TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

MAJOR PROJECT REPORT

ON

**INDOOR RADIO LOCALIZATION AND NAVIGATION SYSTEM DESIGN**

(EG777EX)

By:


AMIT GHIMIRE(069/BEX/402)

KALYAN BHETWAL(069/BEX/415)

KALYAN PANDEY(069/BEX/416)

RAJAT BHATTARAI(069/BEX/432)


A PROJECT SUBMITTED TO THE DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING IN PARTIAL FULLFILLMENT OF THE REQUIREMENT FOR THE BACHELOR'S DEGREE IN ELECTRONICS & COMMUNICATION ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

LALITPUR, NEPAL

AUGUST,2016

## LETTER OF APPROVAL

The undersigned hereby certify that they have read, and recommended to the Institute of Engineering for acceptance, this project report entitled "Indoor Radio Localization and Navigation System Design" submitted by Amit Ghimire, Kalyan Bhetwal, Kalyan Pandey and Rajat Bhattarai in partial fulfilment of the requirements for the Bachelor’s Degree in Electronics & Communication Engineering.

Supervisor
Dr. Nanda Bikram Adhikari
Department of Electronics & Computer Engineering,
Institute of Engineering, Pulchowk Campus,
Tribhuvan University, Nepal

Internal Examiner
Name of Internal Examiner
Title, Affiliation

External Examiner
Name of External Examiner
Title, Affiliation

Dr. Nanda Bikram Adhikari
Deputy Head
Department of Electronics & Computer Engineering,
Institute of Engineering, Pulchowk Campus,
Tribhuvan University, Nepal

Dr. Dibakar Raj Pant
Head
Department of Electronics & Computer Engineering,
Institute of Engineering, Pulchowk Campus,
Tribhuvan University, Nepal

DATE OF APPROVAL:

# COPYRIGHT

## ACKNOWLEDGEMENT

Firstly, we would like to express our gratitude to Department of Electronics and Computer Engineering, Pulchowk Campus for providing us with the platform and opportunity for successful completion of the project.

We are indebted to Asst.Prof.Dr Nanda Bikram Adhikari for his guidance and valuable suggestions.He has always been kind and generous to us,helping in each and every steps of the project without whom it would have been very difficult for us to proceed forward. He has been a source of inspiration for us.

We are also thankful to Asst.Prof.Dr.Surendra Shrestha and Mr.Dinesh Baniya Khestri for helping us throughout the project.

Finally, we appreciate all the valuable suggestions provided to us by teachers, friends and family about the project. We have great respect for their continuous encouragement and support.

## SUMMARY

Positioning in outdoor environment has been successful with advent of GNSS. GNSS has been de-facto standard for outdoor positioning. Indoor positioning system is still in process of development. This project presents a system for tracking people or objects having Wi-Fi enabled mobile device in an area where three or more wireless access points are present.

We have implemented two approaches i.e. trilateration and RSSI fingerprinting.A detailed discussion about relative advantage between them on the basis of system complexity and accuracy is included in this report.

In trilateration based approach we estimate the position of mobile device from known AP using Log distance Path Loss Model.

The RSSI fingerprint approach can be viewed into two parts. Offline finger printing phase and localization phase.

In offline training phase the signal strength of stationary Access point in known location is measured. These values are sorted in data base creating a table.

During localization phase, position is estimated by applying WKNN on RSSI fingerprint data created in offline training phase. The accuracy of the system is further improved using Sequential Monte Carlo Method by fusing RSSI and motion sensor data.

The proposed system can be used for Navigation purpose inside indoor complex structure. This system has been developed and tested for ground floor of KOICA-ICTC building in Pulchowk. Users with android phones can view their position in a map in web application.

It will also be very helping in designing the location aware system.

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AP | Access Point |
| API | Application Programming Interface |
| CDMA | Code Division Multiple Excess |
| FCC | Federal Communication Commission |
| GCP | Ground Control Points |
| GDAL | Geospatial Data Abstraction Library |
| GNSS | GLobal Navigation Satellite System |
| GSM | Global System for Mobile |
| GP | Gaussian Process |
| GPR | Gaussian Process Regression |
| HTML | Hyper Text Markup Language |
| HMM | Hidden Markov Model |
| IEEE | Institute of Electrical and Electronics Engineering |
| IMU | Inertial Measure Unit |
| IPS | Indoor Positioning System |
| JS | Java Script |
| JSON | Java Script Object Notation |
| LQI | Link Quality Indication |
| MEMS | Micro Electro-Mechanical Sensor |
| MAC | Medium Access Control Address |
| MATLAB | Matrix Laboratory |
| PLE | Path Loss Exponent |
| QGIS | Quantum Geographical Information System |
| REST | Representational State Transfer |
| RFID | Radio Frequency Identification |
| RF | Radio Frequency |
| RMS | Root Mean Square |

| | |
|---|---|
| RSSI | Received Signal Strength Indicator |
| RSS | Received Signal Strength |
| RP | Reference Point |
| SSID | Service Set Identifier |
| TIFF | Tagged Image File Format |
| WiFi | Wireless Fedility |
| WkNN | Weighed k Nearest Neighbor |
| WLAN | Wireless Local Area Network |
| UWB | Ultra Wide Band |
| VLSI | Very Large Scale Integration |

# LIST OF SYMBOLS

| | |
|---|---|
| $\sigma_f$ | signal variance |
| $\sigma_n$ | noise variance |
| $w_t^{[m]}$ | weight of mth particle |
| $Z_t$ | sensor Measurement |
| $g$ | acceleration due to gravity$(9.8m/s^2)$ |
| $\theta$ | orientation |

# 1. INTRODUCTION

## 1.1. Background

An Indoor Positioning system is a system to locate object or people inside a building using radio waves, magnetic fields, acoustic signals, or other sensor information collected by mobile devices. An IPS system use different technologies including distance measurement to nearby anchor nodes. They either actively locate mobile devices and tags or provide ambient location or environmental context for devices to get sensed. Use of radio waves for indoor positioning and navigation system is one of the cost efficient and easy to implement technique.

The properties of radio signals like Angle of Arrival, Time of Arrival and RSSI value can be used to develop RF based IPS. Most building nowadyas, where IPS is required, contain Wi-Fi access points. The RSSI of the Wi-Fi signals from those APs can easily be retrieved. Hence, the use of RSSI value of Wi-Fi signal proves to be easy and cost effective solution for IPS.

## 1.2. Problem Statement

Locating peoples and objects inside buildings has emerged as technology having huge market potential with lots of challenges. While GNSS works for outdoor location estimates, no appropriate technology has yet been devised for the indoor scenario.
When we have to position and localize objects in an indoor environment, GNSS miserably fails. GNSS makes use of satellite signals for position estimates. These satellite signals are too weak to penetrate building walls. Hence, those signals suffer a great deal of attenuation thus rendering GNSS unsuitable for indoor localization. Another reason the GNSS is unsuitable for indoor localization is that the radio wave propagation characteristics of GNSS is highly altered by objects like trees, walls, metal furniture as well as the movement of people which are most common in an indoor environment.

Thus, people have been actively researching to find a suitable solution for this problem. The complex nature of indoor environment poses several unique problems and challenges in developing an accurate positioning system that can locate objects within a few meters precision. In recent times,distance estimation based on radio wave signal strength measurement based indoor localization technique has emerged as one of the potential technology for indoor localization.

### 1.3. Motivation

In this present age of automation, location aware devices in indoor environments are becoming increasingly important. The performance of positioning systems for outdoor environments has been excellent with the invent of GPS but many applications (like navigation) targeting mass market require efficient positioning ability in all environments.Therefore indoor positioning has become a field of active research.

Following the performance of satellite based location services, outstanding examples of applications using outdoor positioning are developed. A very good example is the navigation service in Google Maps application which helps user reach their destination in map(also showing their current location).The vision of an similar, even better, application efficiently functional in indoor environment has been the major motivating factor for our project.

### 1.4. Scope of Work

Our project aims in building an android application using which a mobile can be located in an indoor environment. The application takes information from required sensors(Wi-Fi signal receiver, accelerometer and magnetometer) and sends those data to server which calculates the position of the device in an indoor map by processing these data and returns the position to the device. The user interface is a web application where the map of the indoor environment is loaded and position of the mobile phone is tracked using a marker.

## 2. OBJECTIVES

The objective of our work is to develop a system for indoor navigation using existing infrastructure i.e WLAN and Smart phones.

Following are some of the major objectives:

1. Fully understand the propagation of radio signal in a cluttered indoor environment.

2. Study about radio propagation nature of AP.

3. Create a WiFi heat map and RSSI finger print data base.

4. Analyze accuracy,complexity,reliability and cost effectiveness of the proposed system.

5. Compare trilateration, RSS based fingerprint system and ’RSS + motion’ sensor on the basis of error metrics and tracking efficiency.

# 3. THEORY

## 3.1. Wireless Technologies for Indoor Localization

The wireless technology used for indoor localization can be classified by the frequency it uses. Since the frequency of the wireless technology affects its abilities like coverage, wall penetration, and resistance to obstacles. In this paper, we classify them into 3 categories: long distance wireless technology, middle distance wireless technology, short distance technology.

### 3.1.1. Long Distance Wireless Technologies

FM (Frequency Modulation) is used world-wide for regional radio broadcasts. In most regions, it uses 87.5 to 108.0 MHz radio spectrum. Using the VHF (very high frequency) which is far less than the WiFi and other modern wireless technology, FM is less affected by weather or obstacles like walls. Since the ubiquity of FM, there is no need to build extra beacon infrastructure using FM for indoor localization. And FM receiver is cheap and has lower power consumption hence better battery life. However, the FM station is very far away and FM has large wave length (around 3m), which means that the signal strength of FM signal does not dramatically change in short distance. Hence FM works better for large area. Since different FM stations use FDMA to share the spectrum, multiple channel signals can be used to reduce the variance or error introduced by single channel signal. GSM/CDMA has been used in cellular network communication. The GSM/CMDA frequencies in different regions are different. Generally it falls in 850MHz, 900MHz, 1800MHz, and 1900MHz bands. The GSM/CDMA network is already covered in most buildings, hence there is no or less need for extra infrastructure. Unlike FM, GSM has a relatively small propagation distance in indoor environment. However, GSM/CDMA is heavily patented, so it is hard to do modification or extensions based on GSM/CDMA which limits the future development on it.

### 3.1.2. Middle Distance Wireless Technology

WiFi is one of the most used wireless technology. It follows a series of standards in IEEE 802.11. It uses two license-exempt bands: 2.4 GHz, and 5 GHz. Most buildings like super mall or office building have already deployed WiFi hotspots that provide whole building coverage as network access point. And most commercial products, like phones, laptops and tablets, support WiFi. That means the infrastructure cost and user device cost can be very low. Additionally, WiFI based localization can be easily adopted by buildings and users. With all the advantages, WiFi is a mainstream technology in literature for indoor localization. ZigBee is a specification based on IEEE 802.15.4 standard. It uses 868 MHz band in Europe, 915 MHz band in the USA and Australia, and 2.4 GHz in other regions. ZigBee is used for long distance transmission between devices in wireless mesh network. It has low cost, low data transfer rate, short latency time, comparing to WiFi standards. In IEEE 802.15.4 standard LQI (Link Quality Indication) is defined to indicate the quality of the link and can be used to derive RSS (Received Signal Strength). And there are integrated chips (CC2430/CC2431) been manufactured to get the RSSI, which makes the implementation of the system easier.

### 3.1.3. Short Distance Wireless Technology

Bluetooth is a personal area network standard. It also uses 2.4 GHz and 5 GHz bands as WiFi does. Bluetooth is widely used for short distance communication like earphones, cell phones. Bluetooth concerns the power consumption, it uses a very low transmission power. So the coverage of bluetooth is shorter than WiFi and other WLAN technology. Hence, bluetooth is not suit for localization for large area. UWB(Ultra-Wide Band), unlike other technology, uses a sub-nanosecond radio pulse to transmit data in a wide range of bandwidth (normally greater than 500 MHz). Its transmission can be regarded as background noise to other wireless technologies, hence in theory, it can use any spectrum without interfere with other users. It uses small transmission power -41.4dBm/MHz (which is limited by FCC) meaning the power consumption is low. Another advantage of UWB is its immune to

Table 3.1: Characteristics of different wireless technologies used for indoor localization

| Wireless Technology | Range | Dedicated Infrastructure | Power Consumption | Disadvantages |
|---|---|---|---|---|
| FM | 100 km | Yes | Low | Signal changes with distance |
| GSM/CDMA | 100 m $\sim$ 10 km | Yes | Unknown | Highly patented |
| WiFi | 35 m (Indoor) | Yes | High | High variance signal |
| ZigBee | 30$\sim$60 m | Yes | Low | Need dedicated infrastructure |
| Bluetooth | 10 m | Yes | Low | Cover range is limited |
| UWB | Few meters | Yes | Low | Cover range is limited |
| RFID | 1m | Yes | Low | Cover range is limited |

multi-path problems in theory. RFID(Radio Frequency Identification) is a simple technology with a history of more than 50 years. It composes of two parts: tag and reader. The reader uses radio-frequency electromagnetic field to read the data in the tag and get the identification of the object the tag attached to. The tag can either have battery or not which makes it an active tag or a passive tag. The passive tag can be very cheap and have a long lifetime which is ideal for cost-sensitive scenario. However, the passive tag RFID suffers from both tag collision and reader collision problems. Tag collision happens when a reader reads more multiple tags, and reader collision happens when the coverage of two readers overlaps and read the tag at the same time. The communication range of RFID is very short,around (1-2m), this increase the labor works for pre-deployment to cover the a huge area.

### 3.2. Path Losses

The reduction in intensity of radio wave as it propagates away from a source is called path loss or attenuation. Path loss normally includes propagation losses caused by the natural expansion of the radio wave front in free space (which usually takes the shape of an ever-increasing sphere), absorption losses (sometimes called penetration losses), when the signal passes through media not transparent to electromagnetic waves, diffraction losses when part of the radio wave front is obstructed by an opaque obstacle, and losses caused by other phenomena.

The signal radiated by a transmitter may also travel along many and different paths to a receiver simultaneously; this effect is called multipath. Multipath waves combine at the receiver antenna, resulting in a received signal that may vary widely, depending on the distribution of the intensity and relative propagation time of the waves and bandwidth of the transmitted signal. The total power of interfering waves in a Rayleigh fading scenario vary quickly as a function of space (which is known as small scale fading). Small-scale fading refers to the rapid changes in radio signal amplitude in a short period of time or travel distance.

In a communication system, a received signal will differ from the transmitted signal due to various transmission impairments. The most significant transmission impairments for LOS transmission are [14]:

### 3.2.1. Attenuation

The strength of a signal falls off with distance over any transmission medium. This reduction in strength or attenuation is logarithmic for guided media. Whereas attenuation is a more complex function of distance and the makeup of the atmosphere for an unguided media.

### 3.2.2. Free space loss

In any wireless communication, the signal disperses with distance. A receiving antenna will receive less signal power the farther it is from the transmitting antenna. Assuming all the sources of impairments are nullified the transmitted signal attenuates over distance because the signal is being spread over a larger and larger area. This form of attenuation is known as free space loss.

### 3.2.3. Fading

Fading refers to the time variation of received signal power caused by changes in the transmission medium or path. Fading is the most challenging technical problem in designing a communication system. In a fixed environment, fading is affected by changes in atmospheric conditions. Whereas in a mobile environment where either the receiving or transmitting antenna is in motion relative to the other, the relative location of various obstacles changes with time, causing complex transmission effects.

### 3.2.4. Multipath

Multipath is defined as a propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths. The direct and reflected signals are often opposite in phase, which can result in a significant signal loss due to mutual cancelation in some circumstances. Depending on the differences in the path lengths of direct and reflected waves, the composite signal can be either larger or smaller than the direct signal. Multipath is most troublesome indoors and in areas where many metallic surfaces are present. Multipath is caused by the following propagation mechanisms:

#### 3.2.4.1. Reflection

Reflection occurs when a propagating electromagnetic wave impinges upon an object which has very large dimensions when compared to the wavelength of the propagating wave. Reflections occur from the surface of the earth and from buildings and walls. The reflected waves may interfere constructively or destructively at the receiver.

#### 3.2.4.2. Diffraction

Diffraction occurs when the radio path between the transmitter and receiver is obstructed by a surface that is large compared to the wavelength of the radio wave. The secondary waves resulting from the obstructing surface are present throughout the space and even behind the obstacle, giving rise to a bending of waves around the obstacle, even when a line-of-sight path does not exist between transmitter and receiver.

#### 3.2.4.3. Scattering

Scattering occurs when the medium through which the wave travels consists of objects with dimensions that are small compared to the wavelength, and where the number of obstacles per unit volume is large. Scattered waves are produced by rough surfaces, small objects, or by other irregularities in the channel.

These three propagation effects influence system performance in various ways depending on local conditions and as a mobile unit moves through the medium. Diffraction and scattering are generally minor effects if there is a clear LOS between transmitter and receiver although reflection may have a significant impact. In cases where there is no LOS, diffraction and scattering are the primary means of signal reception.

### 3.2.5. Refraction

Refraction is defined as a change in direction of an electromagnetic wave resulting from changes in the velocity of propagation of the medium through which it passes. This may result in a situation in which only a fraction or no part of the line of sight wave reaches the receiving antenna.

### 3.2.6. Noise

In any transmission event, a received signal will consist of the transmitted signal, modified by various distortions imposed by the transmission medium, plus additional unwanted signals that are inserted by the medium. These unwanted signals are referred to as noise or interference. Noise is the major limiting factor in any communications system performance.

### 3.2.7. Atmospheric absorption

Atmospheric absorption is an additional loss due to the presence of different atmospheric elements such as water vapor and oxygen etc. A peak attenuation occurs in the vicinity of 22 GHz due to water vapor. At frequencies below 15 GHz, the attenuation is less.

## 3.3. Path Loss Factors

Path loss in an indoor radio channel are caused mainly by transmission loss through walls, floors and other obstacles. The reflection from, and diffraction around, objects (including walls and floors) and motion of persons and objects within the room also give rise to impairments in propagation of radio signals. Also, there is temporal and spatial variation of path loss due to these factors.

### 3.3.1. Partition Loss

Buildings have a wide variety of partitions and obstacles which form the internal and external structure. Houses typically use a frame partition to form internal walls and have wood or concrete between floors. Office buildings on the other hand often have large open areas which are constructed so that the space maybe reconfigured easily and use concrete between floors. For indoor communication, the construction materials that make up the obstructions are the largest attenuators.

### 3.3.2. Building Entry Losses

Building entry loss is the additional loss due to a terminal being inside a building. The building shadowing loss is the difference between the median of the location variability of the signal level outside the illuminated face of a building and the signal level outside the opposite face of the building at the same height above ground, with multi path fading spatially averaged for both signals. It can be considered as the transmission loss through a building. Indoor propagation characteristics are affected by reflection from and transmission through the building materials. The reflection and transmission characteristics of those materials depend on the complex permittivity of the materials. Site-specific propagation prediction models may need information on the complex permittivity of building materials and on building structures as basic input data.

### 3.3.3. Other Factors

In addition to the fundamental building structures, there are many other factors contributing to path loss inside a building.

1. Furniture and other fixtures
2. Transmitter and receiver siting
3. Orientation of mobile terminal
4. movement of objects in the room
5. presence and movement of people in the room,etc.

## 3.4. Propagation Models

A propagation model is a set of mathematical expressions ,diagrams, and algorithms used to represent the radio characteristics of a given environment [13]. The prediction models can be either empirical (also called statistical) or theoretical (also called deterministic), or a combination of these two. While the empirical models are based on measurements, the theoretical models deal with the fundamental principles of radio wave propagation phenomena.

In the empirical models, all environmental influences are implicitly taken into account regardless of whether they can be separately recognized. This is the main advantage of these models. Because deterministic models are based on the principles of physics they may be applied to different environments without affecting the accuracy. In practice, their implementation usually requires a huge database of environmental characteristics, which is sometimes either impractical or impossible to obtain. The algorithms used by deterministic models are usually very complex and lack computational efficiency. For that reason, the implementation of the deterministic models is commonly restricted to smaller areas of microcell or indoor environments.

### 3.4.1. ITU Path Loss Model

The ITU model for site-general indoor propagation path loss prediction is[13]:

$$L_{dB} = -20log_{10}(f) + Nlog_{10}(d) + LF(n) - 28dB \tag{3.1}$$

where, N is the distance power loss coefficient
f is the frequency in MHz
d is the distance in meters (d>1 m)
Lf(n) is the floor penetration loss factor
n is the number of floors between the transmitter and the receiver

### 3.4.2. Log-Distance Path Loss Model

Log distance path loss model is a generic model and an extension to Friis Free space model. It is used to predict the propagation loss for a wide range of environments. It is given by:

$$L_{total} = PL(d_0) + 10nLog_{10}(\frac{d}{d_0}) + X_s dB \tag{3.2}$$

Where, $PL(d_0)$ is the path loss at the reference distance, usually as( theoretical) free-space loss in 1 m.
n = Path Loss exponent. See the table below that gives the path loss exponent for various environments.
$X_s$ is a zero-mean Gaussian distributed random variable (in dB) with standard deviation $\sigma$.

This variable is used only when there is a shadowing effect. If there is no shadowing effect, then this variable is zero. Taking log of the Normal (Gaussian)-variable results in the name ”Log-Normal” fading.

To model real environments, usually the shadowing effects cannot be neglected. If the shadowing effect is neglected, the Path Loss is simply a straight line. To add shadowing effect a zero-mean Gaussian random variable with standard deviation âĂŞ$sigma$ is added to the equation. The actual path loss may still vary due to other factors. Thus the path loss exponent and the standard deviation of the random variable should be known precisely for a better modeling.

The Path Loss Exponent (PLE) table given above is for reference only. It may or may not fit the actual environment we are trying to model. PLE estimation is done by equating the observed (empirical) values over several time instants to the established theoretical values.

