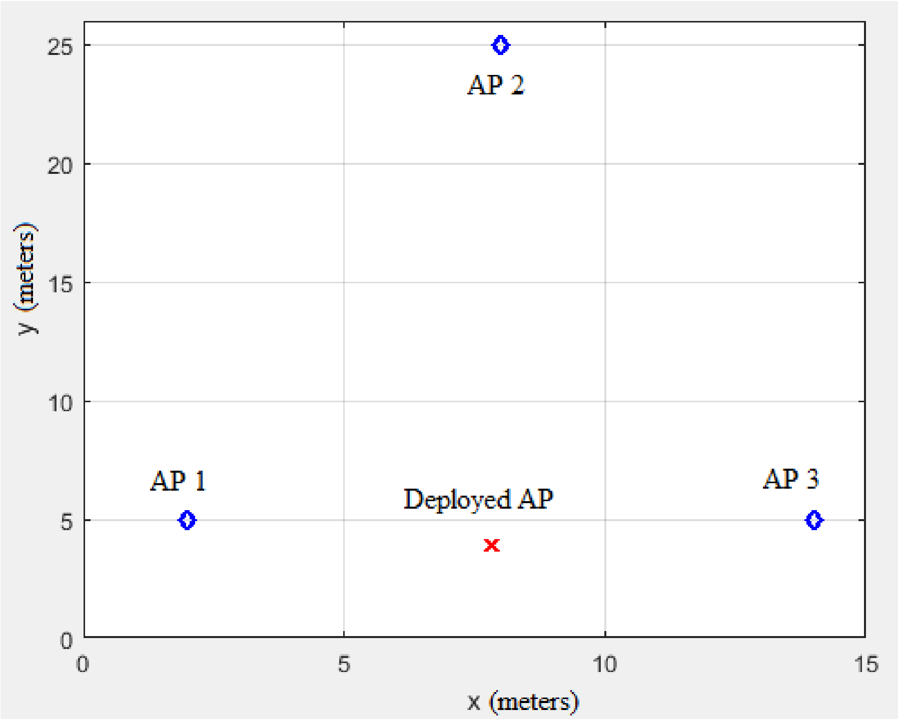

Figure 4.6: Deployment area for Case I: Total 4 APs - Initial: 3, Deployed: 1 - at original locations

initially placed and 1 additional AP has been deployed, the localisation error is 4.669 m as shown in Figure 4.7.

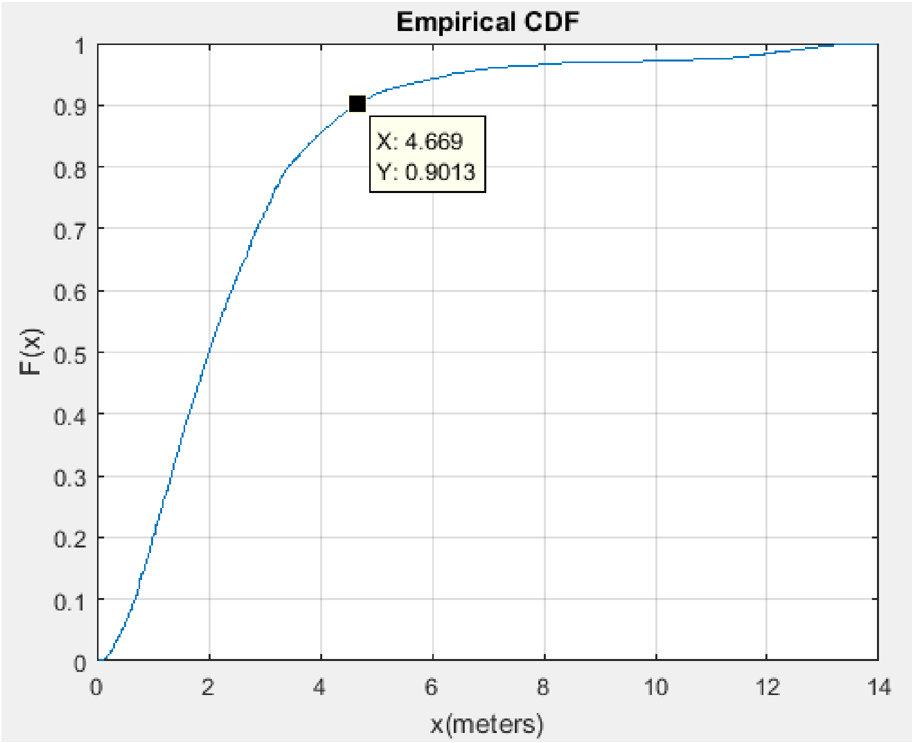

Figure 4.7: Localisation error for Case I: Total 4 APs - Initial: 3, Deployed: 1 - at original locations

With the realisation of deployment area and localisation error computed for Case 1, we move towards Case II.

