Masters Project Final Report

December 2015

Project Title	Privacy and Security Implications on Wireless (Wi-Fi) Tomography		
Student Name	Jithendra Ravinath Ovitigala Sirimanne		
Registration No. & Index No.	2013/MIS/021 13770213		
Supervisor's Name	Dr. Chamath Keppitiyagama		
Please Circle the appropriate	**Master's Program**	**Type**	
	MIS	**Research**	**Implementation**
For Office Use Only			

Privacy and Security Implications on Wireless (Wi-Fi) Tomography

J R O Sirimanne

2015

Privacy and Security Implications on Wireless (Wi-Fi) Tomography

A dissertation submitted for the Degree of Master of Science in Information Security

J R O Sirimanne

University of Colombo School of Computing

2015

Declaration

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Students Name: J R O Sirimanne

Signature: Date: 30th December 2015

This is to certify that this thesis is based on the work of Mr. J R O Sirimanne under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name: Dr. Chamath Keppitiyagama

Signature: Date:

Abstract

This research is aimed to be a proof of concept for a Privacy and Security Implications on Wireless (Wi-Fi) Tomography. Wi-Fi is a popular wireless networking technology used today which uses radio waves to transmit data. For this research Radio tomographic imaging (RTI) technologies are used to prove that the privacy can be breached just by analyzing the wireless signals receive signal strength indicator (RSSI) value. This analysis can be done by anyone who has a simple wireless card that shows the RSSI value. Hence one's privacy can be breached even without him knowing his movements are being analyzed by an outsider.

First it is required to prove that there is an identifiable significant difference in the RSSI values when there is a moving object obstructing the wireless signal. For this real world scenarios are developed, data is gathered and analyzed. I will be using a statistical approach to solve this problem.

Secondly this research tries to solve how accurately we can localize that moving object inside a scenario or simplify identify the objects whereabouts. Probability is an area that can be used for problems like this. So using a probability based approach is needed.

This thesis is organized to show my research methodologies and knowledge obtained by analyzing the data in the area mentioned above.

Acknowledgement

Firstly, I would like to express my sincere gratitude to my advisor Dr. Chamath Keppitiyagama for the continuous support for my research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Secondly I would like to express my appreciations to Mr.Asanka Sayakkara who provided the necessary scripts to gather data and his continued guidance for analysis of gathered data and also for the support given to me to complete this thesis.

I would like to express a special thanks to all the staff members of the Network Operating Center of University Of Colombo School Of Computing for their immense support on providing the technical support and required hardware to make this research a success.

Beside those my heartiest appreciation should go to all the lecturers, assistant lecturers, demonstrators and other staff members in the University Of Colombo School Of Computing who sincerely helped me in numerous ways.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

GPS	Global Positioning System
RADAR	Radio Detection and Ranging Systems
RFID	Radio Frequency Identification
MRI	Magnetic Resonance Imaging
RSSI	Receive Signal Strength Indicator
RTI	Radio Tomographic Imaging
MAC	Media Access Control
dBm	decibel-milliwatt
WLAN	Wireless Local Area Network
LOS	Line of Sight
MIMO	Multiple-Input and Multiple-Output

Table 1 - List of Abbreviations

Chapter 1

1. Introduction

At present day a very large portion of the communication is done using radio waves. These waves use electromagnetism to generate and propagate through space at the speed of light. There are many military and general purpose applications of radio waves that are used for tracking humans, objects and even gestures.

Some of these technologies are Global positioning System (GPS), Sonars, Radio Detection and Ranging Systems (RADAR). From these GPS can accurately show your position anywhere on earth. High grade GPS systems can have an accuracy level of up to 4 meters. Sonars and RADAR can track both moving and stationary objects. Without RADAR modern travel systems like air travel will be unthinkable. Most of these technologies are first developed for military purpose and then generalized for public use.

RFID or Radio Frequency Identification is another immerging technology that uses radio waves to track goods and merchandize. RFID uses a tag which has a circuit that transmits data via radio waves and this data can be read from a RFID reader to identify tagged goods. Apart from that it is also being used for access control since RFID tags are small and they can be easily added to access control cards.

X-Ray is the most popular imaging technology used in medical applications. X-Rays are also electromagnetic waves which uses its penetration power for imaging. Magnetic Resonance Imaging (MRI) is a revolutionary imaging technology that is used in medical field applications. MRI uses a magnetic field and pulses of radio wave energy to make pictures of organs and structures inside the body.

All these technologies use special transmitters and receivers to archive their task. While these are very effective and efficient technologies, the cost factor of implementing such a system is a huge challenge.

The word tomography is derived from the Greek word *tomos* which means sections or sectioning and the process of tomography involves the generation of narrow sections through an object. Tomography is also one of the emerging technologies in this area which is used to localize passive objects or in other words objects which does not carry any tracking devices inside an area of interest. Because of that Radio tomographic imaging (RTI) is also called “device free localization” [14] “passive localization” [15], or “sensor less sensing” [16].

Wi-Fi is the popular wireless networking technology that uses radio waves to provide wireless high-speed Internet and network connections. Wi-Fi connections support millions of people in homes, businesses, and public locations around the world supporting them to be connected at every possible time. According to statistics there is a huge growth in Wi-Fi enabled devices and Wi-Fi hotspots within past few years and it is expected to grow more and more. There is a possibility to track people movements by observing just the signal strength of these Wi-Fi networks.