Table 3.2: PLE for different environments

| Environment | Path Loss Exponent(n) |
|---|---|
| Free Space | 2 |
| Urban area cellular radio | 2.7 to 3.5 |
| Shadowed urban cellular radio | 3 to 5 |
| Inside a buildings - line of Sight | 1.6 to 1.8 |
| Obstructed in building | 4 to 6 |
| Obstructed in Factory | 2 to 3 |

## 3.5. Wireless Local Area Network Technology

A wireless local area network (WLAN) is a wireless distribution method for two or more devices that use high-frequency radio waves and often include an access point to the Internet. A WLAN allows users to move around the coverage area, often a home or small office, while maintaining a network connection. In a typical WLAN infrastructure configuration, there are two basic components:

### 3.5.1. Access Points

An access point or a base station connects to a LAN by means of Ethernet cable. Usually installed in the ceiling, access points receive, buffer, and transmit data between the WLAN and the wired network infrastructure. A single access point supports on average twenty users and has a coverage varying from 20 meters in areas with obstacles (walls, stairways, elevators) up to 100 meters in areas with clear line of sight. A building may require several access points to provide complete coverage and allow users to roam seamlessly between access points.

### 3.5.2. Wireless Client Adapter

A wireless adapter connects users via an access point to the rest of the LAN. A wireless adapter can be a PC card in a laptop, an ISA or PCI adapter in a desktop computer, or fully integrated within a handheld device.

## 3.6. IEEE 802.11 Standard

IEEE 802.11 is a family of specifications for WLANs developed by the Institute of Electrical and Electronics Engineers. The 802.11 standard specifies parameters for both the physical and medium access control (MAC) layers of a WLAN. The physical layer handles the transmission of data between nodes. The MAC layer consists of protocols responsible for maintaining the use of the shared medium. Work on 802.11 began in 1987 within the IEEE 802.4 group. There are three physical layers for WLANs: two radio frequency specifications (RF direct sequence and frequency hopping spread spectrum) and one infrared. Most WLANs operate in the 2.4 GHz license-free frequency band and have throughput rates up to 2 Mbps.

### 3.7. Socket Programming

Sockets provide the communication mechanism between two computers using TCP. A client program creates a socket on its end of the communication and attempts to connect that socket to a server. When the connection is made, the server creates a socket object on its end of the communication. The client and server can now communicate by writing to and reading from the socket. Normally, a server runs on a specific computer and has a socket that is bound to a specific port number. The server just waits, listening to the socket for a client to make a connection request. On the client-side, the client knows the hostname of the machine on which the server is running and the port number on which the server is listening. To make a connection request, the client tries to rendezvous with the server on the server's machine and port. The client also needs to identify itself to the server so it binds to a local port number that it will use during this connection.

### 3.8. Trilateration

In geometry, trilateration is the process of determining absolute or relative locations of points by measurement of distances, using the geometry of circles, spheres or triangles. Trilateration has wide range of application in areas like surveying and navigation, including global positioning systems(GPS). Trilateration can be used for Indoor Positioning System as well.

In two-dimensional geometry, it is known that if a point lies on two circles, then the circle centers and the two radii provide sufficient information to narrow the possible locations down to two. Additional information may narrow the possibilities down to one unique location.

In three-dimensional geometry, when it is known that a point lies on the surfaces of three spheres, then the centers of the three spheres along with their radii provide sufficient information to narrow the possible locations down to no more than two (unless the centers lie on a straight line).

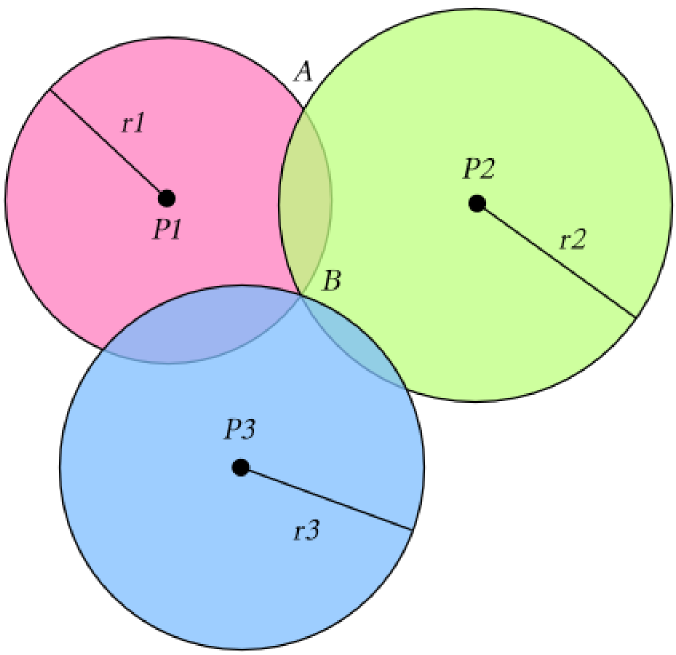



Figure 3.1: Position estimation using trilateration technique

Now, knowing r1, r2 and r3 i.e. distances from center of access points and center coordinates

$(x_i, y_i, z_i)$ of access points, the exact location of the receiver $B(x, y, z)$ can In the fig above, trilateration technique has been used to estimate the location of an object. Point P1 is Access Point 1, P2 is Access Point 2 and P3 is Access point 3. Given that we know the RSSI values of these wireless access points in the location to be determined, we can easily calculate the respective distances r1, r2 and r3 using the log distance path propagation model. The point of intersection of these three circles is the position of the receiver which is point B in the figure can be calculated solving these set of equations.

$$(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 = r1^2 \tag{3.3}$$

$$(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 = r2^2 \tag{3.4}$$

$$(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2 = r3^2 \tag{3.5}$$

These equations reduce to linear set of simultaneous equations which can be easily solved using matrices. The system gives unique solution (x,y,z) which is the required position of user in an indoor environment.

Though trilateration is a simple solution, it doesn't necessarily provide the required accuracy in indoor environment. The RSSI values of access points fluctuate considerably indoors, so there is no definite propagation model that we can apply to the radio signals. So the computed distances r1, r2 and r3 are not much reliable. So we have to adopt another technique called RSSI Fingerprinting for better accuracy.

### 3.9. MEMS

MEMS is an acronym for Micro Electro Mechanical Systems. It is used to create ultra-miniaturized devices, having dimensions from a few microns to a couple of centimeters across. These are quite similar to an IC, but with an ability to integrate moving mechanical parts on the same substrate. These are fabricated using batch fabrication process similar to a VLSI. A typical MEMS is an integrated microsystem on a chip that can incorporate moving mechanical parts in addition to electrical, optical, fluidic, chemical, and biomedical elements.Functionally, MEMS includes a variety of transudation mechanisms to convert signals from one form of energy to another.

Many types of microsensors and microactuators can be integrated with, signal processing, optical subsystems, and micro-computing to form a complete functional system on a chip. MEMS characteristic ability is to include moving mechanical parts on the same substrate. MEMS-based motion, acceleration, and stress sensors are being deployed massively. Integration of multiple functions into MEMS provides higher degree of miniaturization, lower component count, and increased reliability.

Most Android-powered devices have built-in sensors that measure motion, orientation, and various environmental conditions. These sensors are capable of providing raw data with high precision and accuracy, and are useful if you want to monitor three-dimensional device movement or positioning, or you want to monitor changes in the ambient environment near a device.

#### 3.9.1. Magnetometer

The magnetometer sensor in smart phone utilizes the modern solid state technology to create a miniature Hall-effect sensor that detects the Earth's magnetic field along three perpendicular axes X, Y and Z. The Hall-effect sensor produces voltage which is proportional to the strength and polarity of the magnetic field along the axis each sensor is directed. The sensed voltage is converted to digital signal representing the magnetic field intensity.

The magnetometer is enclosed in a small electronic chip that often incorporate another sensor, typically a built in accelerometer, that help to correct the raw magnetic measurements using tilt information from the auxiliary sensor. It is essential for detecting the relative orientation of device relative to the Earth's magnetic north.

### 3.9.2. Accelerometer

With modern MEMS technology, the sensors are easily included in miniature electronic boards.

The accelerometer can detect movement based on double integration of the measured acceleration and addition of the initial position and speed. However, since the Earth exerts a gravity acceleration on all bodies, we can also use the accelerometer to measure tilt. An accelerometer sensor reports the acceleration of the device along the 3 sensor axes. The measured acceleration includes both the physical acceleration (change of velocity) and the gravity. The measurement is reported in the x, y and z fields of $sensors\_event\_t.acceleration$.

## 3.10. Pedestrian Dead Reckoning

Dead Reckoning (DR) determines current position of user from the knowledge of the previous position and the measurements of motion direction and traveled distance

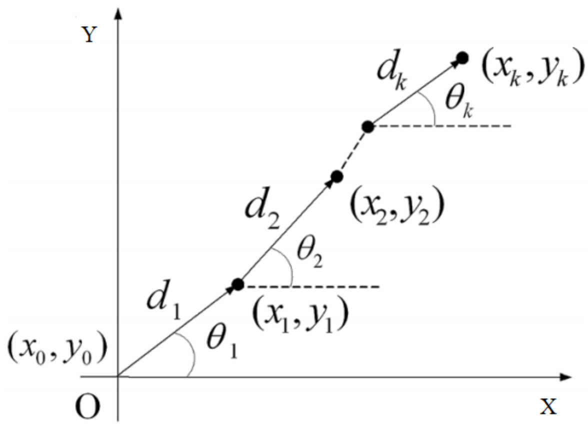



Figure 3.2: Basic Model of PDR

### 3.10.1. Step Detection

The step detection method we have used is quite simple yet powerful. The step detection algorithm is:

1. Sample acceleration at 5Hz.
2. Find rms values of acceleration at each instant.
3. Subtract the offset $g = 9.8m/s^2$ from rms value.
4. if the acceleration is greater than one standard deviation of the signal also called threshold , peak is detected.

The details about threshold selection is described in result section.

### 3.10.2. Step Length Estimation

From results of existing literature and our own experience most the time human walking shows distinct behavior. So, we have assumed the step length of 0.66 m as distance between two detected peaks. One can argue that this may cause high error in localization,but we have used multisensor fusion so that error will be minimized itself.

### 3.10.3. Orientation

The azimuth obtained from magnetometer gives the alignment with magnetic north. Calibrating the value, alignment with local x-axis is obtained.

The equations for pedestrian dead reckoning are:-

$$x_k = x_o + \Sigma_{i=1}^{k} d_i cos(\theta_i) \tag{3.6}$$

$$y_k = y_o + \Sigma_{i=1}^{k} d_i sin(\theta_i) \tag{3.7}$$

### 3.11. Gaussian Process Regression

Gaussian processes (GP) are powerful tools for probabilistic modeling purposes. They can be used to define prior distributions over latent functions in hierarchical Bayesian models. The prior over functions is defined implicitly by the mean and covariance function, which determine the smoothness and variability of the function. The inference can then be conducted directly in the function space by evaluating or approximating the posterior process. Let $D = (x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ be a set of training samples drawn from a noisy process.

$$y_i = f(x_i) + \epsilon \tag{3.8}$$

where each $x_i$ is an input sample in $\Re^d$ and each $y_i$ is a target value, or observation, in $\Re$. $\epsilon$ is zero mean, additive Gaussian noise with known variance $\sigma_n^2$. For notational convenience, we aggregate the n input vectors $x_i$ into a $d \times n$ matrix X, and the target values $y_i$ into the vector denoted y. A Gaussian process estimates posterior distributions over functions $f$ from training data $D$. These distributions are represented non-parametrically, in terms of the training points. A key idea underlying GPs is the requirement that the function values at different points are correlated, where the covariance between two function values, $f(x_p)$ and $f(x_q)$, depends on the input values, $x_p$ and $x_q$. This dependency can be specified via an arbitrary covariance function, or kernel $k(x_p, x_q)$. The kernel function depends upon the choice of user and type of we are applying for. We have chosen squared exponential, or Gaussian kernel.

$$k(x_p, x_q) = \sigma_f^2 exp(-\frac{1}{2l^2}|x_p - x_q|^2) \tag{3.9}$$

Here, $\sigma_f^2$ is the signal variance and $l$ is a length scale that determines how strongly the correlation between points drops off. Both parameters control the smoothness of the functions estimated by a GP. Because $y$ is actually a noisy observation of the true value $f(x)$, we must add a term to account for observation noise in the covariance function:

$$cov(y_p, y_q) = k(x_p, x_q) + \sigma_n^2 \delta_{pq} \tag{3.10}$$

where $\delta_{pq}$ is 1 if p = q and 0 otherwise. The covariance between the outputs is written as a function of the inputs, emphasizing the non-parametric nature of Gaussian process regression. Again for notational convenience, we rewrite the covariances as matrices and vectors, such that $K$ is a matrix of the covariances evaluated at all pairs of training points, and $k$ is a vector of the covariances between $x$ and the $n$ training inputs. This compact form allows us to rewrite equation into:

$$cov(y) = K + \sigma_n^2 I \tag{3.11}$$

We can then generate the posterior distribution over functions for arbitrary points $x_*$ given the training data $X$ and $y$:

$$p(f(x_*)|x_*, X, y) \sim N(\mu_{x_*}, \sigma_{x_*}^2) \tag{3.12}$$

and,

$$\mu_x = k^T (K + \sigma_n^2 I)^{-1} y \tag{3.13}$$

$$\sigma_{x_*}^2 = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_* \tag{3.14}$$

### 3.12. Particle Filter

Particle filter also called Sequential Monte Carlo is non-parametric implementation of Bayes Filter . Bayes Filter is mathematical approach for estimating an unknown probability density function recursively over time using incoming measurements and a mathematical process model. Particle filter is recursive and doesn't require Gaussian posterior distribution

The discussion about implementation of Particle filter is discussed in greater length in methodology section.

### 3.13. Hidden Markov Model

A hidden Markov model (HMM) is Markov Process with hidden states. Markov process are memoryless statistical model.In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but the output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore, the sequence of tokens generated by an HMM gives some information about the sequence of states.

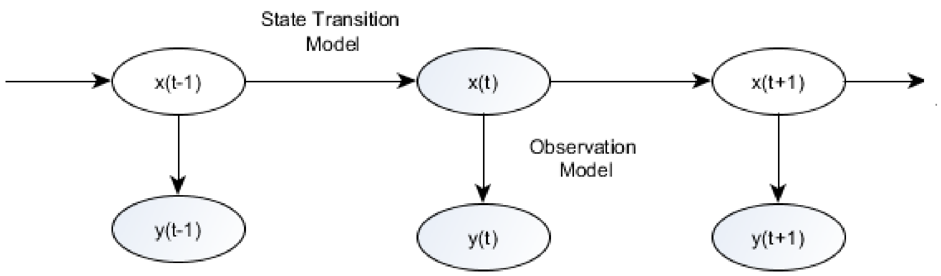



Figure 3.3: Hidden Markov Model

### 3.13.1. State Transition Model or Motion Model

The State Transition Model or Motion Model translates the current state into new new state with higher probability.The equations for state transition is given by:

$$x_t = x_{t-1} + d_t * cos(\theta_t) \tag{3.15}$$

$$y_t = y_{t-1} + d_t * sin(\theta_t) \tag{3.16}$$

Where, $d_t$ and $\theta_t$ are Gaussian Variables.

### 3.13.2. Observation Model

The observation model gives the measurement likelihood of making an observation at different positions in the indoor environment. Our observation model includes WiFi received signal strength and motion sensor readings

The WiFi signal strength measurement likelihood model uses the mean and variance of the signal at each position, calculated by Gaussian process regression.

$$p(z_t^{WiFi}|x_*) = \frac{1}{\sqrt{2\pi\sigma_{x_*}^2}} exp(-\frac{(z_t^{WiFi} - \mu_{x_*})^2}{2\sigma_{x_*}^2}) \tag{3.17}$$

where $z_t^{WiFi}$ is the received signal strength at time $t$, $\mu_{x_*}$ and $\sigma_{x_*}^2$ are the mean and variance at position $x_*$, predicted using equations 3.13 and 3.14.

Since the GP models were learned independently of each other, the resulting likelihood can become highly peaked, which results in overconfident estimates. We take this approximation into account by "smoothing" the likelihood model:

$$p(Z_{1:n}|x_t) = (\prod_{i=1}^{n} p(z_{t[i]}|x_t)^{\gamma} \tag{3.18}$$

Here, n is the number of detected access points and $\gamma \,\epsilon\, [0:1]$ plays the role of a smoothing coefficient. In our experiments we set $\gamma$ to 1/n, resulting in the geometric mean of the individual likelihoods[18].

The motion sensor based likelihood model the motion model equation 3.15 and 3.16 with PDR.

$$p(z_t^{Motion}|x_*) = \frac{1}{\sqrt{2\pi\sigma_{x_*}^2}} exp(-\frac{(z_t^{motion} - \mu_{x_*})^2}{2}) \tag{3.19}$$

where $z_t^{Motion}$ is the position calculated from the motion-sensor-based PDR, $x^*$ is the pose estimated from the motion model defined in HMM.

### 3.14. WKNN

Weighted K nearest neighbors is a version of the K-nearest neighbors (KNN) algorithm improved by introducing a weighted distance factor. If the RPs' distribution is sparse or non-uniform, after calculating the distances and selecting K nearest neighbors, the classification result of KNN is worse than that of WKNN. Torres-Sospedra's [7] extensive experimental results showed that Euclidean distance is not the optimal choice to describe the similarity between fingerprints. On the contrary, the Sorensen distance function can achieve a better accuracy than Euclidean distance. [8] The definition of Sorensen distance function is:

$$L_{(d,i)} = \frac{\Sigma_{i=1}^{n} |RSSI_{MN}^{(i,n)} - RSSI_{RP}^{(i,n)}|}{\Sigma_{i=1}^{n} |RSSI_{MN}^{(i,n)} + RSSI_{RP}^{(i,n)}|} \tag{3.20}$$

where i is a sequence of RPs, n denotes the number of APs with the same MAC between mobile node and RPs, MN denotes mobile node, RP denotes a reference point, then, the K-nearest distance will be selected as the nearest neighbors. Finally, the distances corresponding to the K neighbors are used to calculate the normalized weights using [9].

$$w_i = \frac{\frac{1}{L_{(d,i)}}}{\frac{1}{\Sigma_{i=1}^{K} L_{(d,i)}}} \tag{3.21}$$

According to the normalized weight and RPs' coordinates, a coarse estimation using the KWNN-based WiFi fingerprint can use the following formula[9]:

$$P_{i(x,y)_{KWNN}} = \{\Sigma_{i=1}^{K} x_i.w_i, \Sigma_{i=1}^{K} y_i.w_i\} \tag{3.22}$$

The value of k need to be optimized in order to increase the accuracy of the system and minimize the error.The optimization of k-value is discussed in result section.

# 4. LITERATURE REVIEW

A lot of work has been been done and many research papers are published each year in the field of indoor positioning system. Although, there are many architecture proposed for indoor positioning system RSS based approach is the most popular technique in the industry as well as in academia. This is because of the fact that we don't need any additional infrastructure for positioning. The existing WLAN structure present at various point in building is sufficient for the tracking of the users with mobile device. RSS can be of Bluetooth, RFID, Zigbee or Wi-Fi. The literature published can be divided into two major category: triangulation based and Finger print based system. We present a brief discussion of these literature below.

## 4.1. Triangulation

Triangulation uses the geometric properties of triangles to estimate the target location. It has two derivations: lateration and angulation. Lateration estimates the position of an object by measuring its distances from multiple reference points. So, it is also called range measurement techniques. Instead of measuring the distance directly using received signal strengths (RSS), time of arrival (TOA) or time difference of arrival (TDOA) is usually measured, and the distance is derived by computing the attenuation of the emitted signal strength or by multiplying the radio signal velocity and the travel time. Roundtrip time of flight (RTOF) or received signal phase method is also used for range estimation in some systems. Angulation locates an object by computing angles relative to multiple reference points.

## 4.2. Finger printing based approach

The method is based on analysis of the unique RSS value at different points in an area. There are two stages for location fingerprinting: offline stage and online stage (or run-time stage). During the offline stage, a site survey is performed in an environment. The location coordinates and respective signal strengths from nearby AP are collected. During the online stage, a location positioning technique uses the currently observed signal strengths and previously collected information to figure out an estimated location. The main challenge to the techniques based on location fingerprinting is that the received signal strength could be affected by diffraction, reflection, and scattering in the propagation indoor environments.

Our approach is based on both trilateration and RSSI fingerprinting .The comparison between both approaches on accuracy, system complexity and cost is compared.

In commercial WiFi routers it is almost impossible to accurately measure time of arrival or time difference of arrival. Also the routers don't show real time response for the round time of flight measurement. So, triangulation method isn't best approach in determining position of mobile device.

Whereas finger printing technique doesn't require any real time response from Access points, so it is easier for us to do location estimation but creating a fingerprint database is very tedious task and takes huge amount of time.

In addition to position estimation based on the fingerprinted data we have used motion sensor. The motion sensor also called micro electromechanical system (MEMS) gives us acceleration and orientation of the mobile device. RF fingerprinting techniques have the problem of signal fluctuation due to the multipath fading effect in indoor environment. The motion-sensor-based PDR approach suffers from the fact that the motion sensors equipped in the mobile device are low cost MEMS, which have relatively low accuracy. Thus, the integration drift will cause the positioning error to accumulate over time. The fusion of motion sensor and RF fingerprinting technique gives the best estimate of the position.

# 5. SYSTEM BLOCK DIAGRAM

The complete overview of our proposed system is shown in the figure 5.1.We have used 3 methodology in for localization.Trilateration and WKNN algorithm only uses rssi and strength,where as MCL method uses step detection and orientation along with wifi RSSI.