Consider Figure 4.8 in which the locations of two of the initial APs have been changed to new locations. The final locations of APs for this case are now (2,13), (8,25) and (14,13). From Figure 4.8, one additional AP can be seen placed at (6.125 1.375).

For this case the localisation error has dropped to 4.581 meters at the 90^{th} percentile as shown in the CDF plot in Figure 4.9.

The result verifies the statement that an alignment of Access Points close to a regular polygon (in this case a square or diamond) leads to improvement in localisation accuracy. This has been proved in [3] with navigation satellites for GNSS.

In Figure 4.10 for Case III, the locations of all three Access Points have been changed, with the placement of two additional APs in effect. The final locations of APs are; Initial: (3,20), (7,1), (11,23), and deployed: (11.625,2.125), (6.3125,2.875).

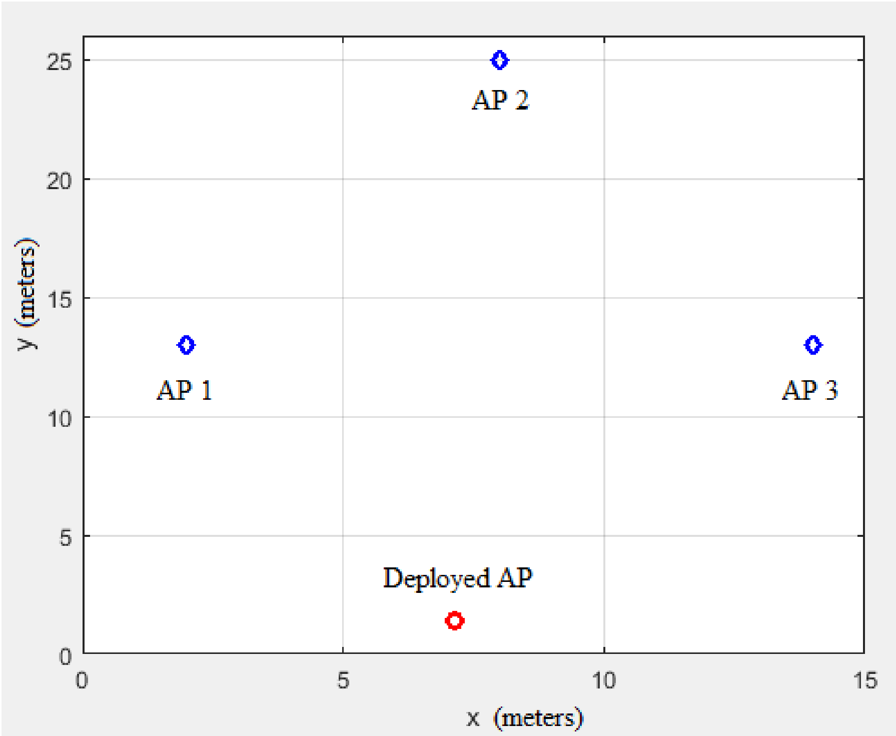

Figure 4.8: Deployment area for Case II: Total 4 APs - Initial: 3, Deployed: 1 - at new locations