1.1 Hardware Devices

For this study an ordinary wireless access point and a Laptop with wireless (Wi-Fi) interface that is capable of publishing signal strengths in decibel (dB) values will be used as hardware devices. By this I’m trying to generalize this study and match it maximum to a real world scenario.

1.1.1 Wireless access point

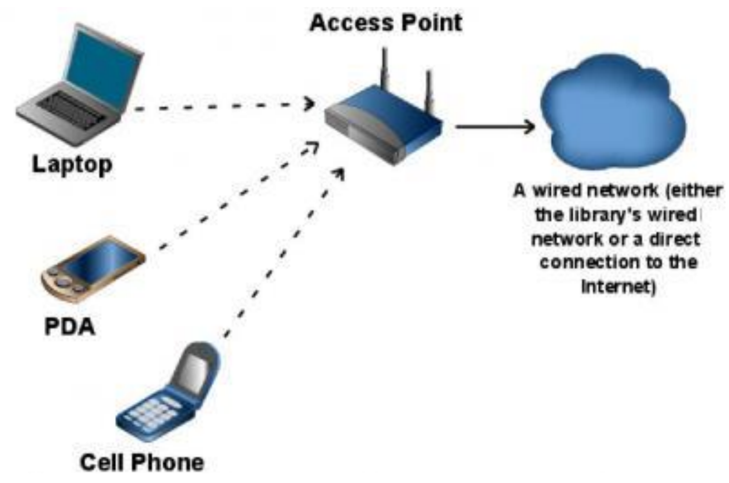


Figure 1- Wireless Access point Simplest usage

(Source: http://www.techsoupforlibraries.org/files/images/wireless_network.img_assist_custom.jpg)

In computer networking, a wireless access point (AP) is a device that allows wireless devices to connect to a wired network using Wi-Fi, or related standards. The AP usually connects to a router (via a wired network) as a standalone device, but it can also be an integral component of the router itself. [8].

Figure 2 - DSL-2750U Wireless N ADSL2+ Wi-Fi Router

(Source: http://115.124.123.225/new/products/DSL-2750u/DSL-2750u.png)

For this study I used a D-Link Wireless N300 ADSL Modem Router (Model Number DSL-2750U) which is a very common wireless access point used in Sri Lanka. The DSL-2750U Wireless N ADSL2+ Wi-Fi Router connects a group of users to the Internet, allowing multiple computers at home or the office to share an integrated high-speed ADSL2/2+ interface. It provides high-performance 802.11n wireless access for wireless networked computers, 4 built-in Ethernet ports, firewall protection, and QoS for smooth and secure download/upload of photos, files, music, video, and e-mail over the Internet. [9]

1.1.2 Signal Strength Capturing device

Captured received signal strength indicator (RSSI) values are the main input data for this study. As mentioned earlier [10] hardware manufacturers are not oblige to provide RSSI value. So when we choose a computer/laptop to capture the RSSI values the wireless card should advertise the RSSI value. Here for the data capturing we have used a general purpose Dell Laptop (model no: Dell Inspiron 14R (N4010) Laptop) with DW1501 Wireless-N Wlan Half-Mini Card which is again a very common device in Sri Lanka market. This wireless card supports IEEE 802.11b/g/n single band mode in 2.4GHz mode. [12]

Figure 3 - Dell 14R n4010 General Purpose laptop

(*http://1.bp.blogspot.com/-jPEiZYrgbc8/VAx1_-yiOtI/AAAAAAAAAF0/DGvRA2iAIRw/s1600/dell-inspiron-14r-disassembly.png*)

Figure 4 - Dell Wireless DW1501 Mini Card

(*Source: http://www.wireless-driver.com/wp-content/uploads/2011/11/DellWireless1501802.11bgMiniCard.jpg*)

1.2 Motivation

There are many security implementations to preserve privacy of a person when connected and using the network. Research shows when an object moves inside a wireless area that objects causes the received signal strength indicator (RSSI) to be dropped. By tracking these drops we can map the path and movements of the object. Apart from providing internet and network access Wi-Fi is used in sensor networks to transfer data gathered from sensors. These sensor networks power and enable modern concepts like smart homes, smart power, smart cities, Internet of things etc. With all these usage of Wi-Fi increases rapidly and in a small area there can be many Wi-Fi networks. Issue with this is by using wireless tomography technologies we can track movements of people unknowing to them that someone is tracking them.

Currently there are no security implementations to safeguard people for above type of privacy breaches and many are unaware about the security issues related to available tomographic techniques.

1.3 Aims and Objectives

Wireless internet is a very common method of providing Internet and networking facilities to people and devices. With such common used technology it should be safe to use. In this research, I will design, test and evaluate diffident real world scenarios of wireless implementations to determine and prove the effects to wireless signal strength when a human is present and find the scenarios that are vulnerable to privacy breaches.

1.4 Research Question

With this study I'm searching to answer two questions that arise with above mentioned problems.

1. Is there a significant amount of wireless signal strength drop when there is a human inside the line of sight of the wireless access point and data gathering computer?
2. How accurately we can identify human presence by observing received signal strength indicator (RSSI) values.