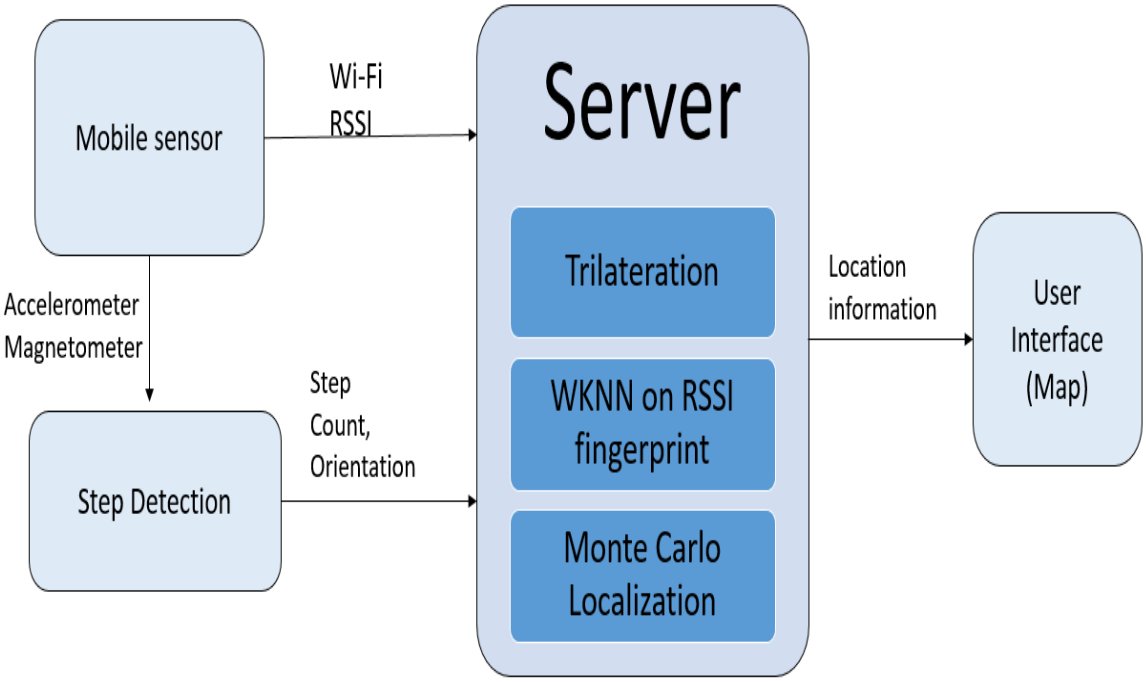



Figure 5.1: System Block Diagram

### 5.1. Monte Carlo localization

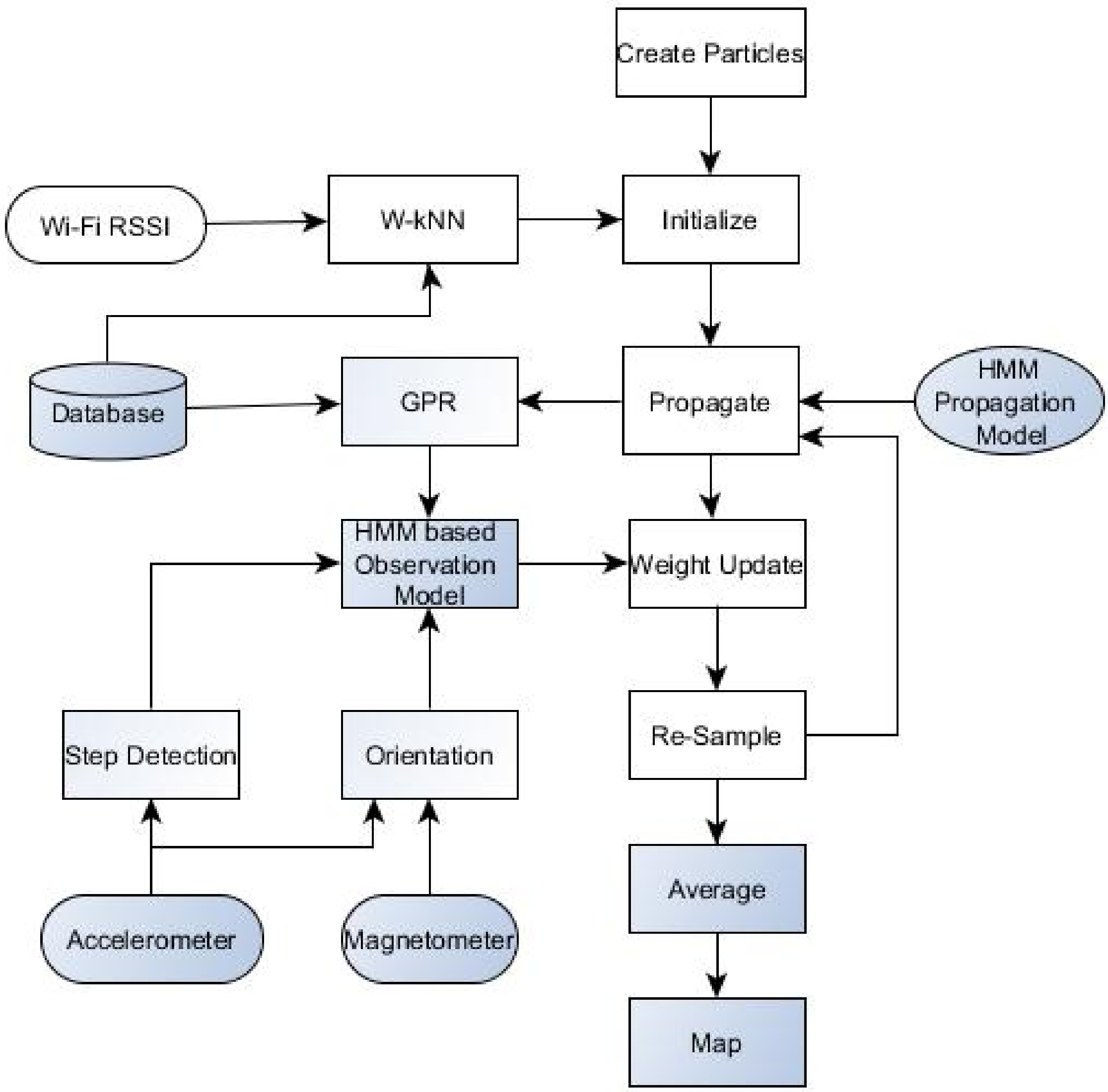



Figure 5.2: Complete Overview of System based on Monte Carlo Localization

# 6. METHODOLOGY

## 6.1. Site Selection

The first step of our project was site selection. An algorithm for measurement of RSSI of AP's was implemented in Linux Machine and Android.Upon testing, we found out that KOICA ICTC building of our campus has the most reception of many access points.

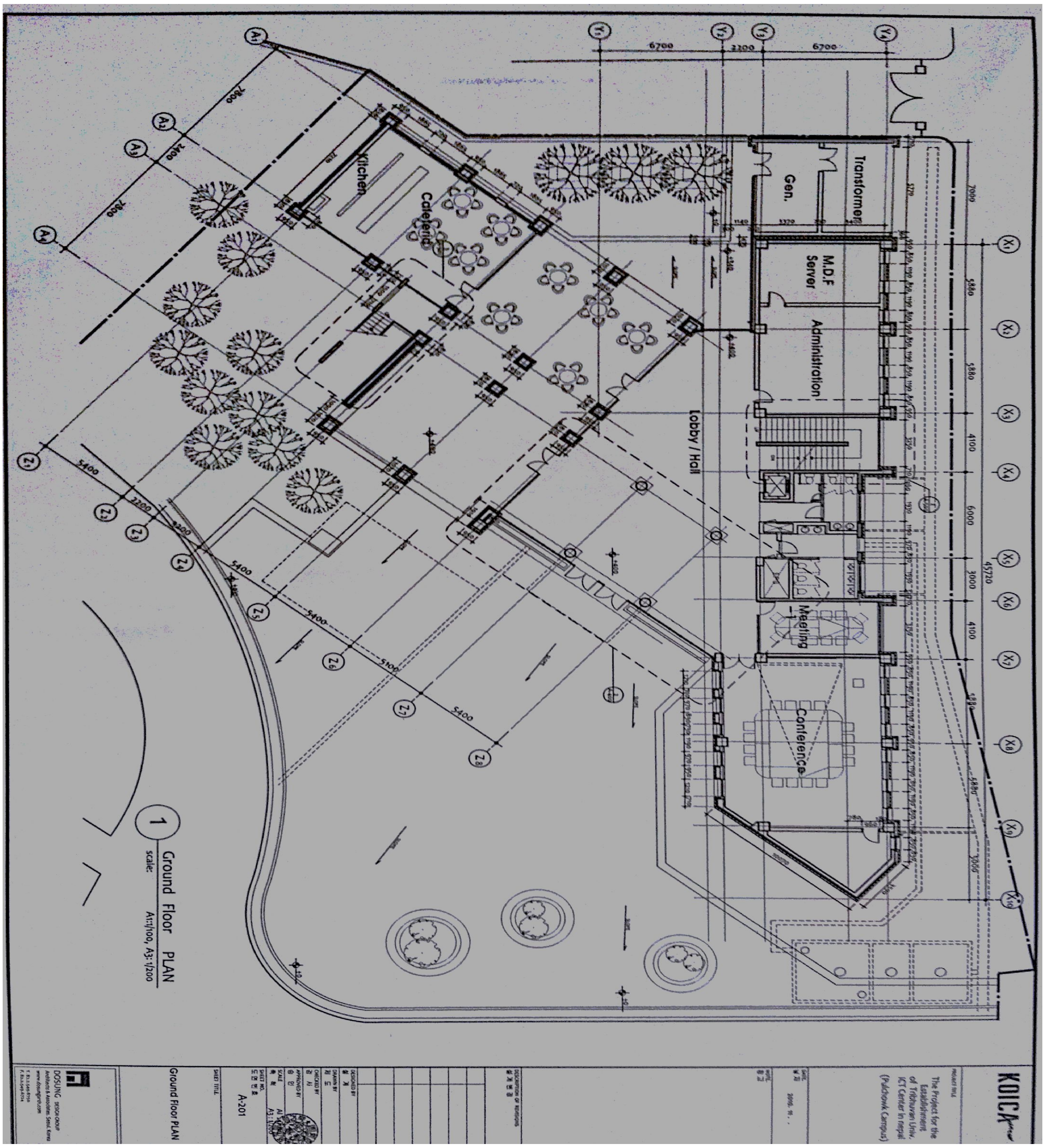



Figure 6.1: KOICA ICTC ground floor Plan

Among them, five access points which gave the good link quality and consistent rssi measurement were used for rssi fingerprinting. Two other access points were created by making wifi hotspots in our laptops and android phones making a total of seven access points for fingerprinting so that better result can be obtained. The other reason for selection of KOICA ICTC as our site is the architecture of building i.e it is composed of RC, brick structures and glasses with big open lobby and long corridors suitable for our study. Likewise, the easy availability of the floor plans of the building also prompted us to select KOICA ICTC as our site for study.The fig 6.1 shows map of the KOICA ICTC ground floor.

### 6.2. Mapping

Not much has been done in the field of indoor maps. Google maps has started indoor maps service in some selected countries which doesn't include Nepal. It works with crowd sourcing i.e. collecting floor plans from people and mapping it into the world map. We acquired floor plans of our site KOICA ICTC and started mapping.

#### 6.2.1. Drawing Floor Plans

The vector image of ground floor was drawn in vector graphics software InkScape.Fig 6.2 shows the vector map.

#### 6.2.2. Geo Referencing

Georeferencing is the process of assigning real-world coordinates to each pixel of the raster. After drawing the image, we used Quantum GIS to georeference the image. QGIS is a geographic information system application that lets data viewing, editing and analysis functionalities.

For georeferencing in QGIS, Georeferencer GDAL' plugin is used. We used sample coordinates also called GCPs (Ground Control Points) method where world coordinates of

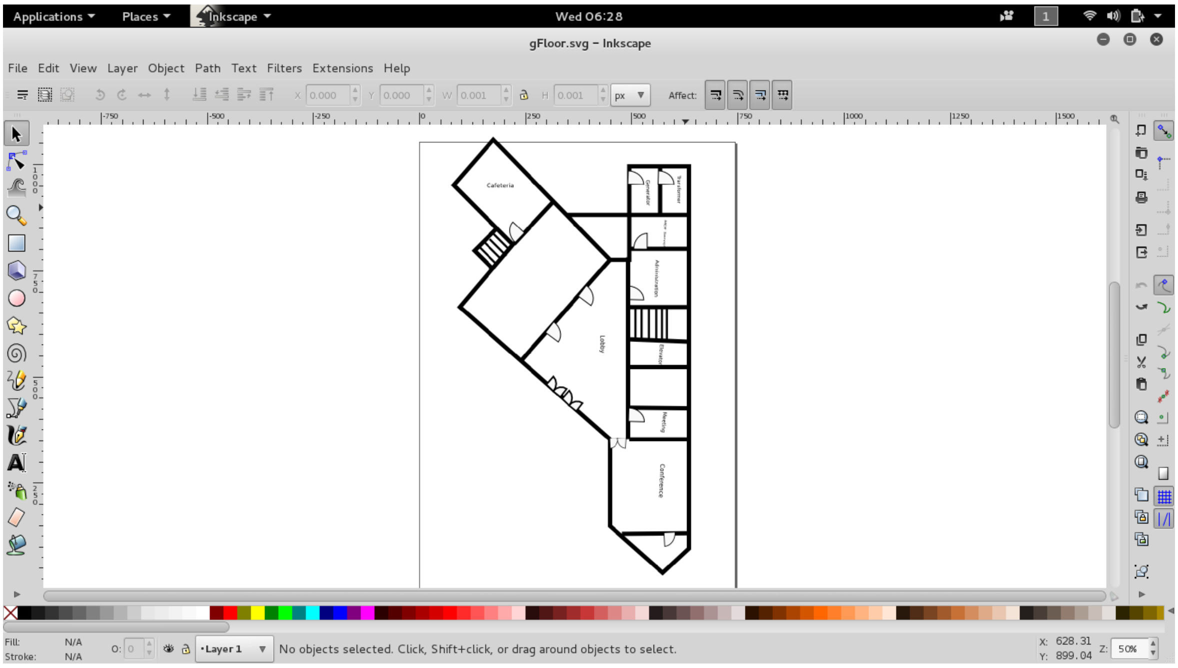



Figure 6.2: Drawing Map in InkScape

some points or pixel are determined and then image is warped and made to fit in world coordinates. Thin plate spline interpolation method was employed on our control points for georeferencing. Geo referencing generates .wld file and GeoTiff file.GeoTIFF is a metadata standard that embeds georeferencing information along with map projection, coordinate systems, ellipsoids, datums informations and so on to establish the exact spatial reference for file.Fig 6.3 shows georeferencing in QGIS

### 6.2.3. Overlaying the Geo-referenced image in world map

Finally, we overlaid the image into the map using Mapbox Studio. Mapbox is a mapping platform for developers for making custom maps. The GeoTIFF file works as a tile layer which is to be overlaid into a basic world map. Mapbox Studio offers various styling tools for customizing maps. Our project doesn't demand a lot of work in outdoor maps. So a basic map with very less styling has been used.Fig 6.4 shows overlaying and styling in Mapbox.

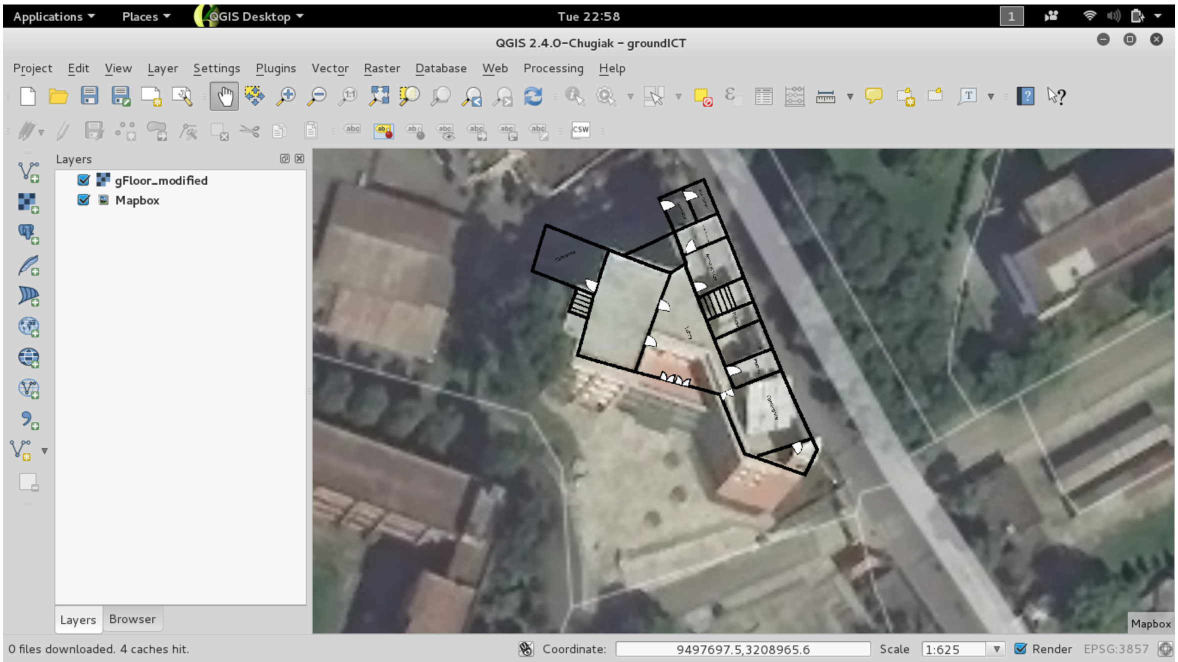



Figure 6.3: Georeferencing in QGIS

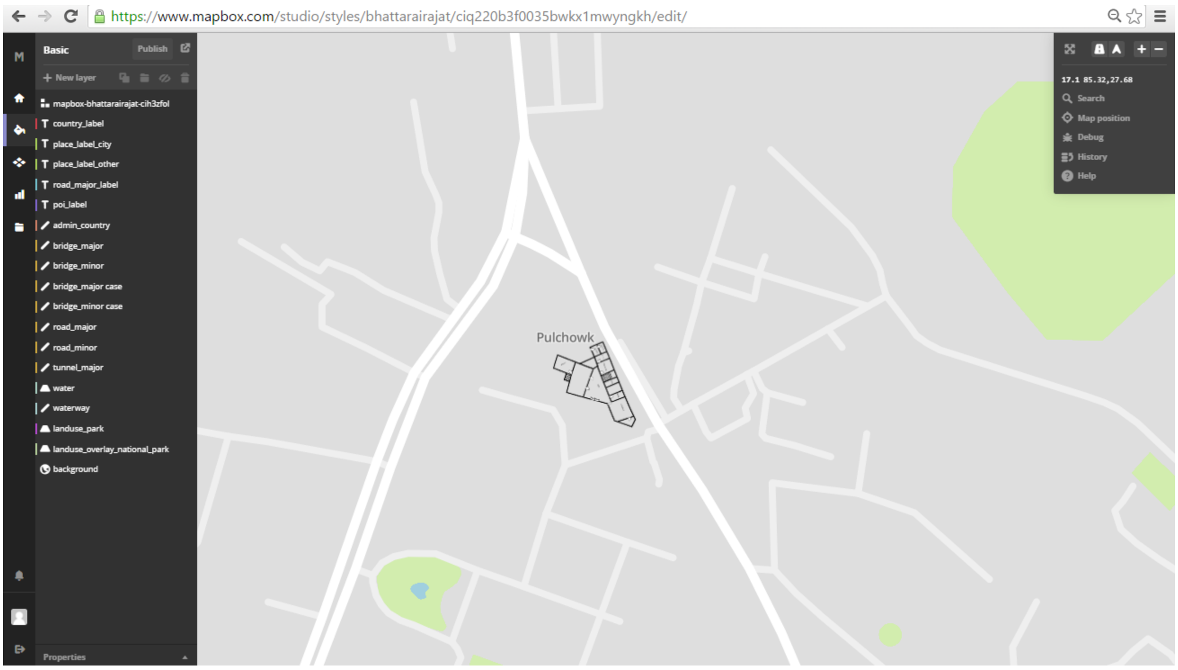



Figure 6.4: Styling and overlaying in Mapbox

### 6.2.4. Rendering the Map

The rendering of map was done in javascript as a web application. Mapbox GL JS was used for this purpose. Mapbox GL JS is a JavaScript library that uses WebGL to render interactive maps from vector tiles and mapbox styles. A marker was added to denote the location of user. The json-server provides the location data as json object. The data is retrieved from the server every two seconds and the marker points accordingly. And hence the localization occurs in the map.

The coordinates in the rendered map are identified by their latitude and longitude. However, since we have used our own local cartesian coordinate system in rssi database , a suitable formula is needed to convert them into latitude and longitude. So, we identified some key points in the map and using appropriate matrix transformation formula, we arrived at a formula that would give us the latitude and longitude value for given value of x and y.