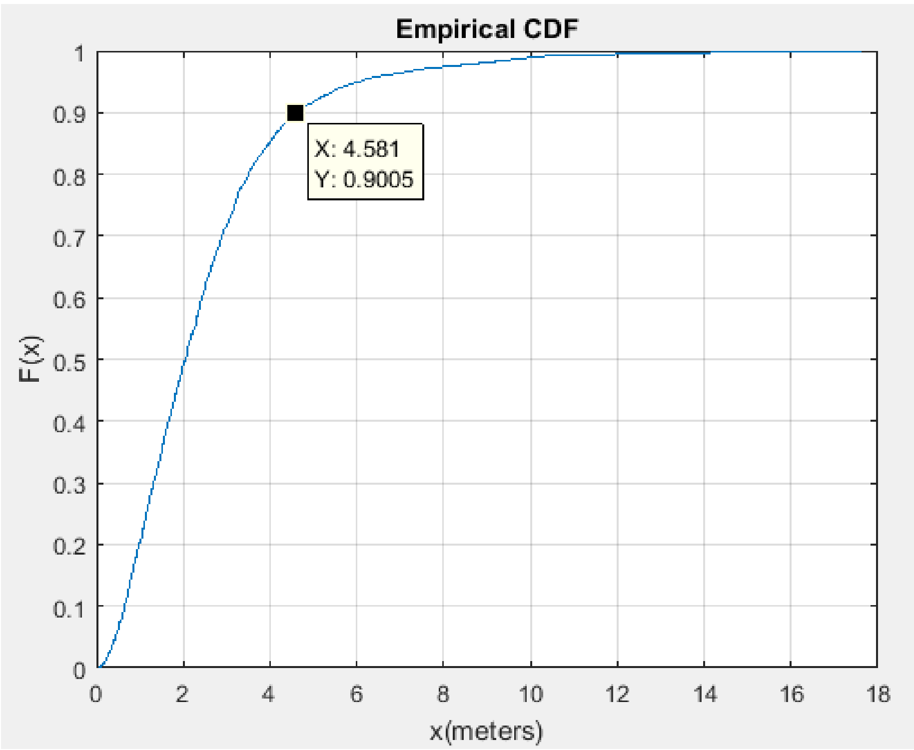

Figure 4.9: Localisation error for Case II: Total 4 APs - Initial: 3, Deployed: 1 - at new locations

The initial AP locations have been selected at random while the deployed AP locations

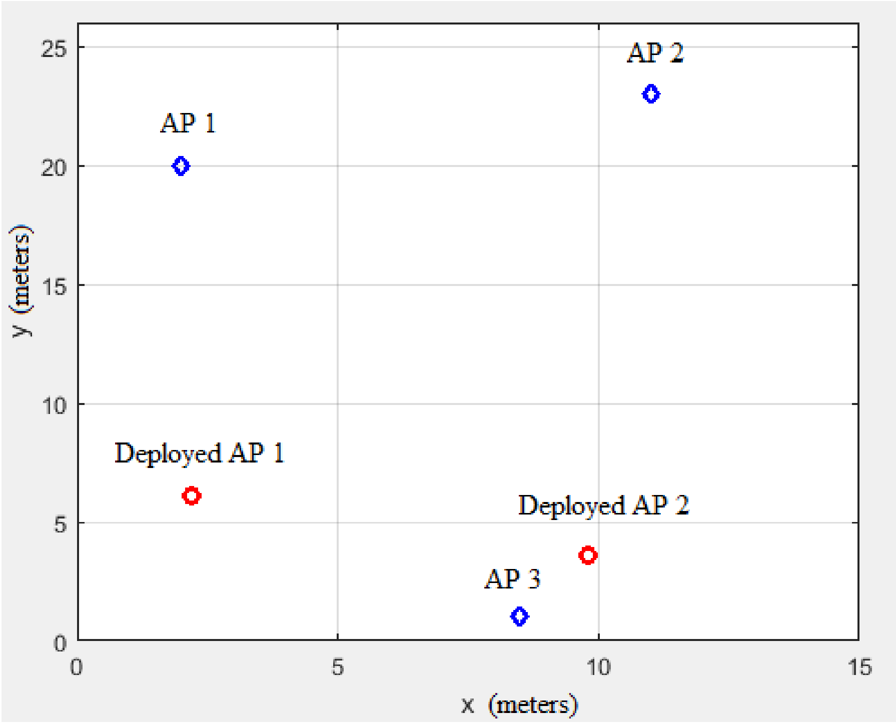

Figure 4.10: Deployment area for Case III: Total 5 APs - Initial: 4, Deployed: 2 - at new locations

are generated from the Genetic Algorithm optimisation. The new locations, however not complying by the regularisation of geometry as described in [3], have yet resulted in the lowest error for this case. The reduction in error is also due to addition of one more AP as compared to the previous cases of three initial and one additionally deloyed AP.

The localisation error for three initial and two deployed APs is shown in figure 4.11. Because of the addition of two APs, the localisation error is further reduced to 4.228 meters. While a large number of Access Points improves the localisation error, a reduction in EDOP will be possible when the APs are regularised in geometry. Also an increase in number of Access Points beyond the requirement is excessive and is better avoided for achieving lower complexity.