1.5 Scope and Limitations

This study will be carried out using the indoor scenarios and does not consider about outdoor scenarios. Also this will be a study to prove the concept of object tracking and localization based on the RSSi value of the Wi-Fi signal using general purpose devices. Tracking and localization only consider about the changes in RSSI value when the signal is obstructed by an object. This will not consider the other factors like environmental variations, Wi-Fi signal overlaps due to other transmitting devices and radio wave noise. Also this object and data gathering device is in line-of-sight with the transmitting base station.

Chapter 2

2 Literature Review

This section will cover the related literature for this research.

2.1 Decibel

The decibel (dB) is a logarithmic unit that expresses the ratio of two values of a physical quantity, often power or intensity. One of these quantities is often a reference value and in this case the decibel expresses the absolute level of the physical quantity. The number of decibels is ten times the logarithm to base 10 of the ratio of two power quantities or of the ratio of the squares of two field amplitude quantities. One decibel is one tenth of one bel, named in honor of Alexander Graham Bell; however, the bel is seldom used. The definition of the decibel is based on the measurement of power in telephony of the early 20th century in the Bell System in the United States. Today, the unit is used for a wide variety of measurements in science and engineering, most prominently in acoustics, electronics, and control theory. In electronics, the gains of amplifiers, attenuation of signals, and signal-to-noise ratios are often expressed in decibels. The decibel confers a number of advantages, such as the ability to conveniently represent very large or small numbers, and the ability to carry out multiplication of ratios by simple addition and subtraction. By contrast, use of the decibel complicates operations of addition and subtraction. [7]

2.2 Received Signal Strength Indicator (RSSI)

In telecommunications, received signal strength indicator (RSSI) is a measurement of the power present in a received radio signal. RSSI is usually invisible to a user of a receiving device. However, because signal strength can vary greatly and impact functionality in wireless networking, IEEE 802.11 devices often make the measure available to users. In an IEEE 802.11 system, RSSI is the relative received signal strength in a wireless environment, in arbitrary units. RSSI is an indication of the power level being received by the antenna. Therefore, the higher the RSSI number, the stronger the signal.

There is no standardized relationship of any particular physical parameter to the RSSI reading. The 802.11 standard does not define any relationship between RSSI value and power level in mW or dBm. Vendors and chipset makers provide their own accuracy, granularity, and range for the actual power (measured as mW or dBm) and their range of RSSI values (from 0 to RSSI_Max). One subtlety of the 802.11 RSSI metric comes from how it is sampled, RSSI is acquired during only the preamble stage of receiving an 802.11 frame, not over the full frame. [10]

2.3 Privacy

Privacy has many meanings. The most general is freedom from interference or intrusion, the right "to be let alone," a formulation cited by Louis Brandeis and Samuel Warren in their groundbreaking 1890 paper on privacy. [4] This recognizes that each person has a sphere of existence and activity that properly belongs to that individual alone, where he or she should be free of constraint, coercion, and even uninvited observation. As we would say today, each of us needs our own "space." Most would recognize the protected sphere to include personal opinions, personal communications, and how one behaves behind closed doors, at least as long as these do not lead to any significant threats to society. Many would also include behavior within the family and other intimate relationships in that sphere. [3]

2.4 Wi-Fi

Figure 5- Wi-Fi logo

(Source: https://upload.wikimedia.org/wikipedia/commons/thumb/3/32/Wi-Fi_Logo.svg/2000px-Wi-Fi_Logo.svg.png)

Wi-Fi uses radio waves to create networks and transmit data. Radio waves are electromagnetic waves. Electromagnetic waves have a large range of frequency starting from as low as 3 kHz and ranging up to 300 GHz. Wi-Fi is a local area wireless computer networking technology that allows electronic devices to network, mainly using the 2.4 GHz UHF and 5 GHz industrial, scientific and medical (ISM) radio band. [5] Wi-Fi standards are defined by IEEE 802.11. IEEE 802.11 is a set of MAC and physical layer specifications for implementing WLAN computer communications. They are created and maintained by the IEEE LAN/MAN Standards Committee (IEEE 802). The base version of the standard was released in 1997, and has had subsequent amendments. The standard and amendments provide the basis for wireless network products using the Wi-Fi brand. While each amendment is officially revoked when it is incorporated in the latest version of the standard, the corporate world tends to market to the revisions because they concisely denote capabilities of their products. As a result, in the market place, each revision tends to become its own standard. [6]

Electromagnetic waves can be affected by many reasons when they are traveling through space.

- Signal Frequency
- Transmission medium
- Objects encountered

These reasons will result in reflection (the wave partially bounces of an object), refraction (change of direction when passing from one medium to another), absorption (loss of energy when an object is hit), diffraction (when waves are bend and spread around an obstacle), scattering (wave bounces off in multiple directions) and polarization (orientation of the oscillations of the waves can change upon interaction) of the signal.