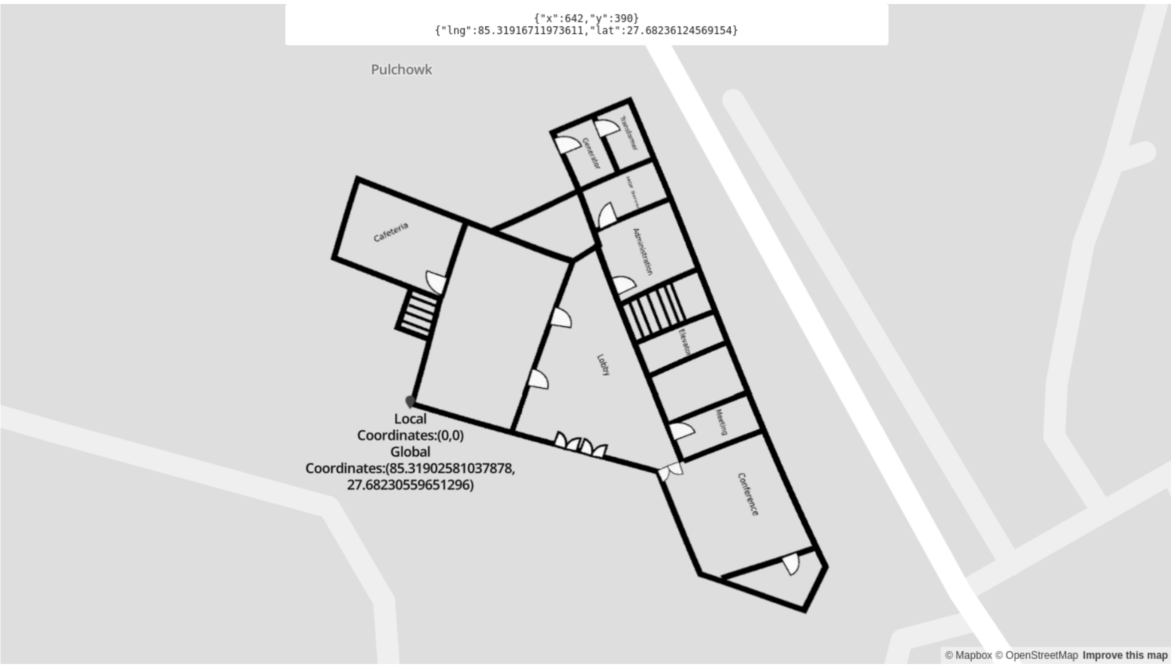



Figure 6.5: Local and World Coordinate of a reference point 1

At the end we arrived at the following formulas which would transform local coordinates in x and y to global coordinates in longitude and latitude.

$$longitude = 85.31902581037878 + 0.00010058283 * x/9 + 0.00005498528 * y/18 \quad (6.1)$$

$$latitude = 27.68230559651296 - 0.00002375195 * x/9 + 0.00016091938 * y/18 \quad (6.2)$$

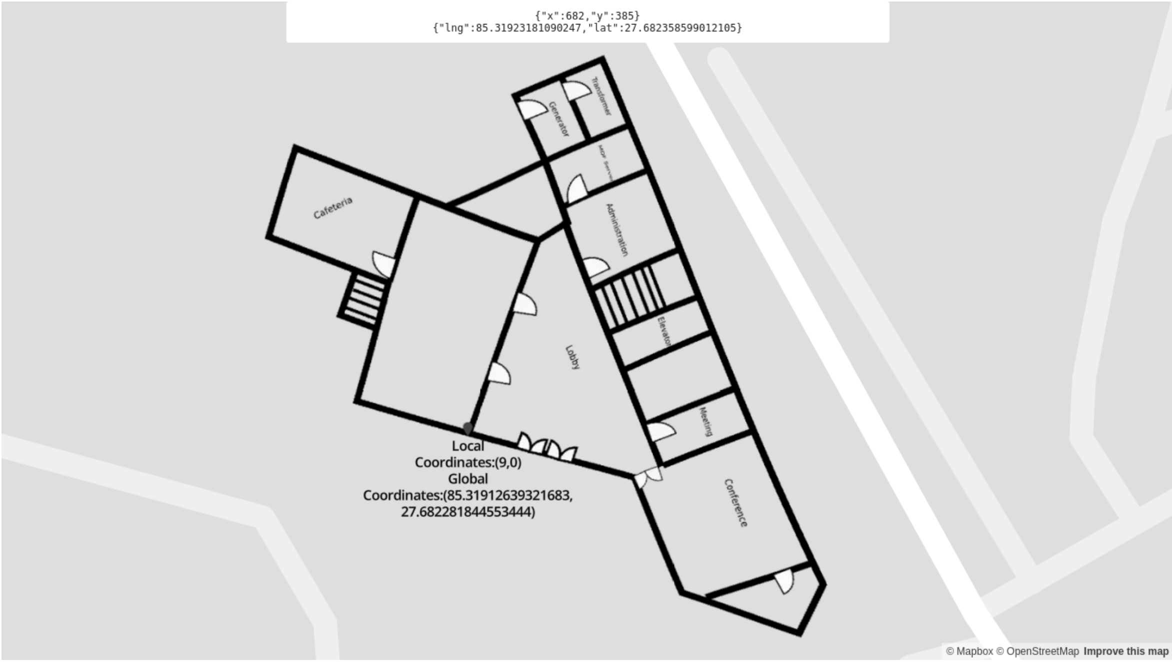



Figure 6.6: Local and World Coordinate of a reference point 2

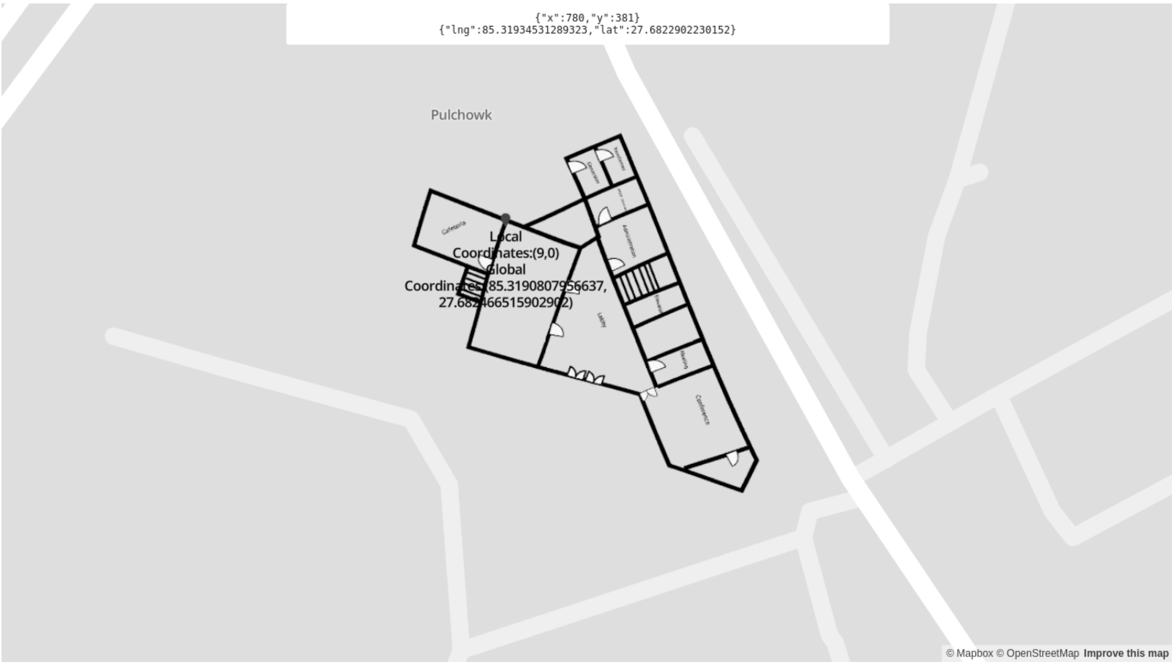



Figure 6.7: Local and World Coordinate of a reference point 3

### 6.3. Client Server Communication

Relaying data between client and server in real time is the vital portion of the project. The main purpose of client server communication in our project is to relay the real time RSSI data obtained by the android device to the server where the server can process the RSSI data to produce the location of the android device.

#### 6.3.1. Client Side

On the client side, we have the android application. The android application has a java socket client which can send data to the python socket server. The java socket client runs on the separate thread than the main thread so that it can communicate data asynchronously to the server. In every two seconds, the java socket client sends data which contains information about SSID, MAC address and RSSI of all the access points available at that point.

#### 6.3.2. Server Side

On the server side, we have two servers running. One of them is python socket server and the other is json-server.

##### 6.3.2.1. Python Socket Server

Python Socket Server receives data from the java socket client. We have set up the python socket server so that is always up and can continuously receive data from the java client. However, since java socket client transmits the data in a single string, proper parsing mechanism is required to extract the necessary information. So, we have devised a mechanism to parse the string received from the java client. At the end of parsing, we get a nice python list of RSSI values and MAC addresses. Thus obtained RSSI values and MAC addresses are used by other modules for further processing.

##### 6.3.2.2. Json-Server

The Json-Server is used to server the JSON data to our HTML file. Every two seconds the HTML file sends a GET request to the Json-Server and the Json-Server replies with the JSON data from db.json file which contains information about current latitude and longitude position of the android device. The db.json file is modified by the python program every few seconds so that the db.json contains the current latitude and longitude position of the android device.

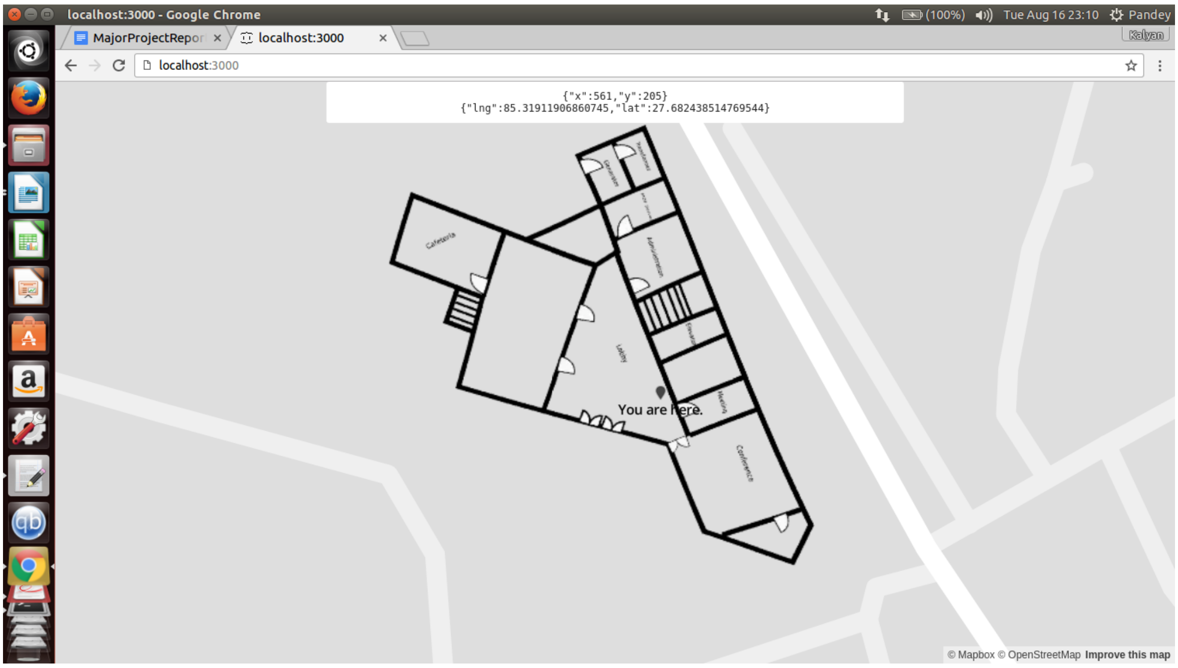



Figure 6.8: Json-Server serving data at localhost:3000

### 6.4. RSSI Fingerprinting

The next step is to create the RSSI Fingerprint database. For the purpose of recording the RSSI fingerprints, we divided our test site i.e. the ground floor of KOICA ICTC building into the grid of 3m X 3m. Now, at each grid point location we noted the signal strength of 7 access points. The distance between two points is chosen in such a way that there is diversity in measured signal strength. We collected the signal strengths of about 50 different points.The measured signal strengths of all the 7 access points are stored in a CSV file format. The CSV file will later be read by a python program for comparing present RSSI values to the values present in the CSV file.

Table 6.1: Access Points and Their Mac Address

| S.N | Access Point SSID | MAC Address |
|---|---|---|
| 1 | wifi-hotspot | ac:72:89:83:48:7a |
| 2 | CIT1 | d4:ca:6d:8c:4b:97 |
| 3 | UBT_ICTC | 00:26:66:86:27:d0 |
| 4 | CIT2 | d4:ca:6d:8d:4b:bb |
| 5 | CIT4 | 00:0c:42:c8:5f:22 |
| 6 | tu_ICT_Center | 34:a8:4e:4e:a6:10 |
| 7 | CIT Connect-2 | 00:14:d1:23:da:b3 |

### 6.5. Android Application Development

Next step involves developing an android application. Our android application provides an easy to use interface for the user to navigate in an indoor environment. The android application has been developed using Android Studio. The main steps involved in android application development were main layout design, sensor data acquisition, wifi data acquisition.

#### 6.5.1. Main layout Design

Our app consists of a main activity with an edit text field and a button. In the text field, user can enter the address of the json-server. On pressing the button, java socket client will be started in a separate thread which will start streaming RSSI data immediately to the server. On pressing the button, users will also be redirected to the web browser where they can view their current position in the map.

#### 6.5.2. Sensor Data Acquisition

Android provides sensor framework which lets you access many different kind of sensors. Not all type of sensors are available in every android device. However, accelerometer and magnetometer are generally present in many android devices. So using the sensor

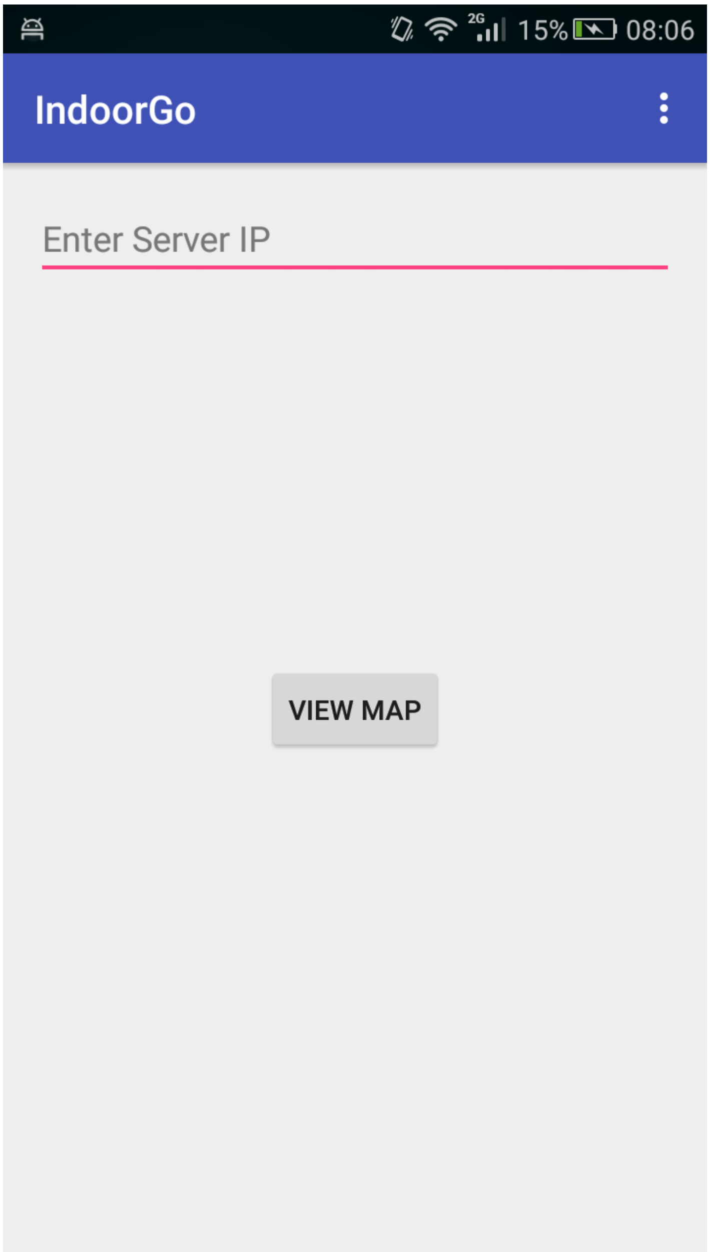



Figure 6.9: Main Layout Design

framework, we acquired the data of accelerometer and magnetometer. We also applied low pass filter to the accelerometer data as well as magnetometer data for stable data values. Now that we have accelerometer and magnetometer data, next we calculated orientation of the device using those data. The accelerometer data was also used to calculate total steps taken by the user. The steps were calculated by following the procedures mentioned in the concerned theory.

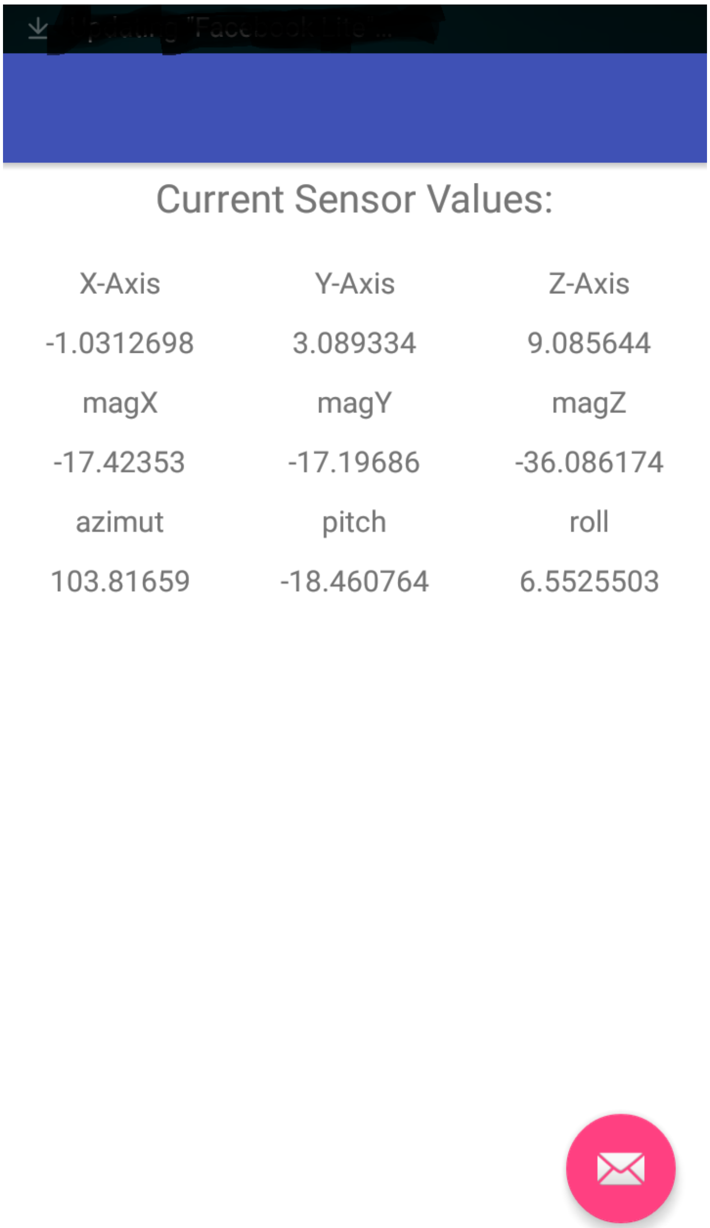



Figure 6.10: Android app showing sensor values

### 6.5.3. WiFi RSSI Data Acquisition

Android provides WiFiManager API to manage all aspects of WiFi connectivity. Using the WifiManager API. we can scan all the available wireless networks and gather information about them. We noted SSID, MAC address and RSSI value of all the available access points using WiFi manager and stored them in a java array which later is converted to strong and sent to python socket server.

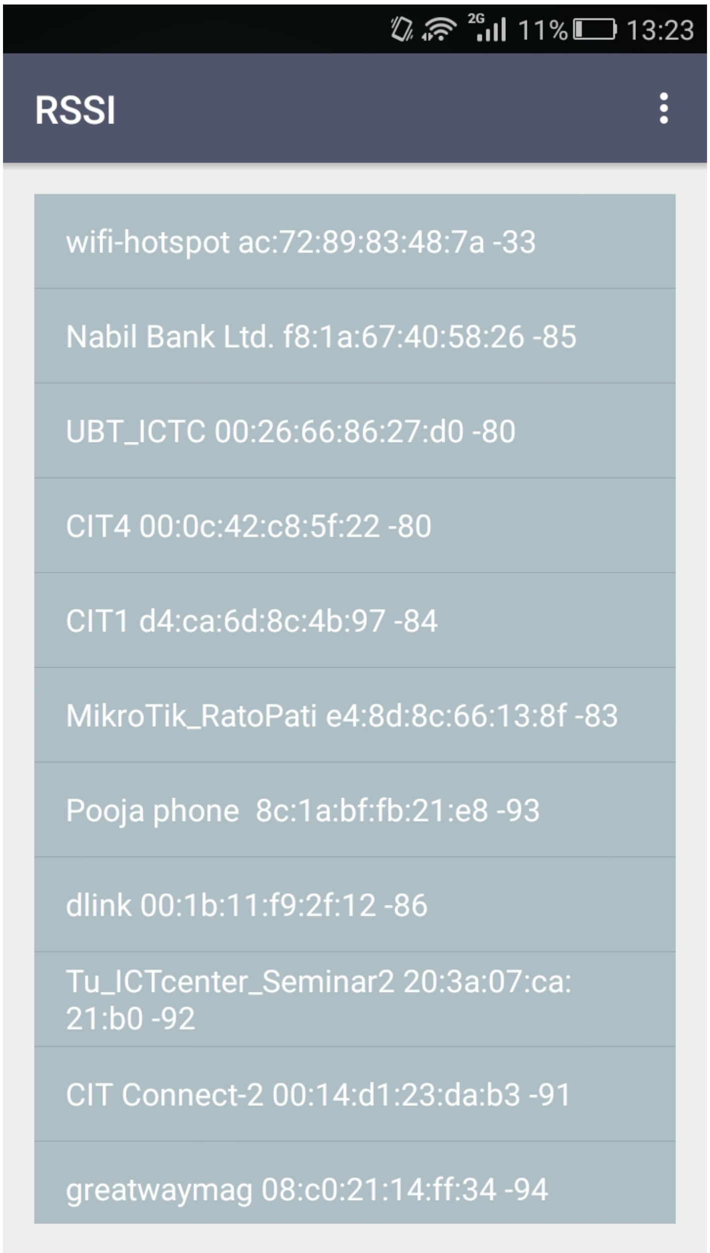



Figure 6.11: Android app showing SSID, Mac Address and RSSI of available access points

## 6.6. Position Estimation(Online Phase)

### 6.6.1. Online Tracking with Particle Filter

The algorithm for particle filter is as given below. A first a set of particles are created, the particles are assigned weight based upon sensor measurement.Low variance resampling selects the particle in state space with higher probability density. The HMM based motion model propagates particle form one state to another and HMM based observation model fusses sensors for weight update[3]&[11].

**Algorithm Particle_filter(**$\mathcal{X}_{t-1}, u_t, z_t$**):**
$\bar{\mathcal{X}}_t = \mathcal{X}_t = \emptyset$
for $m = 1$ to $M$ do
sample $x_t^{[m]} \sim p(x_t \mid u_t, x_{t-1}^{[m]})$
$w_t^{[m]} = p(z_t \mid x_t^{[m]})$
$\bar{\mathcal{X}}_t = \bar{\mathcal{X}}_t + \langle x_t^{[m]}, w_t^{[m]} \rangle$
endfor
for $m = 1$ to $M$ do
draw $i$ with probability $\propto w_t^{[i]}$
add $x_t^{[i]}$ to $\mathcal{X}_t$
endfor
return $\mathcal{X}_t$

An online tracking algorithm using a particle filter is performed according to the following steps:

#### 6.6.1.1. Particle Initialization

The initial position is calculated through the weighted K nearest neighbor (WKNN) method. It searches for K closest matches of known positions in the WiFi received signal space from the offline-built data set. By averaging these K position candidates with adopting the distances in signal space as weights, the initial estimated position is acquired. This initial position estimation is used as the starting point for particles.

#### 6.6.1.2. Particle propagation

Next, we apply the state transition model to guide particle propagation. During the state transition process, we observe that the particle propagation plays an important role, which represents the state transition probability. The more accurate the particles propagate towards the right position, the better the positioning performance will be. We generate new particles by sampling $X_t^{[m]}$ from the distribution $p(X_t|X_{t-1})$.

#### 6.6.1.3. Particle weight update

After particle propagation in each epoch, we weight the sample$X_t^{[m]}$ by the probability $p(Z_t|X_t^{[m]})$ . The weight is calculated as the product of different sensor measurement likelihood function.