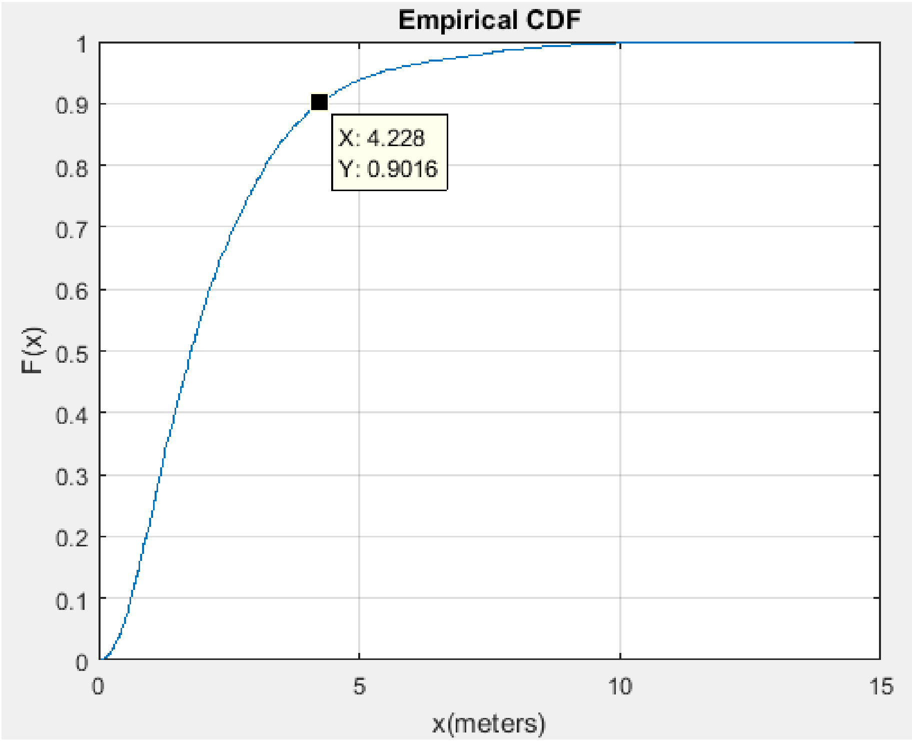

Figure 4.11: Localisation error for Case III: Total 5 APs - Initial: 4, Deployed: 2 - at new locations

Table 4.6 enlists the effect on localisation accuracy with the change in AP locations.

AP locations	CDF
(2,5),(8,25),(14,5),**(7.75,4.125)**	4.669
(2,13),(8,25),(14,13),**(6.125,1.375)**	4.581
(3,20),(7,1),(11,23),**(11.625,2.125),(6.325,2.875)**	4.228

Table 4.6: Effect of changing AP locations on localisation accuracy

From the table, when the geometry of Access Points is not regularised (close to a regular polygon), the localisation error is 4.669.

After the geometry was regularised as shown in Figure 4.8, the localisation error reduced to 4.581 for the same number of APs. This is because of a reduction in EDOP due to a better geometry which results in lower localisation error. [2].

An anomaly case was observed in Figure 10 in which three existing APs are present and two additional APs are placed at locations determined by the simulations. The aim is to check localisation error. Although the geometry formed by the Access Points is not close to a regular polygon, but the localisation error was further reduced to 4.228 because the addition of another AP makes the received RSSI vector larger and makes a better comparison with the RSSI database, making the localisation process easier. However, a rigorous increase in the number of Access Points may lead to similar and repetitive RSSI measurements which will induce errors in online RSSI measurements and should be approached with caution.

Chapter 5

CONCLUSION AND FUTURE WORKS

An RSSI database was used as the framework for GDOP based AP deployment in this research. Localisation of users was carried out as a comparison of three AP deployment scenarios: without the deployment or additional APs, with the deployment and selection of APs based on RSSI and lastly with AP selection based on user-to-AP position vectors formed after the deployment.