2.5 Software Defined Radio

Software-defined radio (SDR) is a radio communication system where components that have been typically implemented in hardware such as mixers, filters, amplifiers, modulators/demodulators, detectors, etc. are instead implemented by means of software on a personal computer. GNURadio is an example to this. These SDR devices are relatively expensive but they have wide range of applications when comes to tracking. SDR's can be used to track a single human, track multiple humans and gesture decoding. Researchers in MIT have done extensive work on tracking using SDR devices. They have shown the ability to detect humans behind walls [17], and then they have shown the possibility of 3D tracking of the position of a human body [18]. The SDR device which they have used is made of USRP N210. Another mid-range device capable of these types of work is HackRF.

2.6 Radio Tomographic Imaging

Tomography refers to imaging by sections or sectioning, through the use of any kind of penetrating wave. Radio tomographic imaging (RTI) is an emerging application which offers a new way to image passive objects in buildings and outdoor environments using received signal strength indicator (RSSI). [1] RTI can be used in emergencies, rescue operations, and security breaches, since the objects being imaged need not carry an electronic device. Using the images to track humans moving through a building, for example, provides a basis for new applications in security systems and “smart” buildings.

2.7 Related Work

2.7.1 RADAR: An In-Building RF-based User Location and Tracking System. [13]

Paramvir at al. [13] have done research on tracking indoor objects using WIFI signal strength way back in year 2000. They have done this by using 3 base stations and 1 mobile node in an indoor test environment. Each base station and the mobile host was equipped with a Digital RoamAbout network interface card (NIC), based on Lucent’s popular WaveLAN RF LAN technology. The network operates in the 2.4 GHz band. The range of the network, as specified in [Roa96], is 200 m, 50 m, and 25 m, respectively, for open, semi-open, and closed office environments. The base stations provide overlapping coverage in portions of the floor, and together cover the entire floor. They have used several data analytic methods to evaluate the accuracy of the user location.

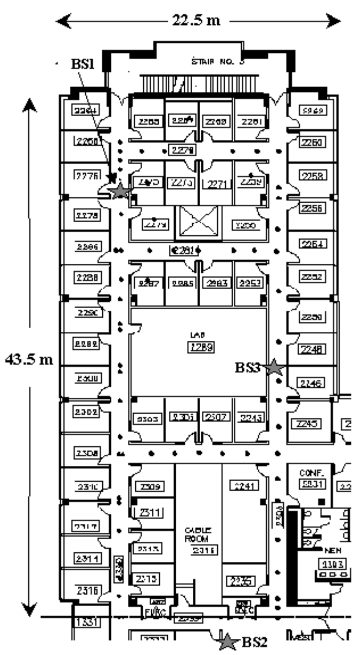


Figure 6 - Map of the floor where the experiments were conducted. The black dots denote locations were empirical signal strength information was collected. The large stars show the locations of the 3 base stations. The orientation is North (up) and East (right).

The mobile host is transmitting 4 UDP packets per second with 6-byte payload size. Then these packets are being captured by the three base stations. Base stations record the timestamp (t), Base station id (bs) and signal strength (ss). The mobile host records timestamp (t), users coordinates in the root (x, y) and the direction (d) user is facing when recording the data.

When the data is analyzed using the signal strength we can construct a graph as shown the figure. Nearest neighbor(s) in signal space (NNSS) technique is used to identify multiple locations and pick the one that best matches the observed signal strength. $\text{Euclidian Distance} = \sqrt{(ss1 - ss'1)^2 + (ss2 - ss'2)^2 + (ss3 - ss'3)^2}$ is used for this purpose.

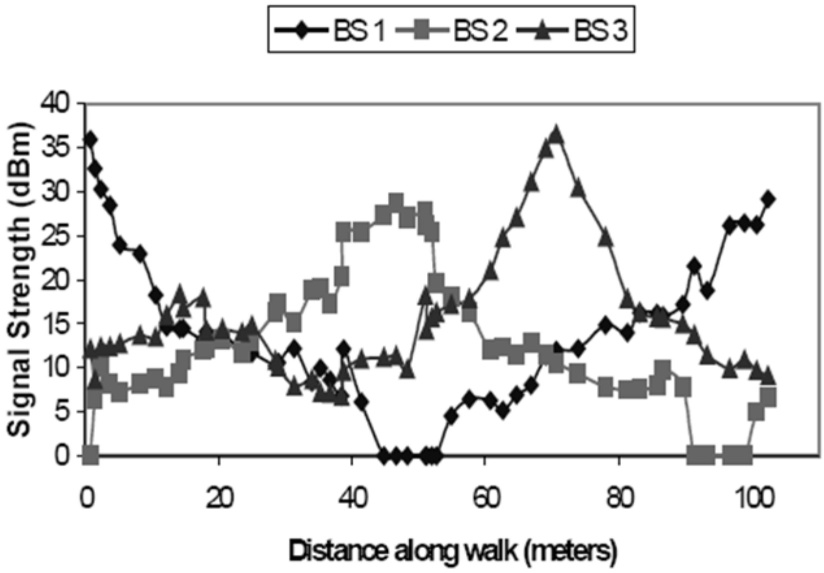


Figure 7 - Signal strength recorded at the three base stations

2.7.2 Radio Tomographic Imaging with Wireless Networks [1]

Joey Wilson and Neal Patwari have done many contributions and research on the field of RTI. By analyzing the received signal strength (RSS) measurements they were able to obtain RTI images of moving objects. They have developed a basic model and image reconstruction technique that has low computational complexity. Their research results show that RTI is capable of imaging the RF attenuation caused by humans in dense wireless networks with inexpensive and standard hardware. Below figure shows an illustration of a RTI network and a reconstruction of a moving object inside that network.