#### 6.6.1.4. Particle resampling

Once the particle weights are updated, we perform importance resampling to update the particles' state by picking a random sample $X_{t+1}^{[m]}$from the sample set $S_t$ according to the importance weight $w_t^{[m]}$. In resampling, the weight of each particle is treated as a probability, where this particular particle is chosen to be at the estimated position. Particles with higher weights will be picked more frequently than others. This is how the resampling is able to eliminate wrongly moved particles and correctly track the user's position.

The resampling step is a probabilistic implementation of the Darwinian idea of survival of the fittest: It refocuses the particle set to regions in state space with high posterior probability. By doing so, it focuses the computational resources of the filter algorithm to regions in the state space where they matter the most[11].

**Algorithm Low_variance_sampler($\mathcal{X}_t, \mathcal{W}_t$):**

$\bar{\mathcal{X}}_t = \emptyset$
$r = \text{rand}(0; M^{-1})$
$c = w_t^{[1]}$
$i = 1$
for $m = 1$ to $M$ do
    $u = r + (m-1) \cdot M^{-1}$
    while $u > c$
        $i = i + 1$
        $c = c + w_t^{[i]}$
    endwhile
    add $x_t^{[i]}$ to $\bar{\mathcal{X}}_t$
endfor
return $\bar{\mathcal{X}}_t$

#### 6.6.1.5. Position estimation

After the resampling process, the estimated position is calculated as the mean of all the re-sampled particles' positions.

### 6.7. Threshold Selection for Step Detection

As described in theory, the peak detection algorithm needs a threshold value.To do the selection,user carrying hand held device was made to travel an area in first floor of KOICA ICTC with varying walking pattern. The selection of appropriate site for testing the algorithm is very important since the peaks values for step detection from acceleration data are highly dependent upon surface of floor.The user has to hold mobile device without jerking and changing orientation.The primary objective of user should be to get localized and hence navigate, rather messing with system to cause failure in localization.

The results obtained are discussed in result section.

Figure 6.12: Human Walk Analysis for threshold selection training phase,user walking in straight path with fixed step length,taking few number of turnings,holding device firmly.

### 6.8. Parameter Selection for Gaussian Process Regression

As described in theory,GPR has 3 most important parameter for best fitting of data. The signal variance $\sigma_f$, length scale $l$ and noise variance $\sigma_n$ determines the nature of fit.The GPR algorithm gives mean and variance at the points regression.The objective should to minimize the error in fit. Since we have very few training data we couldn't do much to minimize error.

For better fit we should maximize the log likelihood these parameters. These parameters are also called hyperparameters. Let $= \langle \sigma_n^2, l, \sigma_f^2 \rangle$ denote the hyperparameters we wish to estimate. The log likelihood of the observations is given by:

$$logp(y|X,\theta) = -\frac{1}{2}log(K+\sigma_n^2 I)\frac{1}{2}|log(K+\sigma_n^2 I)| - \frac{n}{2}log2\pi \tag{6.3}$$

Using default optimizer of pyGPs library in python for Gaussian Process Regression, these values were optimized.

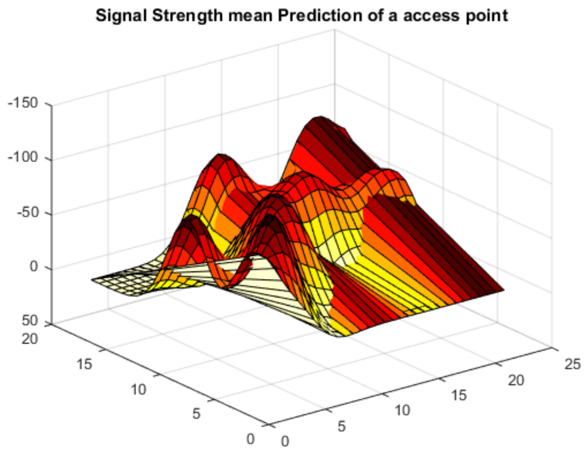



Figure 6.13: Signal Strength Mean prediction of AP1.

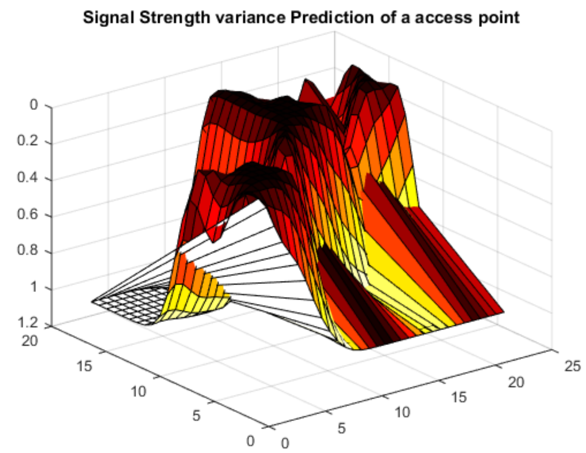



Figure 6.14: Signal Strength Variance Prediction.

# 7. RESULTS AND DISCUSSION

## 7.1. Step Detection

The low cost MEMS sensor data obtained from Android phone is very unreliable.The sensor data is sampled at 5Hz with Matlab Mobile and analysed in Matlab.

The peak detection algorithm is implemented as per described in theory. The threshold value of should be set such that there in no false peak detection. The threshold value is selected one standard deviation away from mean. The threshold should be tuned experimentally to match a person's level of movement while walking, hardness of floor surfaces, etc.

We tested our peak detection algorithm under various circumstances.The results obtained results showed excellent performance.

Table 7.1: Human Walk Analysis

| Walking Pattern | Detected Steps | Actual Steps |
|---|---|---|
| Straight Walk(in IOS) | 11 | 11 |
| Straight Walk(in Android Phone) | 19 | 20 |
| Walk with few turnings | 40 | 42 |

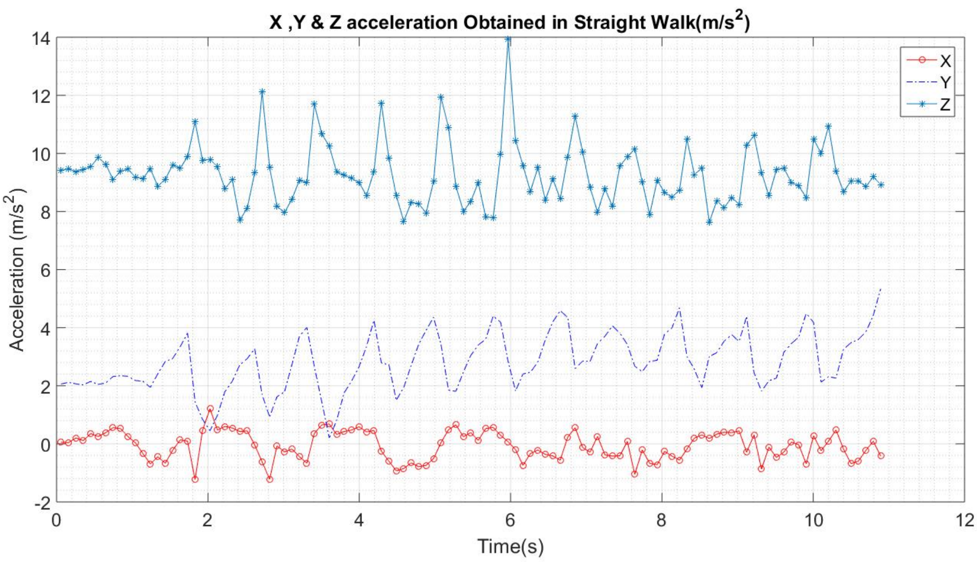



Figure 7.1: x,y and z acceleration obtained from mobile sensor sampled at 5Hz.

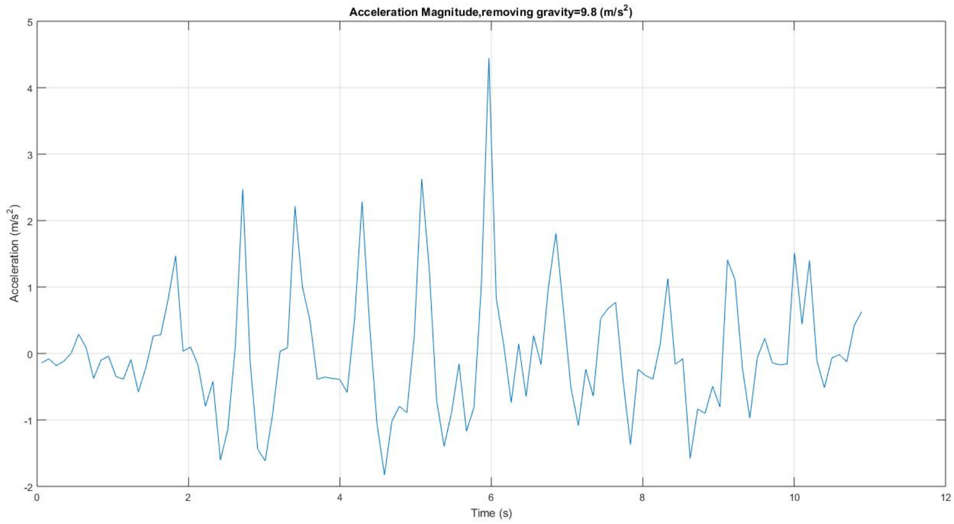



Figure 7.2: sum of root mean square of accelerations removing gravity offset('g').

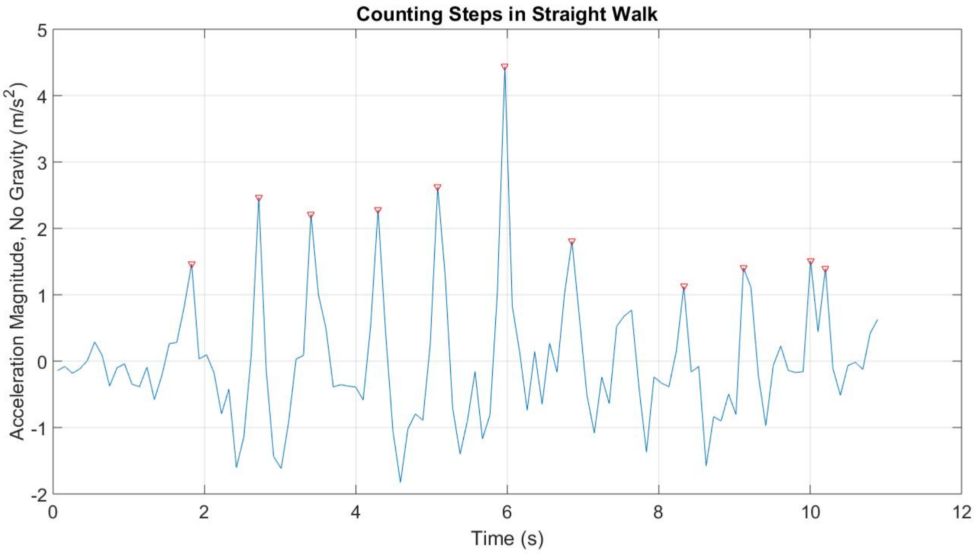



Figure 7.3: Peak detected with threshold above one standard deviation of signal(peaks = 11).

Since , we have taken root mean square of acceleration,the values of acceleration can be peaked if user moves the phone randomly or changes orientation of phone. In such case, false peak detection may occur.False peak detection can be avoided by increasing the threshold value but this is the promising solution. It is advised to the users to hold mobile firmly while navigating.

### 7.2. K Parameter Selection for W-KNN

-The data obtained from Wi-Fi Fingerprint is divided in training and test set.The data is split randomly shows that there is now bias in test and training set.25% of data is taken as test case and 75% as training set.The average error obtained for $k = \{1, 2, ....., 20\}$ in x and y measurement are observed for 1000 times. The curves below shows k parameter vs error means.

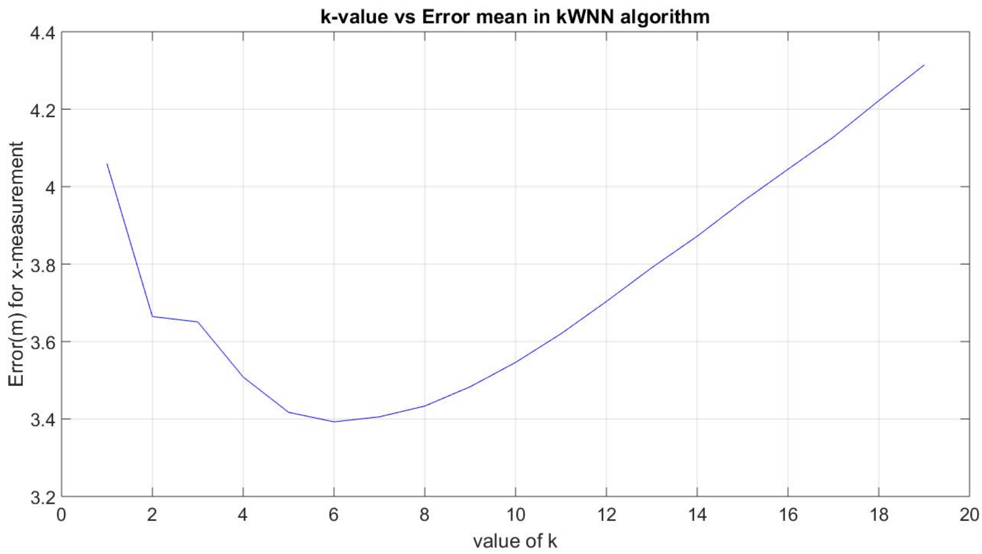



Figure 7.4: x-error vs k-value showing minimum error for k = 6.

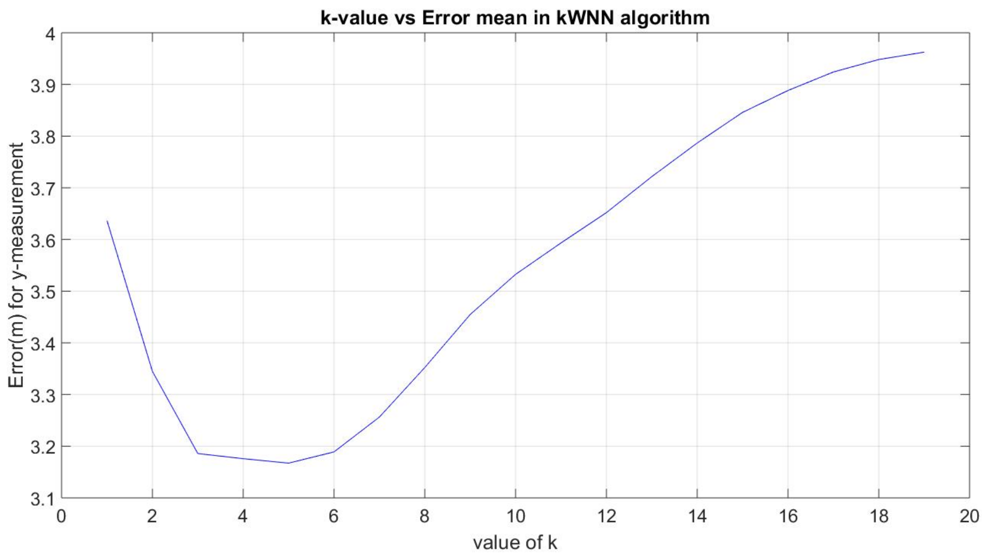



Figure 7.5: y-error vs k-value showing minimum error for k = 5.

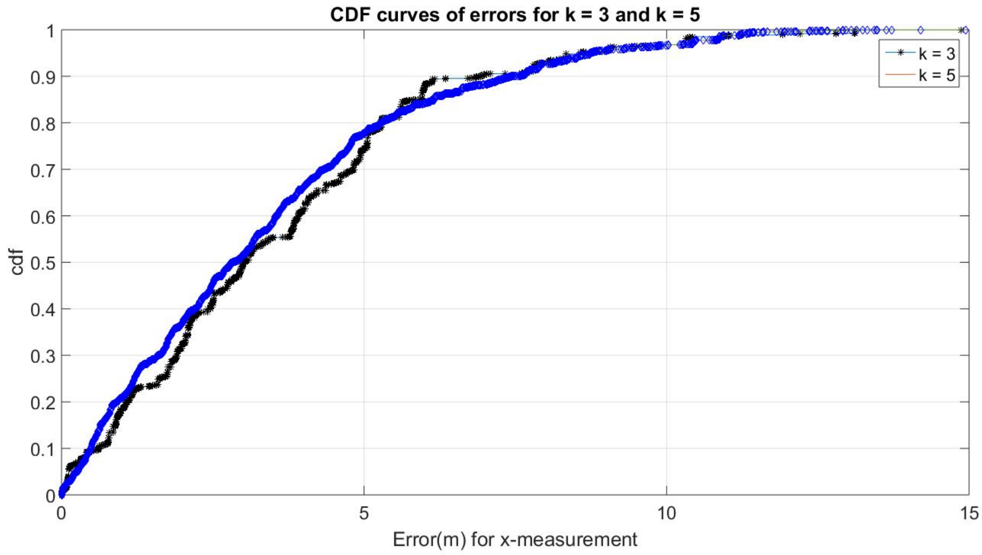



Figure 7.6: Cdf comparison showing probability of error in x for k = 3 and 5.

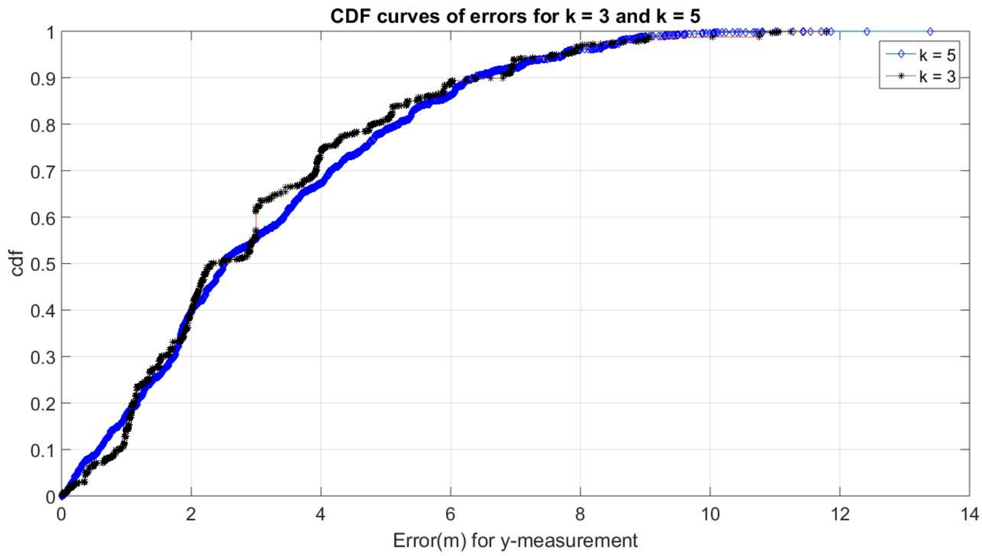



Figure 7.7: Cdf comparison showing probability of error in y for k = 3 and 5.

Fig 7.1 shows x error 3.39m at k = 6 and Fig 7.2 show y error 3.16m. These are minimum errors obtained. So it obvious to choose k = 5. But Fig 7.3 and Fig 7.4 comparison of Cdf curves for k = 3 and k =5 shows almost similar performance is error for more than 60% of time. It doesn't create much difference in error whether we choose k = 3 or 5. For our convenience we have chosen k = 3.

### 7.3. Signal Strength Variation Visualization

In Matlab,we generated Signal Strength variation plot of all sampled access points over the area. The surfaces were smoothed by Biharmonic Spline Interpolation.