A discrepancy was observed in the cumulative distributions of the three cases, with the user-to-AP position-vector based AP selection giving the least error as compared to either the RSSI-based selection or localisation using only the existing number of APs. This is a productive result, since the 3-tuple numerical set of APs is dynamically optimised in accordance with proximity to each of the localisable user in signal space (in case of RSSI-based selection) or the physical space (in case of position-vector based selection), with an improvement in localisation accuracy observed in the latter case. Furthermore, a performance analysis showed the variation in localisation error with the change in simulation parameters like K, L and AP locations. The performance was best with 3 initial and 2 deployed APs with new AP locations. Hence the more regular the geometry formed by APs was, the lesser was the localisation error. This can be attributed to a better signal reception and lesser time delay from one AP to another AP. This work can be extended to anomaly cases where one or more of the APs become disfunctional due to any reason and the signal strengths from the remaining set are left to localise the users, in which a need for selection mechansim can arise for the APs.

To summarise, indoor location systems have originated in various forms and techniques among which Wi-Fi based systems are robust and cost-effective because they require little modification in the existing systems.

Localisation error reduction has been the focus of many research works and improving the error significantly with little modification in the existing system has a major impact on such system's efficiency. Nowadays, hand-held mobile devices can easily allow interfacing with Wi-Fi technologies and in indoor environments where location estimate is the main concern, a Wi-Fi coverage is almost readily available. Such systems when crowned with an intelligent deployment of Access Points not only increase the localisation accuracy but also reduce the overall cost requirement when combined with a hybrid approach towards Access Point selection at the users' end.

BIBLIOGRAPHY

[1] Du, Y., Yang, D., Xiu, C., Huang, Z. and Luo, H.,"An Optimization Model for the Additional AP Placement in the Existing Wi-Fi Indoor Positioning System", *Proceedings of the ION 2013 Pacific PNT Meeting*, Honolulu, Hawaii, Apr 2013, pp. 605-609.

[2] Nuria Blanco-Delgado, Fernando D. Nunes and Gonzalo Seco-Granados," Relation between GDOP and the Geometry of the Satellite Constellation", *1st International Conference on Localization and GNSS*, 2011.

[3] M. Mironovova and H. Havlis, "Calculation of GDOP Coefficient", *CTU in Prague: Faculty of Electrical Engineering*, 2011.

[4] P. Bahl and Venkata N. Padmanabhan,"User location and tracking in an in-building radio network", *Technical Report, Microsoft Research*, 1999.

[5] U. K. Chakraborty and C. Z. Janikow, "An Analysis of Gray versus Binary Encoding in Genetic Search",*University of Missouri*, 2003.

[6] K. M. Leung,"k-Nearest Neighbor Algorithm for Classification",*Polytechnic University*, Nov 2007.

[7] D. Li, B. Zhang and C. Li,"A Feature-Scaling-Based k-Nearest Neighbor Algorithm for Indoor Positioning Systems",*IEEE Internet of Things Journal*, vol. 3, No. 4, Aug 2016.

[8] Santiago J. Flores Asenjo,"Caracterizacin del canal radio mvil en el interior de edificios con mltiples plantas mediante tcnicas de lanzado de rayos", *PhD. Thesis. Department of Communications, Polytechnic University of Valencia*, Spain, May 1998.

[9] H. Liu, H. Darabi and P. Banerjee, J. Liu," Survey of Wireless Indoor Positioning Techniques and System",*IEEE Transanctions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 37, No. 6, Nov 2007.

[10] J. Bardwell," Converting Signal Strength Percentage to dBm Values",*Wild packets (Article)*, Nov 2002.

[11] Ishan Agrawal,"Indoor localization of mobile users",*CA report*, ISHAN.

[12] ABI Research,"Indoor location in retail: Where is the money", *Research: Location Technologies Market Research*, May 2015. [Online]. Available: https://www.abiresearch.com/market-research/service/location-technologies/

[13] S. He and S.-H. Gary Chan,"Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons",*IEEE Communications Surveys and Tutorials*, Vol. 18, No. 1, First quarter 2016.

[14] [Online] https://www3.aerohive.com/solutions/technology/location-based-services.html

[15] B. Bobescu and M. Alexandru,"Mobile Indoor Positioning using Wi-Fi Localisation",*Review of the Air Force Academy*, No.1 (28), 2015.

[16] Omar Costilla Reyes,"Dynamic Wi-Fi Fingerprinting Indoor Positioning System",*Thesis Prepared for the Degree of Master of Science, University of North Texas*, Aug 2014.

[17] C. Feng,W. Sy Anthea Au, S. Valaee and Z. Tan,"Compressive Sensing Based Positioning Using RSS of WLAN Access Points",*State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University*, Project No.863 (2007AA01Z277).