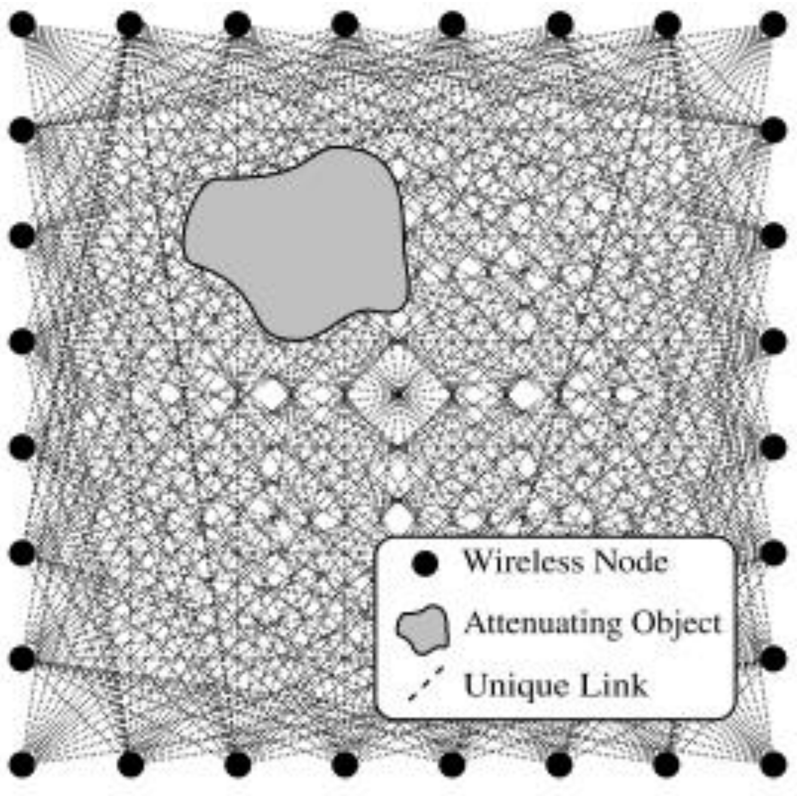


Figure 8 - An illustration of an RTI network. Each node broadcasts to the others, creating many projections that can be used to reconstruct an image of objects inside the network area. [2]

(Source: http://span.ece.utah.edu/uploads/RTINetwork.png)

2.7.3 WiSee[20]

WiSee, is a novel approach to gesture recognition that leverages Wi-Fi signals to enable whole-home sensing and recognition of human gestures. WiSee make use of the property of Doppler shift to archive its target. The Doppler shift is the change in frequency of a wave for an observer moving relative to its source. When applied this to WiSee when a user moves his hand in a wireless environment that movement causes a very small change to the frequency of the wireless signal. Wi-Fi uses 2.4GHz band and 5GHz band for transmissions. When user moves his hand at the rate of 0.5 m/s the frequency shift it generates is 17Hz.

The WiSee is implemented in GNURadio using the USRP-N210 hardware. In an ordinary household there may be multiple people who can affect the wireless signals at the same time. WiSee uses the MIMO capability that is inherent to Wi-Fi (802.11n), to focus on gestures from a specific individual. MIMO provides throughput gains by enabling multiple transmitters to concurrently send packets to a MIMO receiver. By treating the wireless reflections from each human as signals from a wireless transmitter, then individuals can be separated using a MIMO receiver.

WiSee can correctly detect nine gestures with 94% average accuracy from 900 tested gestures. By increasing the receiver antennas WiSee can improve its accuracy and make to track multiple users. Gestures recognized by Wi-See are shown in below figure.

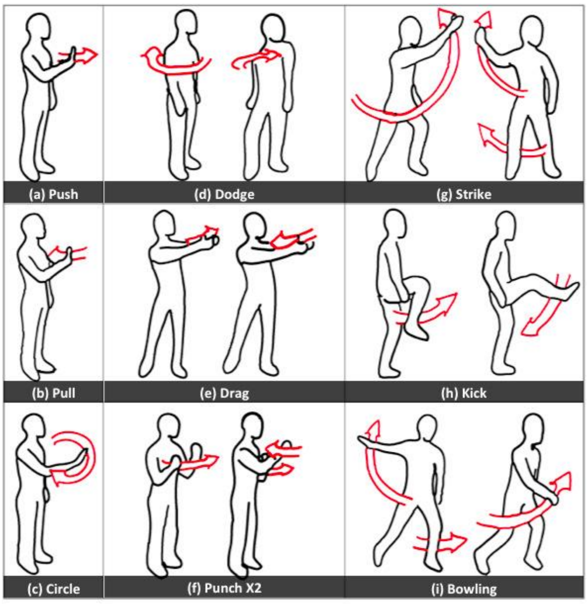


Figure 9 - Gestures that are identified by WiSee

2.8 Related Products

2.8.1 See through wall radar system

This is a commercially available product developed by Camero-Tech Ltd in their XAVER™ product line. This is a very easy to use handheld device which can be quickly deployed in to situations. Unit cost of the product is around $6000USD. It is used by the USA military and law enforcement agencies to gain tactical advantages. This product has a micro power Ultra-Wideband sensor that operates at very high bandwidth which enables reliable detection. Device transmits a radio wave and waits for the reflected energy to be bounced off from objects (including people and animals). The system provides information regarding the number of people, their location and orientation, as well as information on non-living objects including the layout of the room and major infrastructure elements.