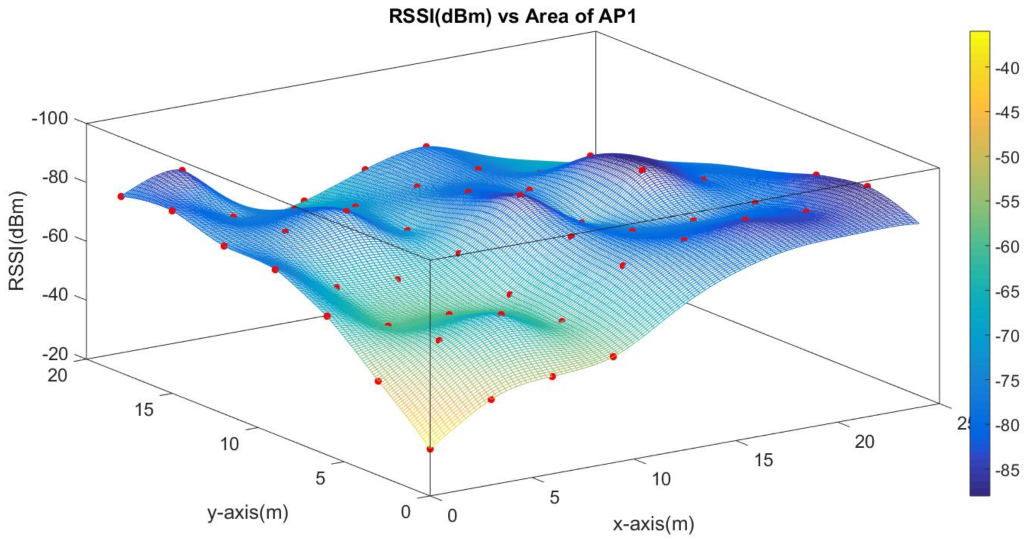



Figure 7.8: Signal Strength(dBm) of AP1 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

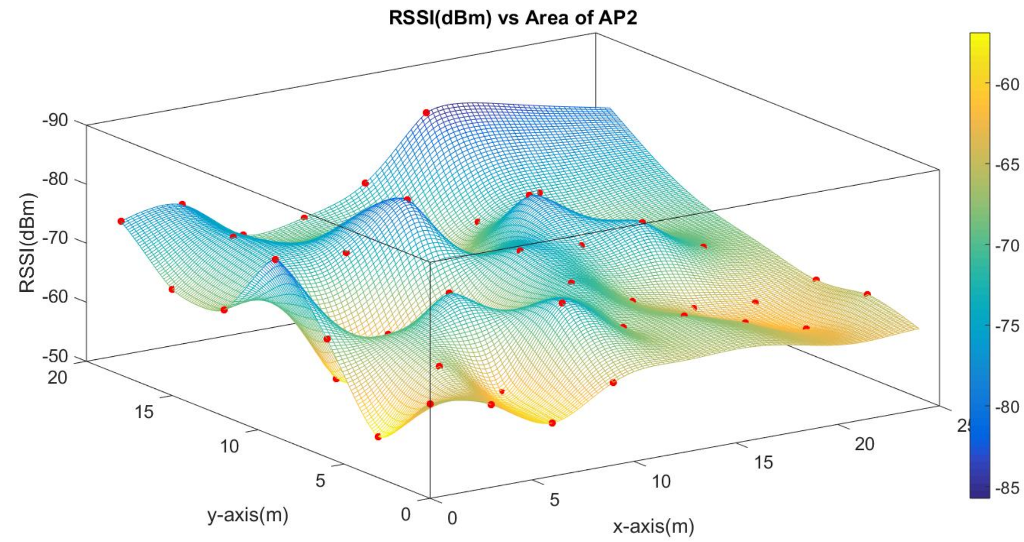



Figure 7.9: Signal Strength(dBm) of AP2 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

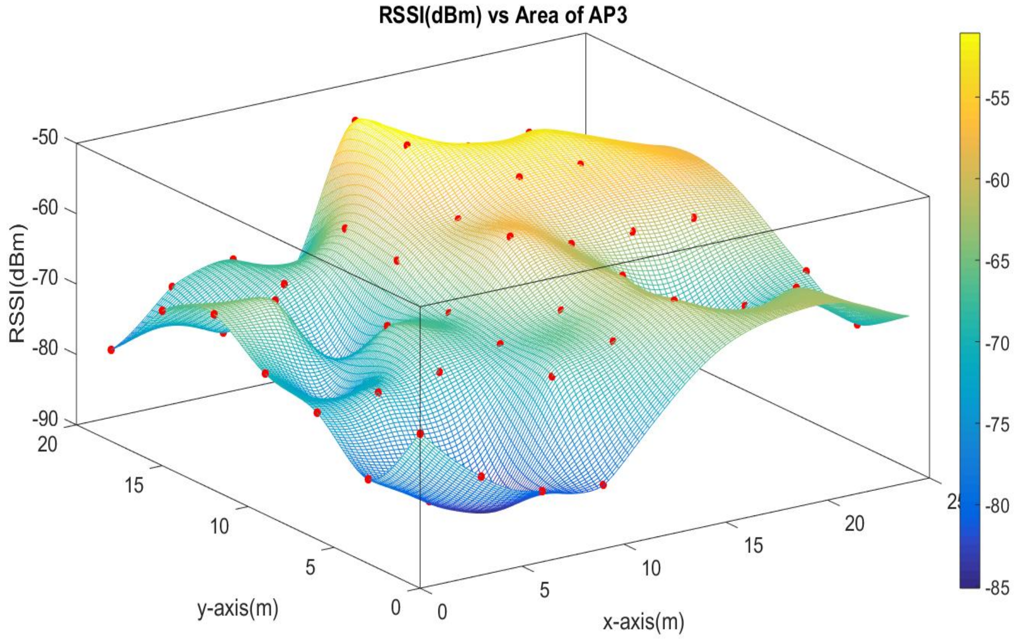



Figure 7.10: Signal Strength(dBm) of AP3 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

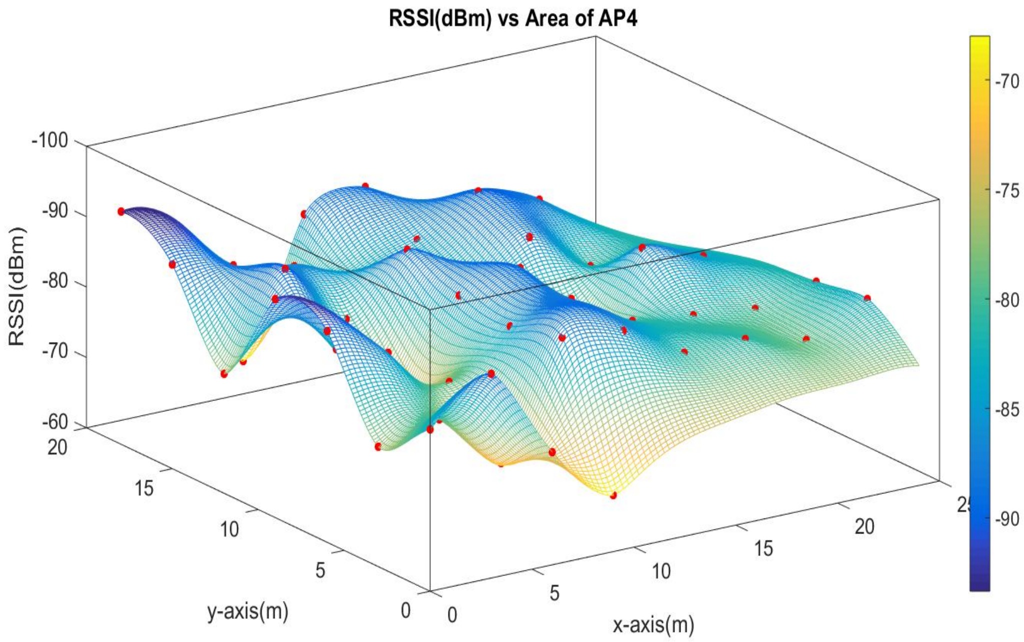



Figure 7.11: Signal Strength(dBm) of AP4 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

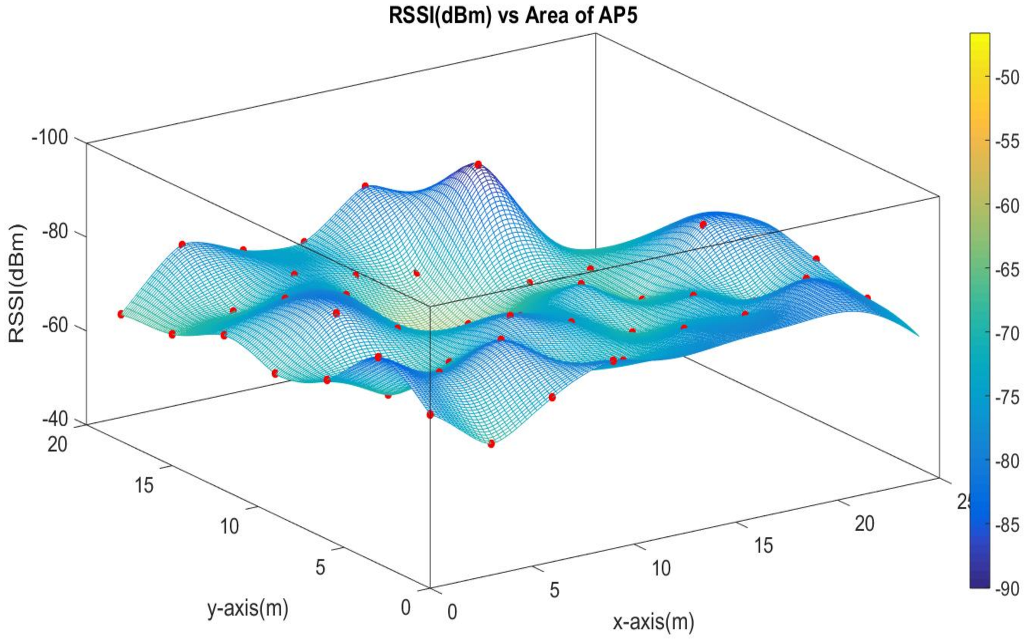



Figure 7.12: Signal Strength(dBm) of AP5 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

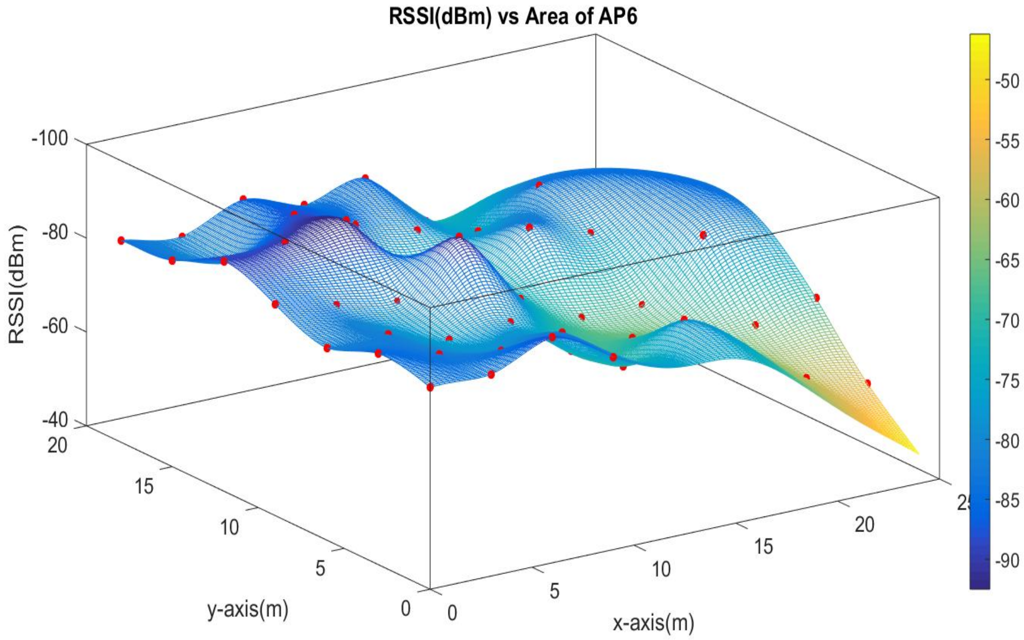



Figure 7.13: Signal Strength(dBm) of AP6 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

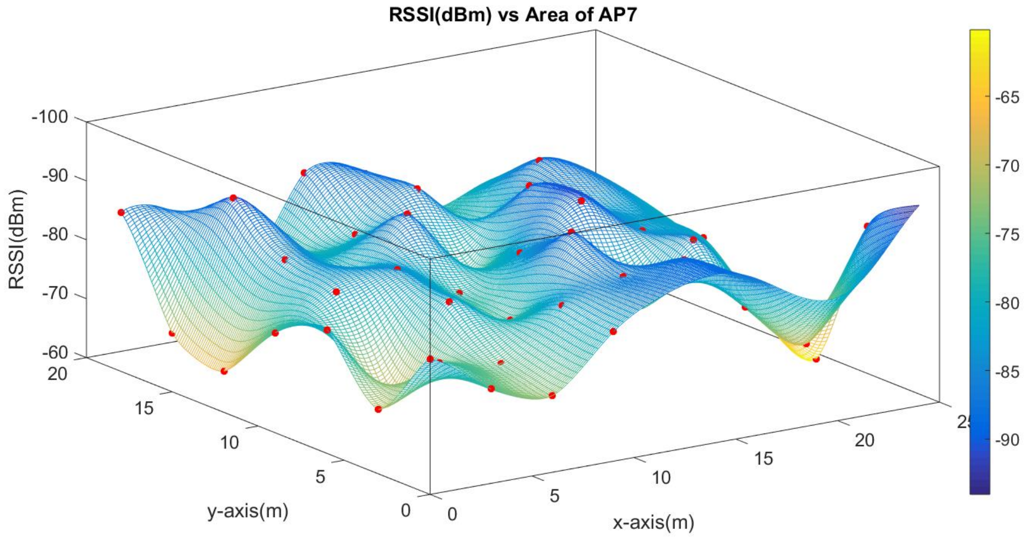



Figure 7.14: Signal Strength(dBm) of AP7 over ICTC Ground Floor 3D surface generated by Biharmonic spline interpolation

### 7.4. Trilateration

The results from trilateration were obtained as below in figures . However, the results were very inaccurate. This is because the RSSI measurements tend to fluctuate according to changes in the environment or multipath fading.

Table 7.2: A data obtained from trilateration

| MAC Address | RSSI Value Obtained | Distance(m) |
|---|---|---|
| 00:26:66:86:27:D0 | -75 | 5.6234132519 |
| AC:72:89:83:48:7A | -72 | 5.01187233627 |
| 48:DB:50:6C:A6:3C | -72 | 14.1253754462 |

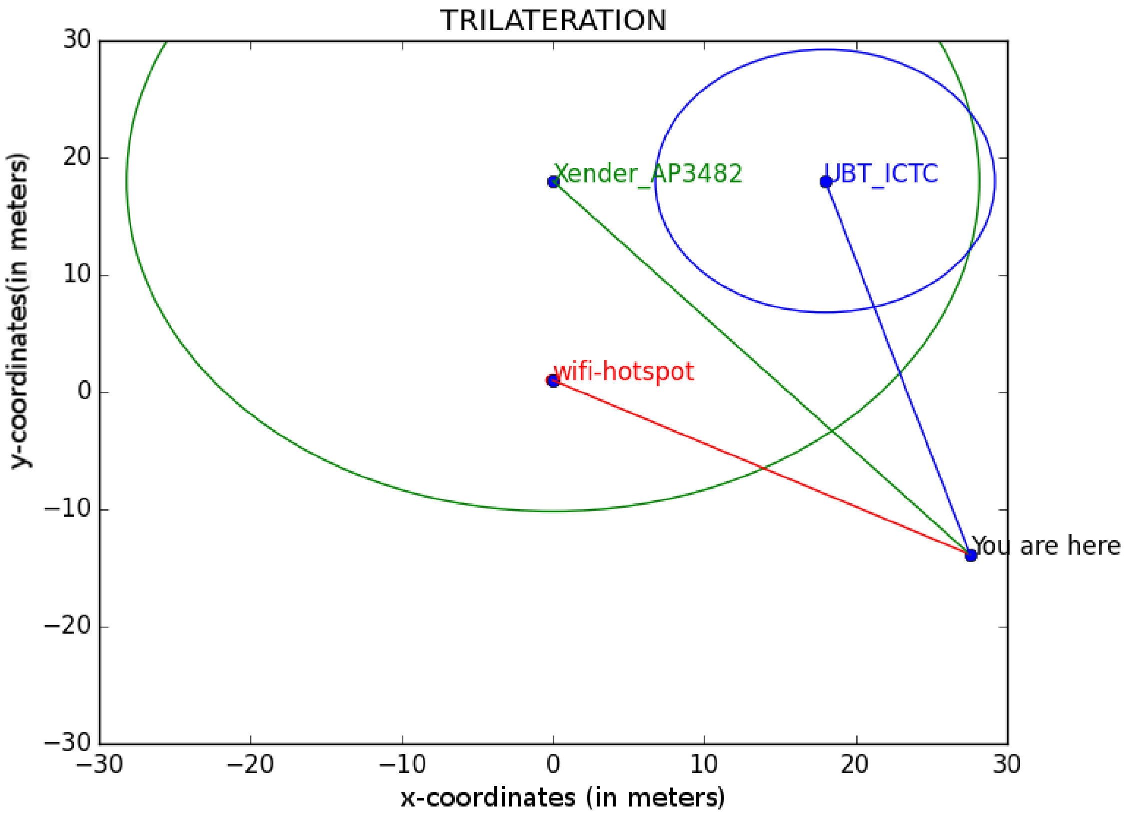



Figure 7.15: Computing 2D location using trilateration.


Figure 7.16: Trilateration result in browser .

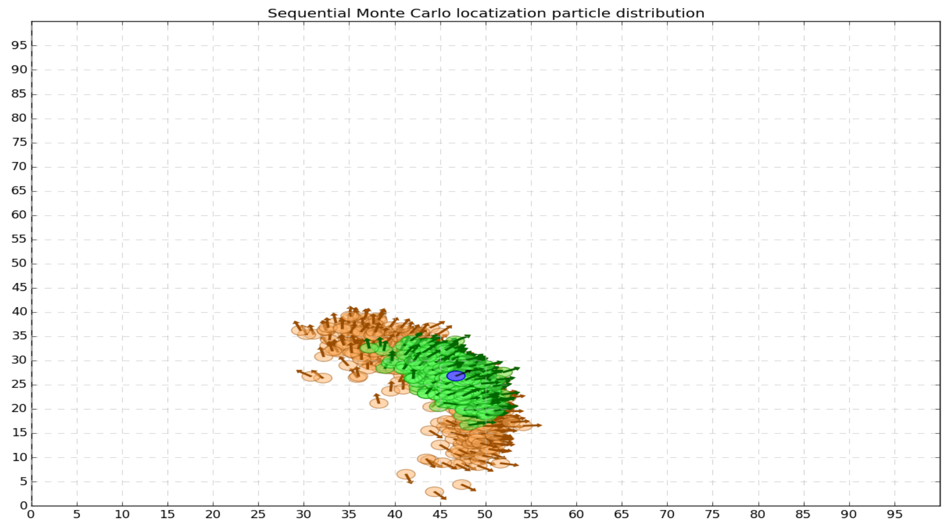
### 7.5. WKNN Based Position Estimation on RSSI Fingerprint Data

A more accurate result was obtained from this method with an accuracy level of about 3.5 meters.


Figure 7.17: Output from WKNN based on RSSI Fingerprinting data.

## 7.6. Sequential Monte Carlo based localization output

Using probabilistic localization algorithm increased localization accuracy to 1-2 m.



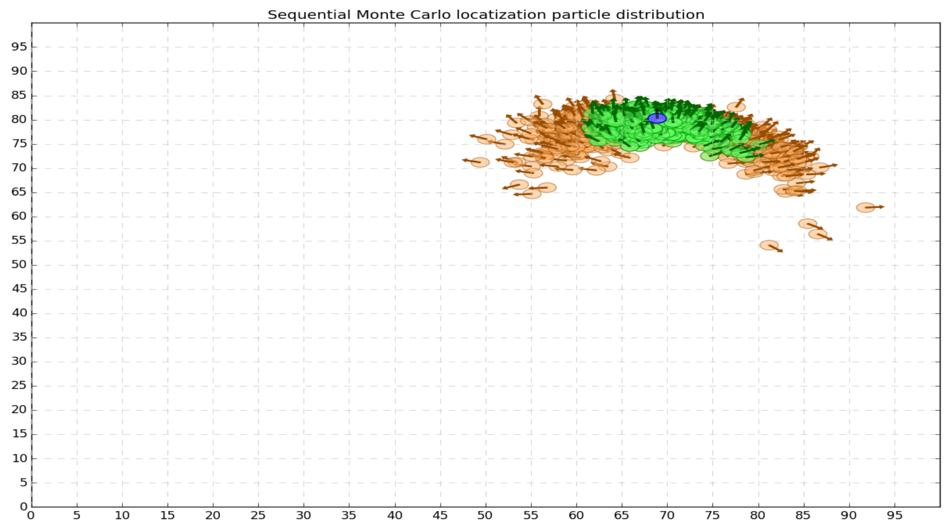




Figure 7.18: During localization phase, particles distribution with brown color particles before re-sampling, green color particles after re-sampling in a world of 100 X 100(simulation) and blue color actual location .

#### 7.6.0.6. Visualization in Real World

Motion of user was tracked using MCL in our site. The results obtained are shown in figure in figure below.

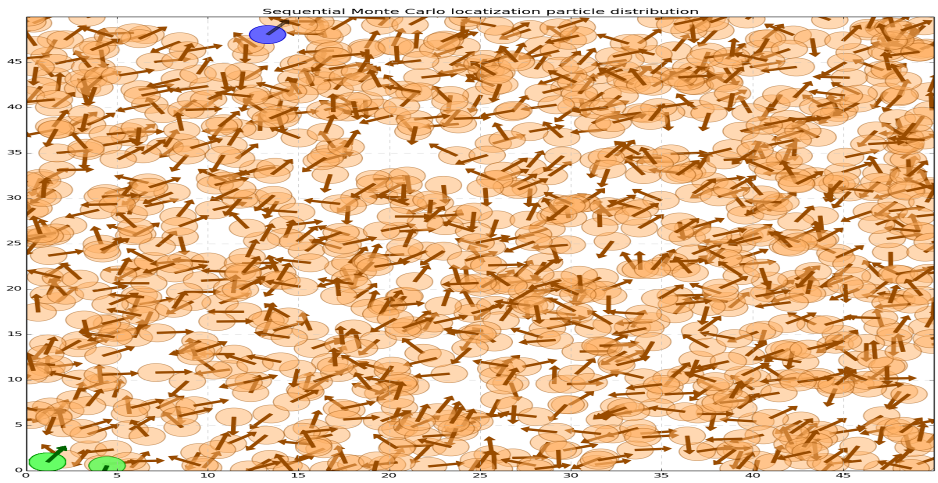



Figure 7.19: Initial Distribution of Particle in real world(world size = 50X50 m).Brown color particle showing initial distribution and green color particle showing actual location

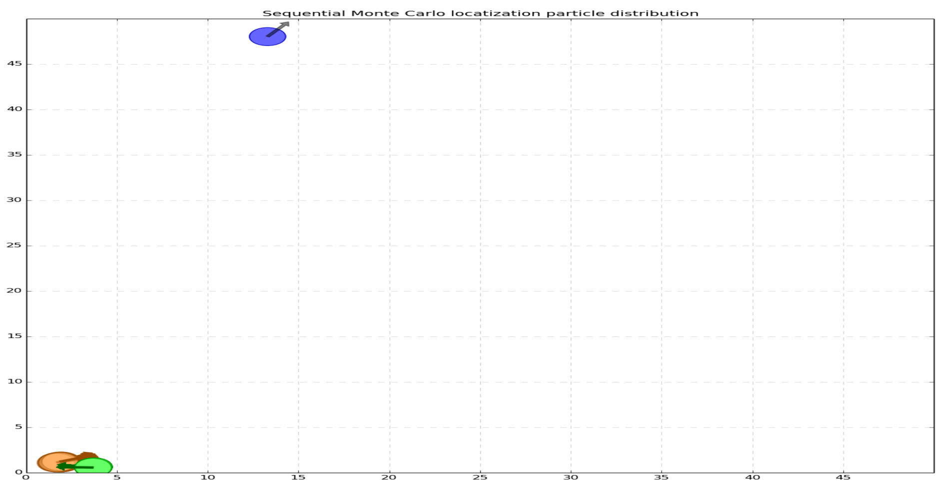



Figure 7.20: Green Color Particle showing actual location after first interation,brown color particle showing states before resampling