[18] K. Farkas, . Huszk and G. Gdor,"Optimization of Wi-Fi Access Point Placement for Indoor Localization",*ISSN: 1339-147X*, Vol. 1, Issue 1, Jul 2013.

[19] Americas Headquarters, Cisco Systems, Inc.,"Wi-Fi Location-Based Services 4.1 Design Guide",*www.cisco.com*,May 20, 2008

[20] L. Jiang,"A WLAN finf=gerprinting based indoor localisation technique",*Thesis presented to the Faculty of The Graduate College, University of Nebraska*, Jul 2012.

[21] B. Pricope,"Positioning using terrestrial wireless systems",*PhD dissertation submitted to department of Electrical Engineering, Jacob's University*, Mar 2013.

[22] S. Sendra, P. Fernandez, C. Turro and J. Lloret,"IEEE 802.11a/b/g/n Indoor Coverage and Performance Comparison",*ISSN: 1348-146X*, Apr 2014.

[23] Ruchus Wireless Inc,"Best practice guide: Wireless Mesh",*Ruchus Wireless Inc*, 2010.

[24] [Online] www.procurve.com"Planning a Wireless Network",*Hewlett-Packard Development Company*, 2006.

[25] G. Ochoa"Error Thresholds and Optimal Mutation Rates in Genetic Algorithms",*University of Sussex*, Dec 2000.

[26] J.G Hwang, K. E. Lee,"Enhanced Indoor Positioning Method based on IEEE 802.11 RSSI Considering DOP in Building Environments",*ICDT 2016 : The Eleventh International Conference on Digital Telecommunications*, 2016.

[27] A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, and D. Wallach"Robotics based location sensing using wireless ethernet",*Proceedings of the Tenth ACM International Conference on Mobile Computing and Networking (MOBICOM)*, 2002.

[28] L. Aalto, N. Gothlin, J. Korhonen, and T. Ojala"Bluetooth and WAP push based location-aware mobile advertising system",*MobiSYS 04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 4958. ACM Press, 2004

[29] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan"The cricket location support system",*Mobile Computing and Networking*, pages 3243, 2000.

[30] A. Ward, A. Jones, and A. Hopper"A new location technique for the active office",*IEEE Personnel Communications, 4(5)*,pages 4247, Oct. 1997.

[31] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose and D. Towsley"Facilitating Access Point Selection in IEEE 802.11 Wireless Networks",*Internet Measurement Conference*, 2005.

[32] N. L. Dortz, F. Gain, P. Zetterberg,"Wi-Fi Fingerprint Indoor positioning System using Probability Distribution Comparison",*ICASSP*, 2012.

[33] K. Grski, M. Groth, L. Kulas,"A Multi-Building WiFi-Based Indoor Positioning System",*Gdansk University of Technology*, 2013.

[34] L. Mainetti, L. Patrono, I. Sergi,"A Survey on Indoor Positioning Systems",*University of Salento*, 2010.

[35] Liu Y, Yang Z, Wang X et al."Location, Localization, and Localizability",*Journal of Computer Science and Technology*, 25(2): 274-297, Mar 2010.

[36] V. Otsason, A. Varshavsky, A. LaMarca, and Eyal de Lara"Accurate GSM Indoor Localization",*Springer-Verlag Berlin Heidelberg*, 2005.

[37] B. Huang, Z. Yao, X. Cui, and M. Lu"Dilution of Precision Analysis for GNSS Collaborative Positioning",*IEEE Transactions on Vehicular Technology*, 2013.

[38] Chafee J., Abel J."GDOP and the Craner-Rao bound",*Position Location and Navigation Symposium, IEEE*, p:663-668, 1994.

[39] Chafee J., Abel J."GDOP and the Craner-Rao bound",*Position Location and Navigation Symposium, IEEE*, p:663-668, 1994.

[40] C. Sheng, Chen, K. S. Chen, J. Fa, Huang, and Y. Ruie, Li"Using Genetic Algorithms to Approximate Weighted Geometric Dilution of Precision",*International Symposium on Computer, Consumer and Control*,2016.

[41] T. Rappaport"Wireless Communications: Principles and Practice",*Prentice Halll*,3rd edition.

[42] K. Kaemarungsi and P. K. Chrysanthis"PhD thesis entitled Design of Indoor Positioning system based on Location Fingerprinting Technique",*University of Pittsburgh*,2005.

[43] R. B. Langley"Dilution of Precision",*University of New Brunswick*, May 1999.