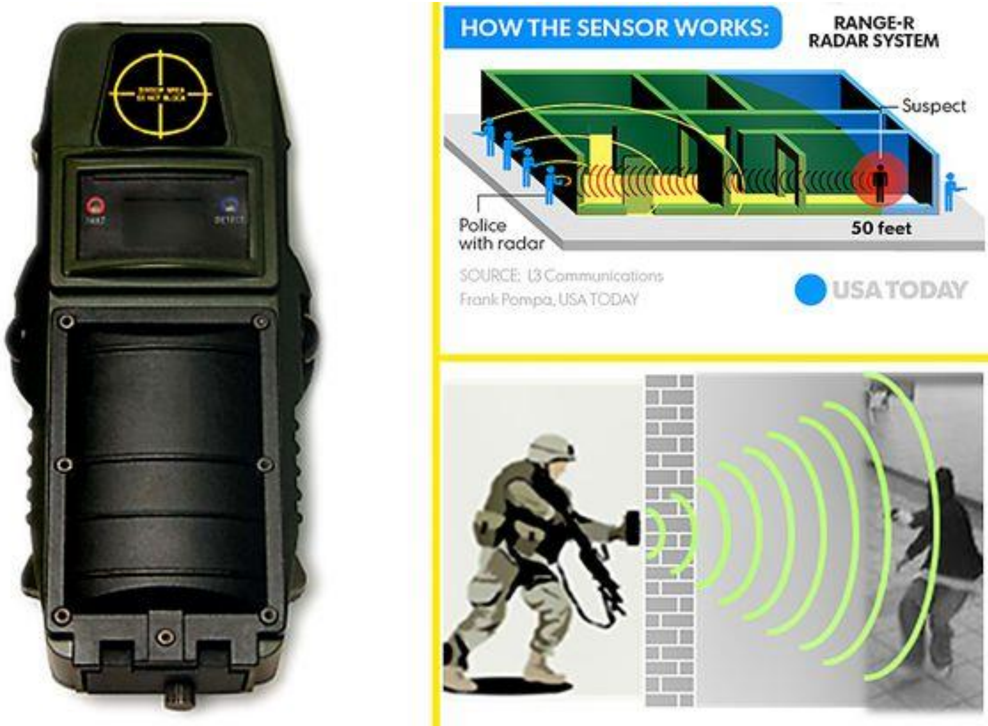


Figure 10 - See through wall radar handheld device and its usage

(*Source:http://i3.mirror.co.uk/incoming/article5010564.ece/ALTERNATES/s615/Radar_MAIN.jpg)*

2.8.2 Google project "soli"

Project soli is an initiative by Google to develop a microchip which can detect human hand movements and gestures. The Soli sensor is able to track sub-millimeter motions at high speed and accuracy. It fits onto a chip. This chip can be produced at scale, and can be used inside even small wearable devices. Radar is the type of the radio wave used in this project. Because radar waves can be fine-tuned for detect tiniest changes in gestures. Each of the movements in the hand generates a unique pattern in the received signal. As shown in the below figure. By analyzing these signal patterns using soli chip programmers are able to detect what the hand is doing and act according to the pre-programed actions in the application.

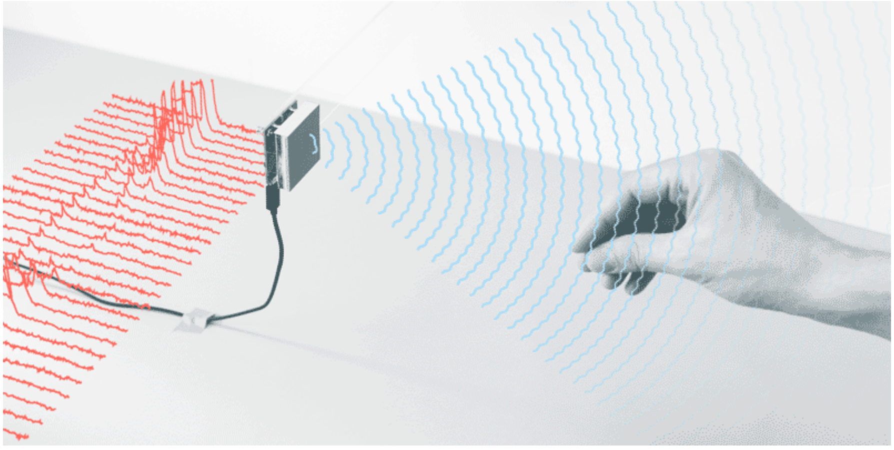


Figure 11 - How the system detects gestures

(*Source:http://o.aolcdn.com/hss/storage/midas/cb197597e76790224b5dcbe34cb1a82c/202060602/0529_soli_2.gif*)

Figure 12 - Practical usages of project "soli" chip

(*Source:http://static1.squarespace.com/static/5438bd1be4b0446f6692b906/t/55767d0fe4b01527f4e85be5/1433828630726/img-googleProjectSoli.png*)

Chapter 3

3 Design and Implementation

In the design stage, scenarios are developed to answer research questions and in the implementation stage developed scenarios are used for gathering test data.