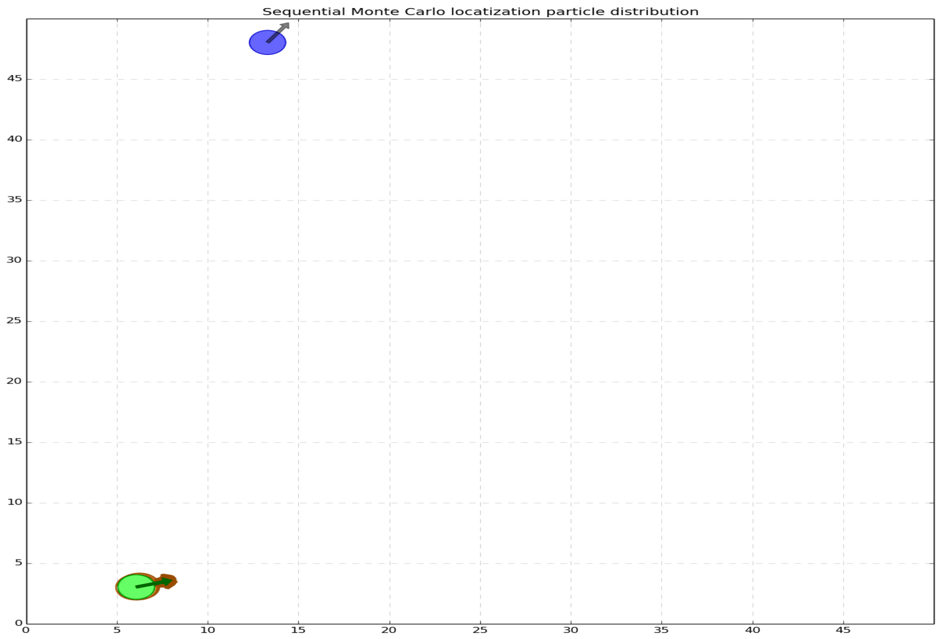



Figure 7.21: Green Color particle showing actual location in 8th iteration

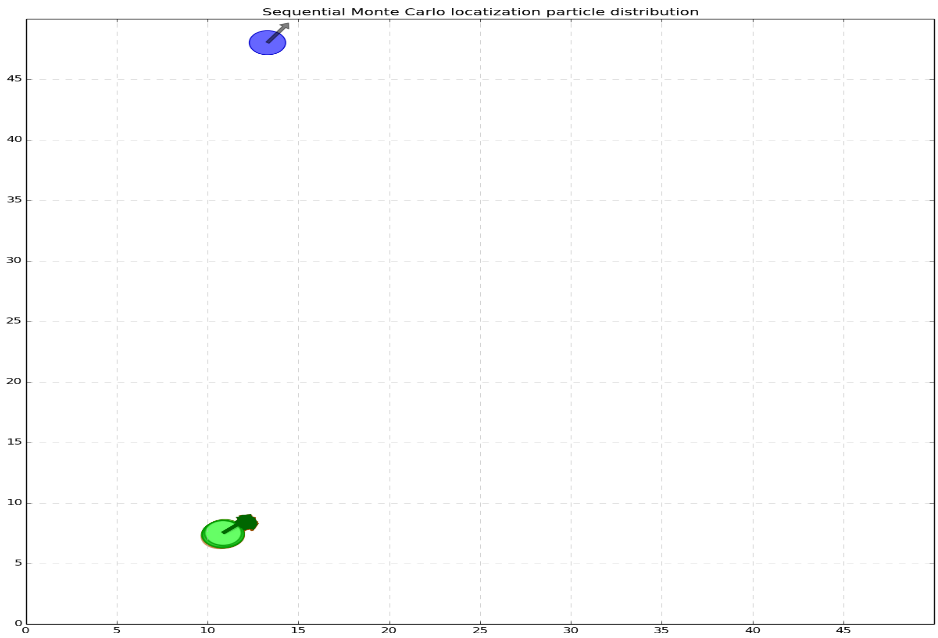



Figure 7.22: Location obtained in 18th iteration when user has taken 18 steps

## 7.7. Comparison

Every system implemented have some constrains.To select one of the localization system one can consider accuracy, reliability, fault-tolerance,cost effectiveness,etc.

Table 7.3: Comparison of Trilateration ,WKNN and MCL

| | Trilateration | WKNN on WiFi Fingerprint Data | MCL |
|---|---|---|---|
| Accuracy | room level accuacy | 3-5 m | 1-2m |
| System Complexity | Low | Medium | High |
| Cost | Low | Moderate | High |
| Reliabllity | Unreliable | Least | High |
| Responsiveness | Fastest | Fast | Slow |

# 8. COST ESTIMATION

The cost given below are excluding non-recurring engineering cost(NRE).

Table 8.1: Cost Estimation

| S.N | Items | Rate(NRs) | Quantity | Cost(NRS) |
|---|---|---|---|---|
| 1. | RF Transceiver or Routers | 3,000 | 3 | 9,000 |
| 2. | Smart Phone | 15,000 | 1 | 15,000 |
| 3. | Laptop | 70,000 | 1 | 70,000 |
| 4. | Stationary | - | - | 2,000 |
| 5. | Miscellaneous | - | - | 3,000 |
| | Total | | | 99,000 |

# 9. GANTT CHART

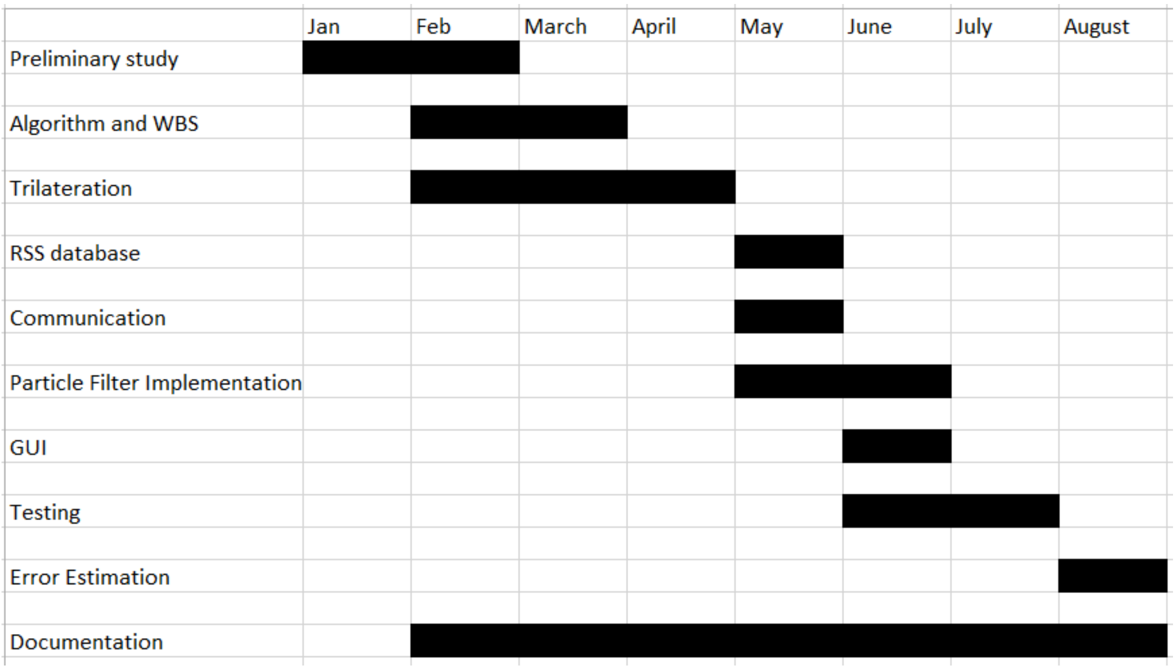



Figure 9.1: Gantt Chart

# 10. LIMITATIONS AND CHALLENGES

## 10.1. Unreliable Sensor Data

The sensors which we are using aren't primarily constructed for using in navigation system .The low cost sensor gives noisy data which in many cases doesn't make any sense. The objective of using low cost sensor is to utilize the existing infrastructure so it would be free of cost or at minimum cost anyone using it.

## 10.2. RSSI Fingerprint Database

Creation of fingerprint database is the most time consuming task in development of system. Also the rssi value obtained depends upon sensing device i.e. different user carrying different mobile device at same location can show different rssi value.In such cases we have to create rssi fingerprint for different receiver. The to reduce time requirement and remove receiver dependence problem we propose the following solution.

### 10.2.1. Crowd Sourcing

Collecting data from large number user with different hand held mobile device over a period of time will help in creating large data set and hence error over time can be minimized. In crowd sourcing, user is given a map of buildings. The user tap in the map where they are over the period of time during navigation inside building. In this period rssi value at different location is logged in database. This is cost effective solution which requires volunteers.

### 10.2.2. Simulation

There are may RF propagation planing tools currently available in market. We can give 3D models of buildings with locations of AP and attenuation factors for various component of building. Using ray tracing we obtain the RSSI value at desired location and store in database. This method is little bit inaccurate but more feasible than crowd sourcing.

### 10.3. Kidnapped Robot Problem

In particular, probabilistic approaches are typically more robust in the face of sensor limitations, sensor noise, environment dynamics, and so on. They often scale much better to complex and unstructured environments, where the ability to handle uncertainty is of even greater importance.

There are several challenges in probabilitic approach of localisation.Consider the following cases:-

1. When object to be localized changes position rapidly and show random behavior.
2. Sensor data shows very unnatural change with large unexpected noise in measurement.
3. All of Sudden no sensor data becomes unavailable.

These problems may or may not arise. In such cases the probabilistic localization algorithm fails. Such problem in probabilistic robotics is called kidnapped robot problem. This is very challenging because the robot believe it knows where it is while it doesn't.

Similar case may arise in tracking of people using probabilistic method.In such case we get position which isn't true but the localization algorithm cannot figure it out. A localization algorithm should be able to identify and take suitable action for recovery of location.

There are many approaches in detection and solution to the problem. Since, this is out of the scope of our project no further discussion is done.

# 11. CONCLUSION AND FUTURE ENHANCEMENT

## 11.1. Conclusion

Indoor localization using wireless technology is an exciting prospect attracting interest of researchers and scholars from around the globe.The vast applications that indoor localization can unravel is undeniable.
Trilateration and fingerprinting were the two methods used for calculating position with Wi-Fi based IPS. Both methods use estimated wireless signal strength to determine the position. However, each determines the location in different ways.

The trilateration technique uses one of the various indoor radio wave propagation model to locate a point so is more flexible as the system calculates position in real-time and is more adaptable to change in surroundings than fingerprinting. In real-world use however, the accuracy of this technique is terrible(not even room level accuracy as shown in the results section). This is because the RSSI measurements are affected in great extent by the surrounding and tend to fluctuate according to changes in the environment or multi path fading. The performance of this technique breaks down to appropriate selection of path loss exponent which is not easy in the case of indoor environment where path loss exponent is disorderly variant.

Fingerprinting technique on the other hand requires a detailed signal strength database for each reference point that can be compared with received signal strength(using WKNN method) in the field. Fingerprinting already considers attenuation in the database creation process, which leads to better accuracy in signal strength data for calculation. When room level accuracy (3-5 m)is wanted, WKNN on Wi-Fi fingerprints could be implemented in combination with multiple well-placed Wi-Fi access points and appropriate selection of value of K(X as shown in results section ). Fingerprinting method proofs to be the most robust as well. When one of the APs fails, it will preserve its accuracy results. However a big flaw of this system is that any changes of the environment,such as adding or removing furniture or buildings, change in weather(rain),etc. may change the Wi-Fi fingerprint that corresponds to each location.

The accuracy of Fingerprinting based systems can be further increased with probabilistic approach. Monte Carlo Localization with multi-sensor fusion technique can be used for this purpose. The Wi-Fi sensor and MEMS(accelerometer and magnetometer in smart-phone) data fused using HMM can predict position more accurately than a stand-alone system using only Wi-Fi signals. Using this method,accuracy of 1-2 m can be achieved.

So, the use of probabilistic approach with sensor fusion can help us achieve desired performance accuracy level for Wi-Fi based indoor positioning systems however the use of this method needs to balance the accuracy and time-commitment for collecting data when creating signal strength database.

### 11.2. Future Enhancement

The project at its current state can show 2D position of an android device in a map.

1. A commercial product with multiple user support.
2. Render 3d maps .
3. Incorporate every floor of building.
4. Integrate pressure Sensor for floor information or height in building.
5. Monitoring system .

## References


[1] Li, Huaiyu, Xiuwan Chen, Guifei Jing, Yuan Wang, Yanfeng Cao, Fei Li, Xinlong Zhang, and Han Xiao. "An Indoor Continuous Positioning Algorithm on the Move by Fusing Sensors and Wi-Fi on Smartphones." *Sensors* 15.12 (2015): 31244-1267. Web.

[2] He, Xiang, Daniel Aloi, and Jia Li. "Probabilistic Multi-Sensor Fusion Based Indoor Positioning System on a Mobile Device." *Sensors* 15.12 (2015): 31464-1481. Web.

[3] Duvallet, F., and A.d. Tews. "WiFi Position Estimation in Industrial Environments Using Gaussian Processes."*2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*(2008): n. pag. Web.

[4] Zhu, Nan, Hongbo Zhao, Wenquan Feng, and Zulin Wang. "A Novel Particle Filter Approach for Indoor Positioning by Fusing WiFi and Inertial Sensors."*Chinese Journal of Aeronautics* 28.6 (2015): 1725-734. Web.

[5] Haute, Tom Van, Eli De Poorter, Pieter Crombez, Filip Lemic, Vlado Handziski, Niklas Wirstrom, Adam Wolisz, Thiemo Voigt, and Ingrid Moerman. "Performance Analysis of Multiple Indoor Positioning Systems in a Healthcare Environment." *International Journal of Health Geographics Int J Health Geogr* 15.1 (2016): n. pag. Web.

[6] Zheng, Lingxiang, Wencheng Zhou, Weiwei Tang, Xianchao Zheng, Ao Peng, and Huiru Zheng. "A 3D Indoor Positioning System Based on Low-cost MEMS Sensors."*Simulation Modelling Practice and Theory* 65 (2016): 45-56. Web.

[7] Torres-Sospedra, Joaquin, Raul Montoliu, Adolfo Martinez-Uso, Joan P. Avariento, Tomas J. Arnau, Mauri Benedito-Bordonau, and Joaquin Huerta. "UJIIndoorLoc: A New Multi-building and Multi-floor Database for WLAN Fingerprint-based Indoor Localization Problems." *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (2014): n. pag. Web.

[8] Chen, Wei, Weiping Wang, Qun Li, Qiang Chang, and Hongtao Hou. "A Crowd-Sourcing Indoor Localization Algorithm via Optical Camera on a Smartphone

Assisted by Wi-Fi Fingerprint RSSI."*Sensors* 16.3 (2016): 410. Web.

[9] Chang, Qiang, Samuel Van De Velde, Weiping Wang, Qun Li, Hongtao Hou, and Steendam Heidi. "Wi-Fi Fingerprint Positioning Updated by Pedestrian Dead Reckoning for Mobile Phone Indoor Localization." *Lecture Notes in Electrical Engineering China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*(2015): 729-39. Web.

[10] Seidel, S.y., and T.s. Rappaport. "A Ray Tracing Technique to Predict Path Loss and Delay Spread inside Buildings."*[Conference Record] GLOBECOM '92 - Communications for Global Users: IEEE*(n.d.): n. pag. Web.

[11] Thrun, Sebastian, Wolfram Burgard, and Dieter Fox.*Probabilistic Robotics*. Cambridge, MA: MIT, 2005. Print.

[12] Rappaport, Theodore S.*Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice Hall PTR, 1996. Print.

[13] Seybold, John S.*Introduction to RF Propagation*. Hoboken, NJ: Wiley, 2005. Print.

[14] Stallings, William. Wireless Communications and Networks. Upper Saddle River, NJ: Pearson Prentice Hall, 2005. Print.

[15] Kim, Jeong Won, Han Jin Jang, Dong-Hwan Hwang, and Chansik Park. "A Step, Stride and Heading Determination for the Pedestrian Navigation System."*Journal of Global Positioning Systems J. of GPS* 3.1&2 (2004): 273-79. Web.

[16] Rasmussen, Carl Edward., and Christopher K. I. Williams.*Gaussian Processes for Machine Learning*. Cambridge, MA: MIT, 2006. Print.

[17] Russell, Stuart J., and Peter Norvig.*Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall, 1995. Print.

[18] Ferris, B., D. Haehnel, and D. Fox. "Gaussian Processes for Signal Strength-Based Location Estimation." *Robotics: Science and Systems II*(2006): n. pag. Web.

[19] Fujii, Kenjiro, Hiroaki Arie, Wei Wang, Yuto Kaneko, Yoshihiro Sakamoto, Alexander Schmitz, and Shigeki Sugano. "Improving IMES Localization Accuracy by Integrating Dead Reckoning Information." *Sensors* 16.2 (2016): 163. Web.

[20] Li, T., S. Sun, and T.p. Sattar. "Adapting Sample Size in Particle Filters through KLD-resampling." *Electronics Letters*49.12 (2013): 740-42. Web.

[21] Richter, Philipp, and Manuel Toledano-Ayala. "Revisiting Gaussian Process Regression Modeling for Localization in Wireless Sensor Networks." *Sensors 15.9*(2015): 22587-2615. Web.