3.1 Scenarios

Data gathering is a key step in this research and scenarios are used to simulate real world examples. Then based on the data we do the analysis and determine the results of the project. First of all there are two main sets of scenario designs; one without a human subject and one with the human subject obstructing the LOS view of the wireless access point and data gathering computer for a period of time. Reason to do is RSSI is not a constant value all the time as shown in the graph below.

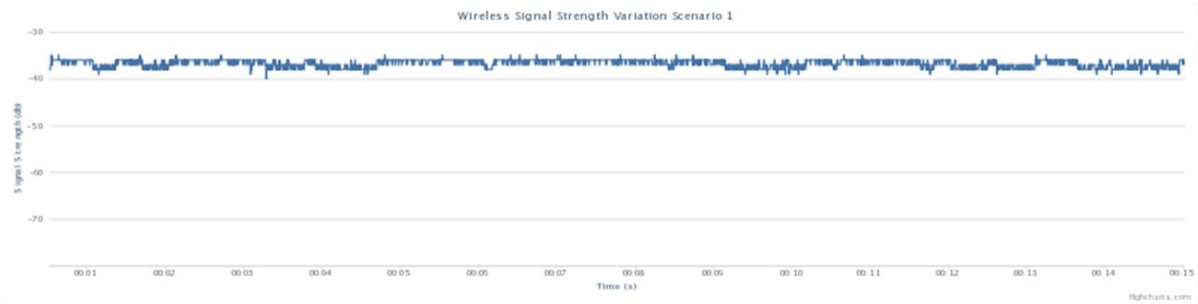


Figure 13 - Wireless signal strength variation with time

In the initial design face primary scenarios are developed to answer first research question(Is there a significant amount of wireless signal strength drop when there is a human inside the line of sight of the wireless access point and data gathering computer?). After initial stages more advanced scenarios are tested in hope to localize and track the passion of the human subject.

For initial scenarios there are three variables

1. Data gathering computer
2. Wireless access point
3. Person

3.2 Scenario Diagrams

In these diagrams we have used symbols to represent objects below table describes the symbols.

Symbol	Resemblance
	Wireless access point
	Data gathering computer
	Human

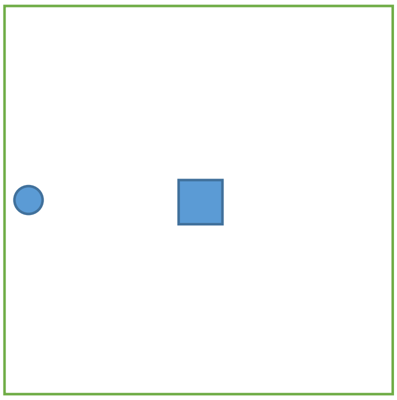
Table 2 - Symbols and resemblance

Scenarios are broken into three sets based on the position of the wireless access point. For an individual set in the initial stage wireless access point will remain stationary and data gathering computer will be the variable.

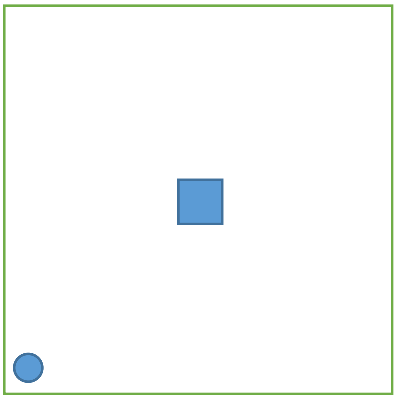
3.2.1 Scenarios Set 1

In the scenario set one wireless access point is stationary in the middle of the room and data gathering computers position is changed around the access point.