## Appendix A: HEAT MAP OF ACCESS POINTS

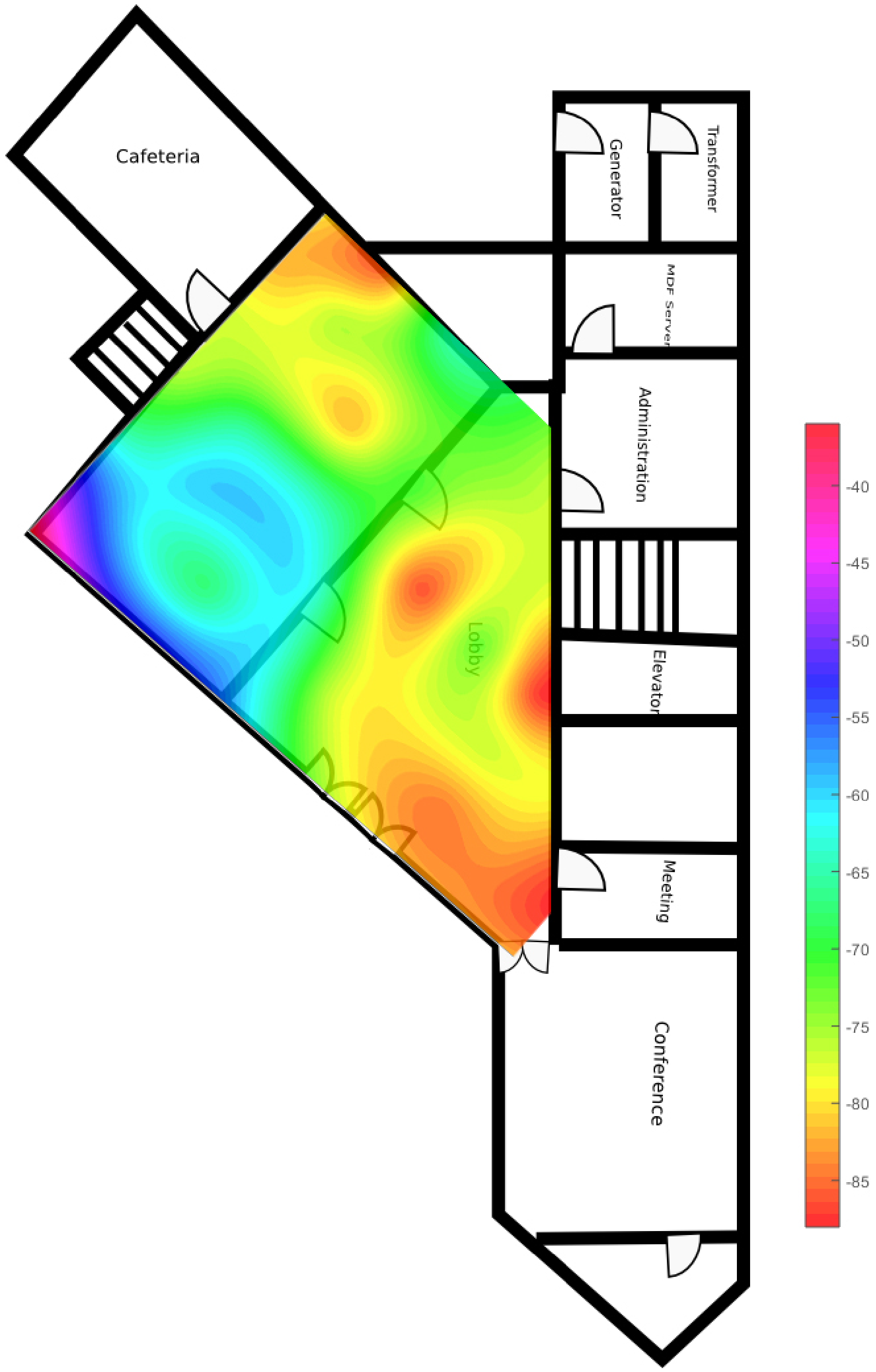



Figure A.1: Heat Map of RSSI value of AP1

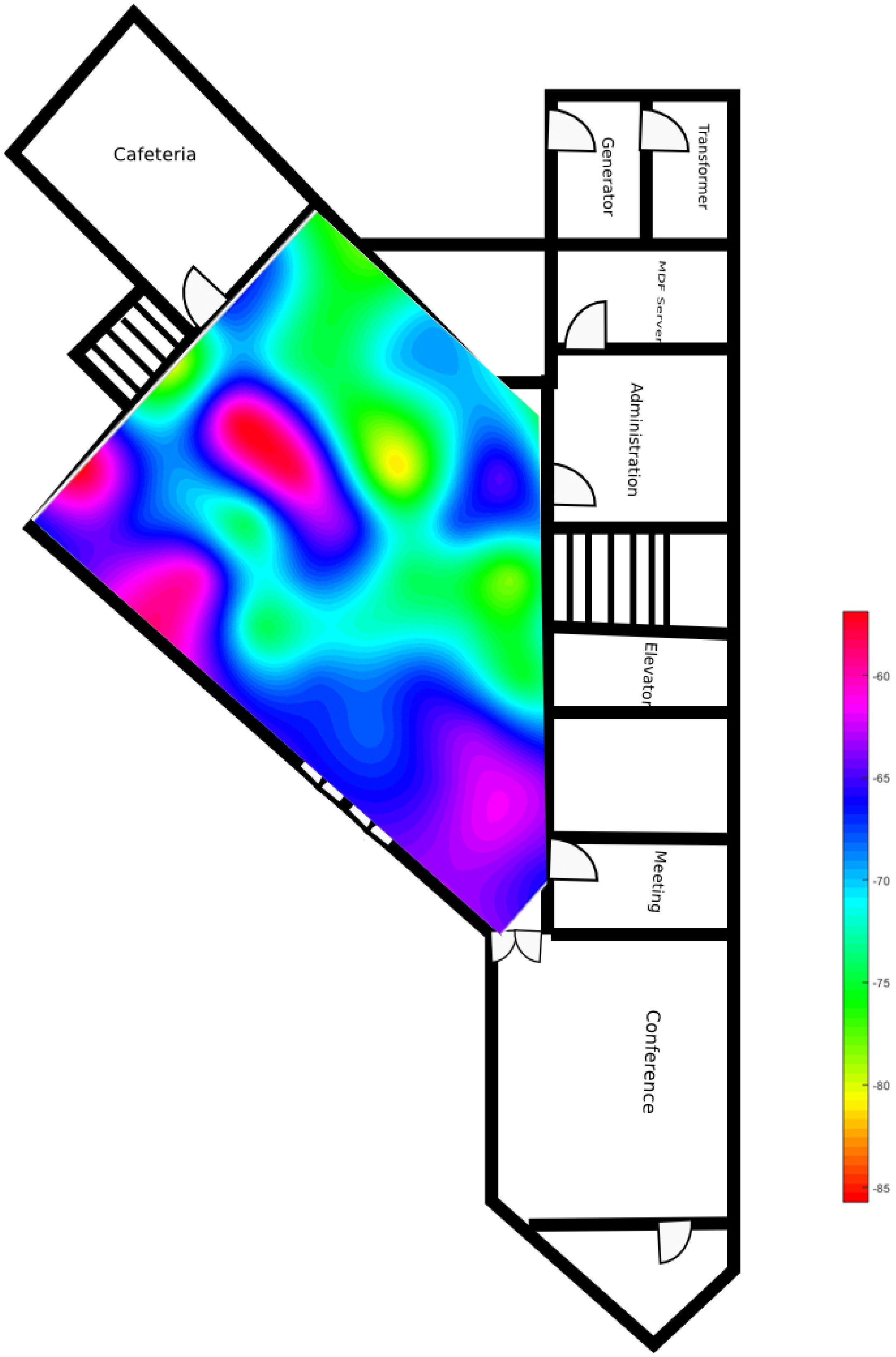



Figure A.2: Heat Map of RSSI value of AP2

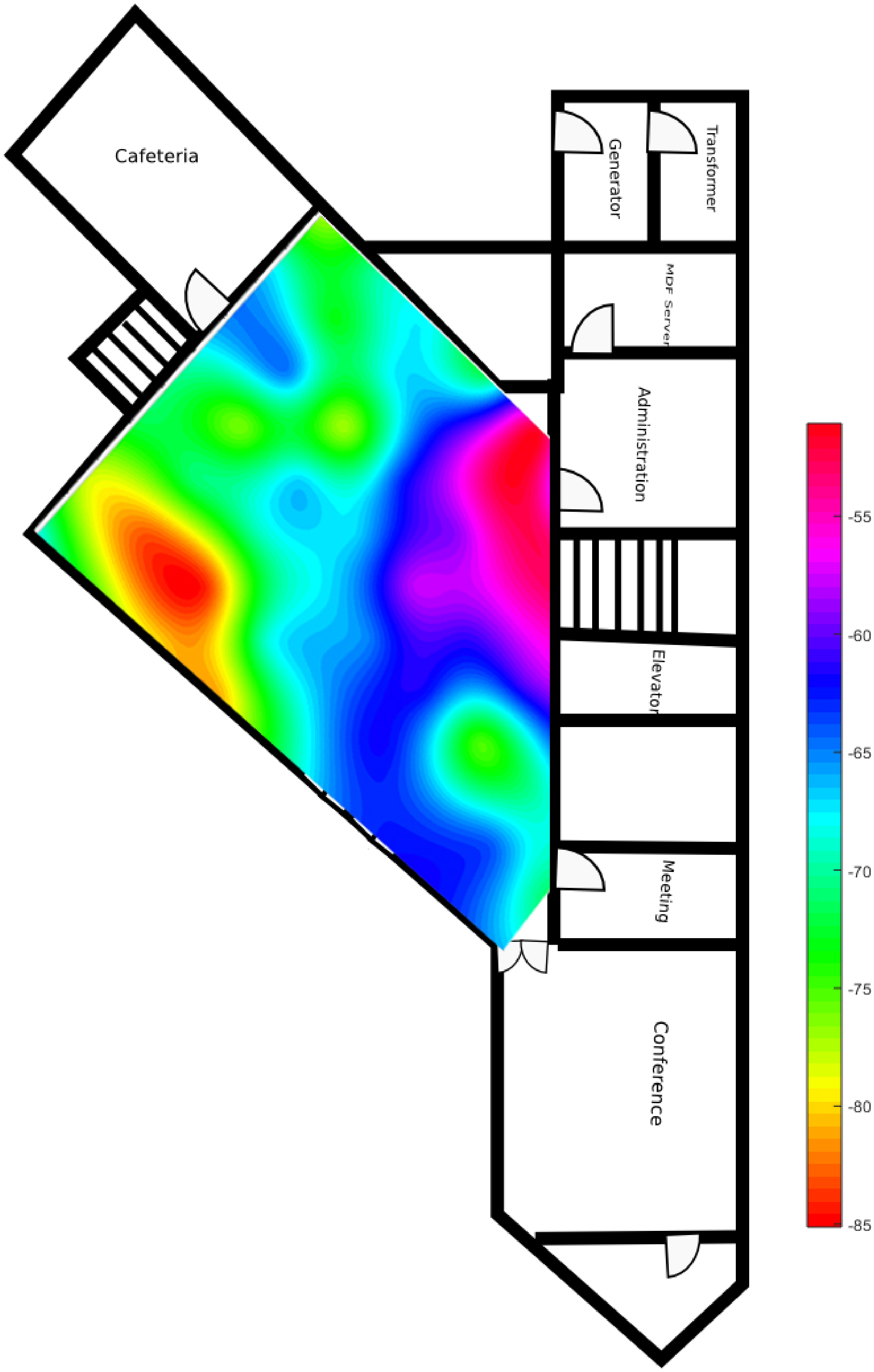



Figure A.3: Heat Map of RSSI value of AP3

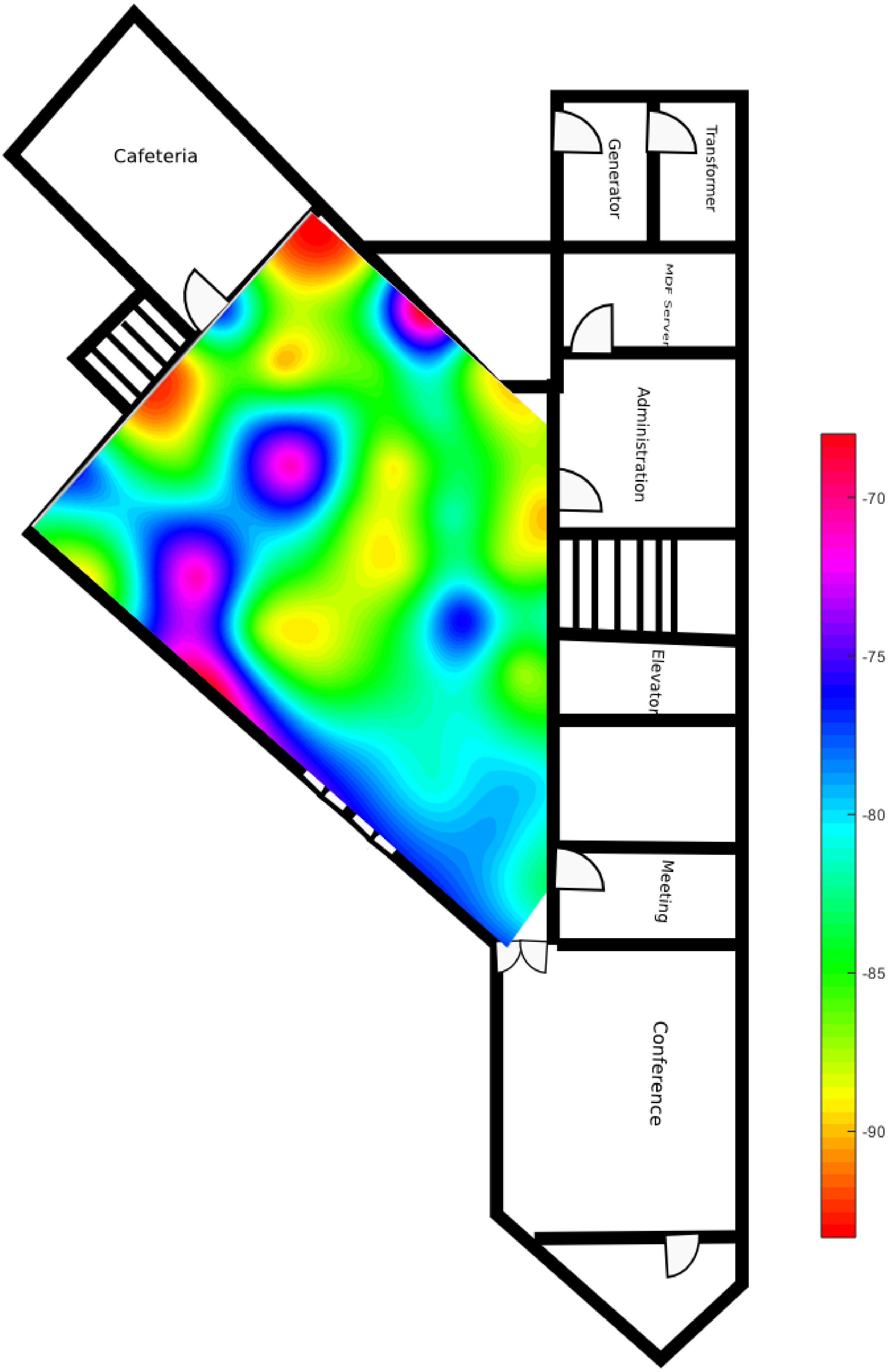



Figure A.4: Heat Map of RSSI value of AP4

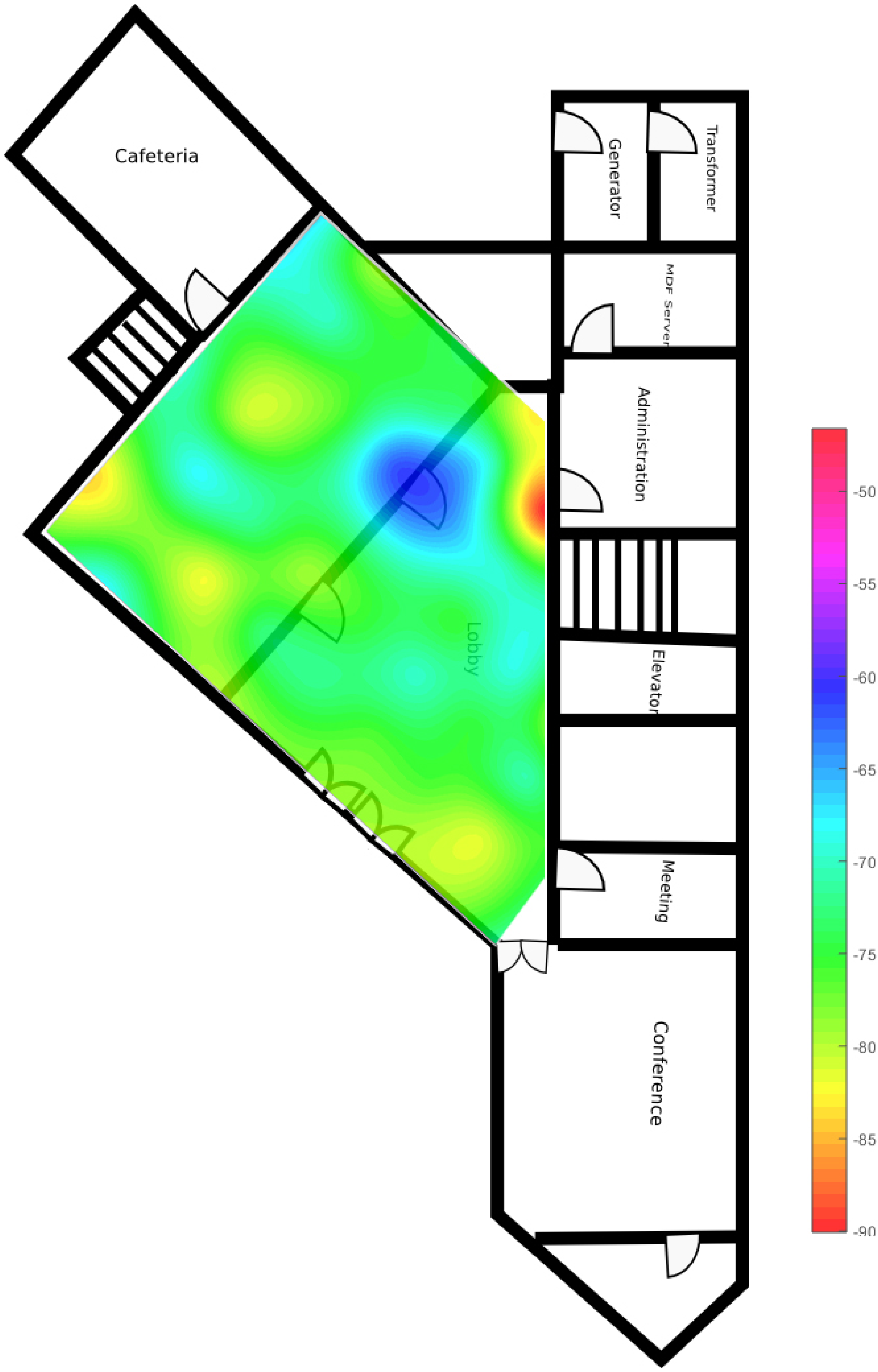



Figure A.5: Heat Map of RSSI value of AP5

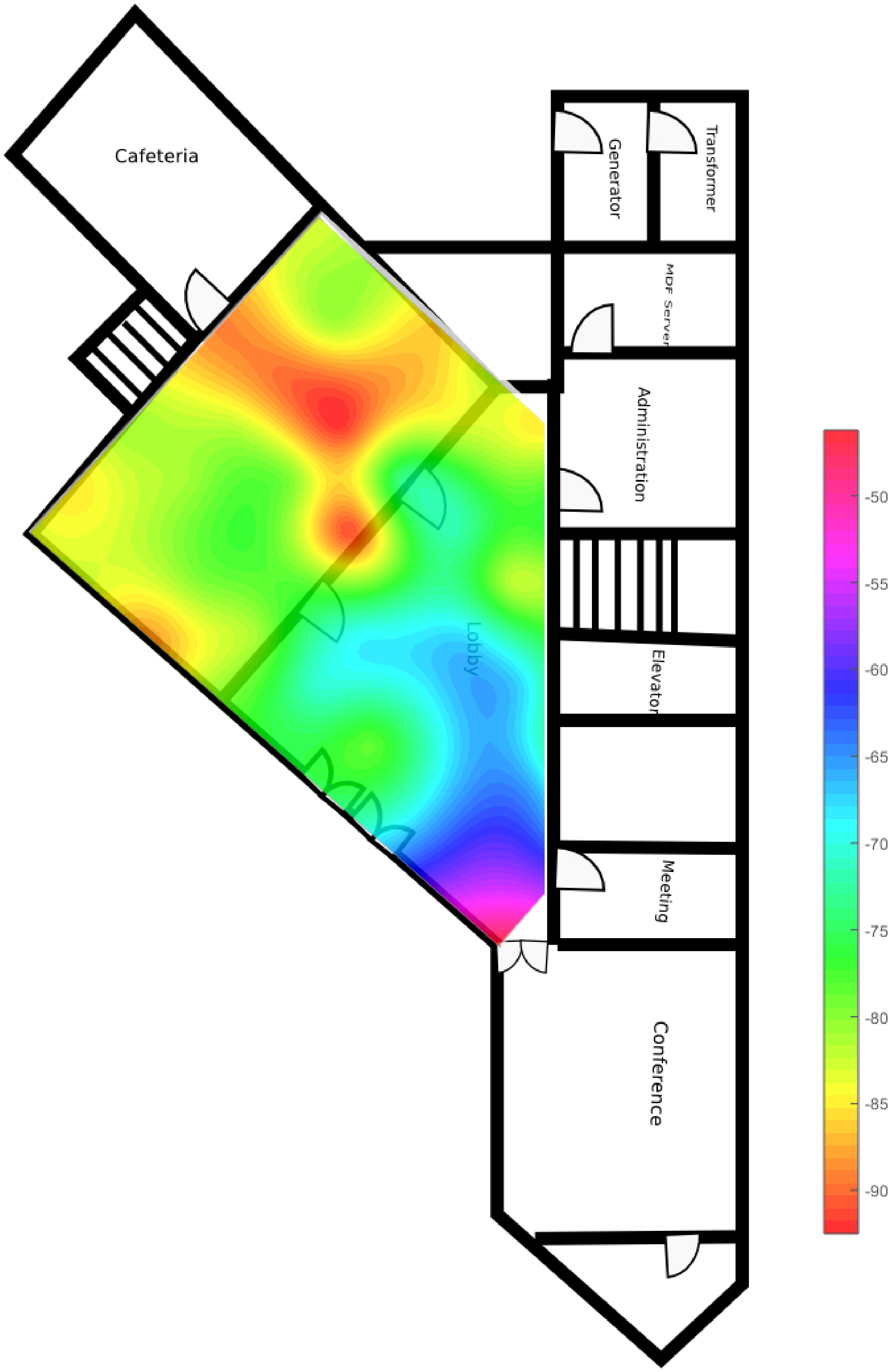



Figure A.6: Heat Map of RSSI value of AP6

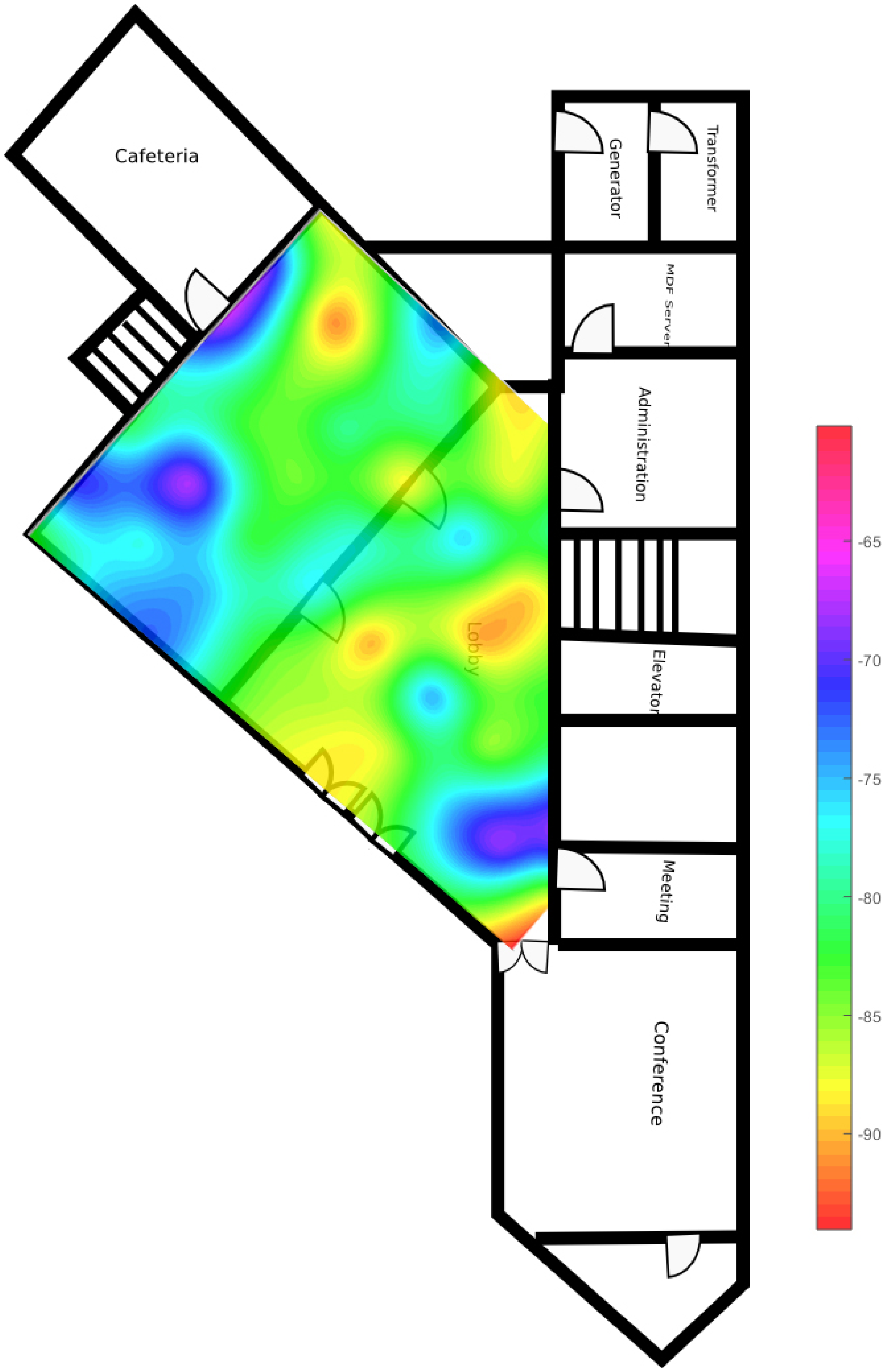



Figure A.7: Heat Map of RSSI value of AP7

# Appendix B: Software Listing

## B.1. Programming Languages

### B.1.1. Python

Used for implementation of WKNN, sequential monte carlo localization, GPR and socket server

### B.1.2. Matlab

Used for graphical analysis of results obtained

### B.1.3. Java

for android application Development

### B.1.4. HTML/JavaScriptUsed

for web application development

## B.2. Libraries, APIs and Classes

### B.2.1. pyGPs

Python library for Gaussian Processes Regression

### B.2.2. Socket

Python module for socket communication implementation

### B.2.3. Mapbox GL JS

Mapbox GL JS is a JavaScript library that uses WebGL to render interactive maps for desktop and mobile platforms

### B.2.4. GDAL

GDAL is a translator library for raster and vector geospatial data formats. As a library, it presents a single raster abstract data model and single vector abstract data model to the calling application for all supported formats. It also comes with a variety of useful command line utilities for data translation and processing.