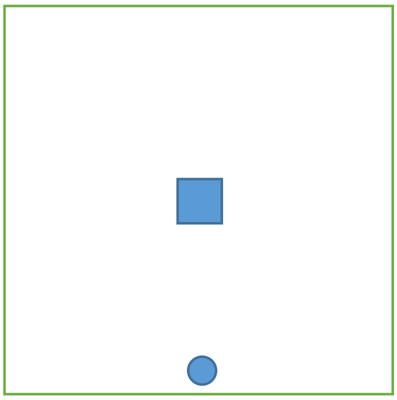
Scenario 1



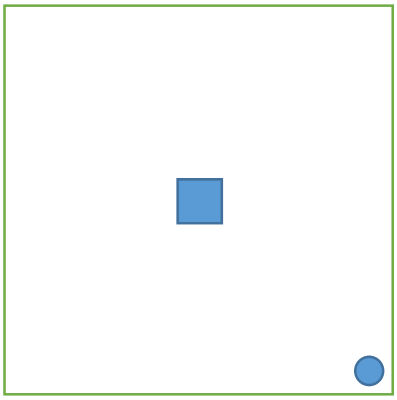
Scenario 2



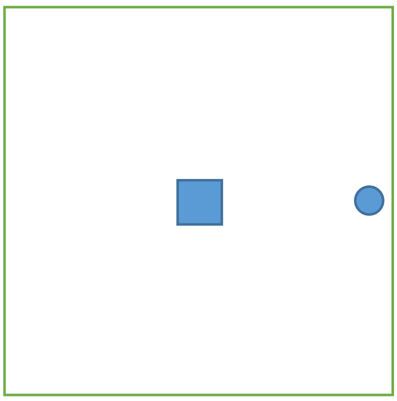
Scenario 3



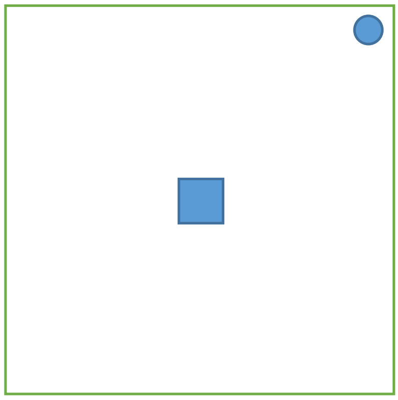
Scenario 4



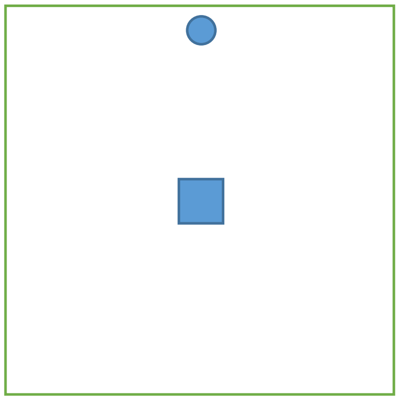
Scenario 5



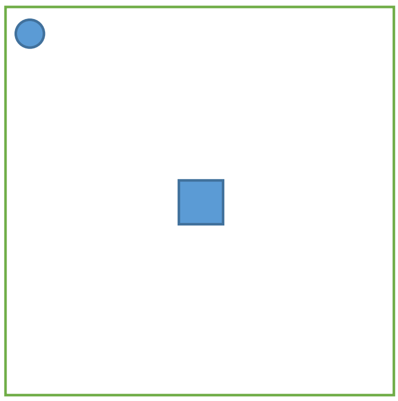
Scenario 6



Scenario 7



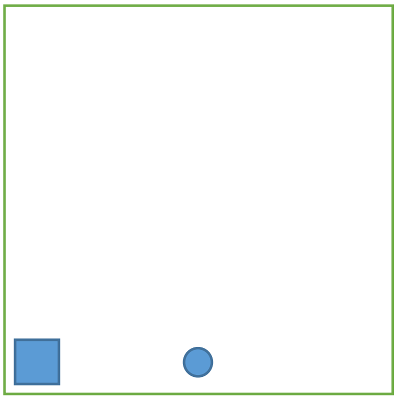
Scenario 8



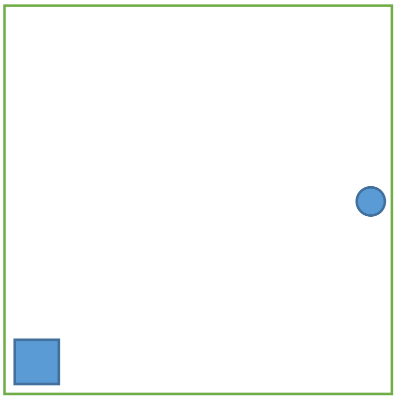
3.2.2 Scenarios Set 2

In the scenario set two wireless access point is stationary in the corner of the room and data gathering computers position is changed around the access point.

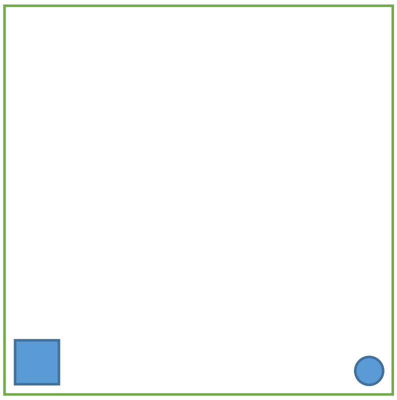
Scenario 9



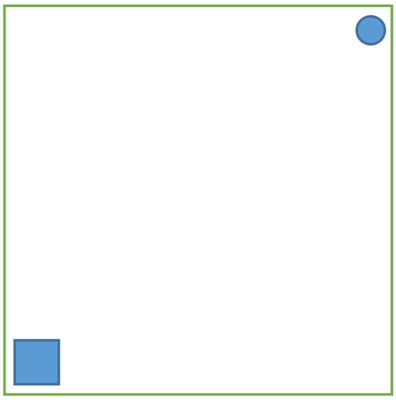
Scenario 11



Scenario 10



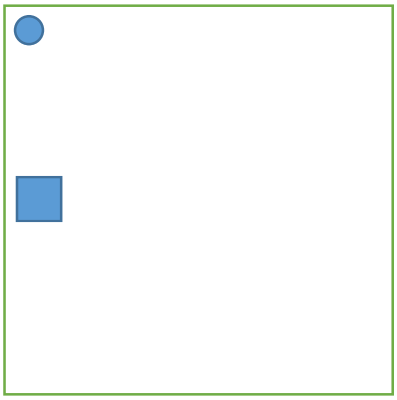
Scenario 12



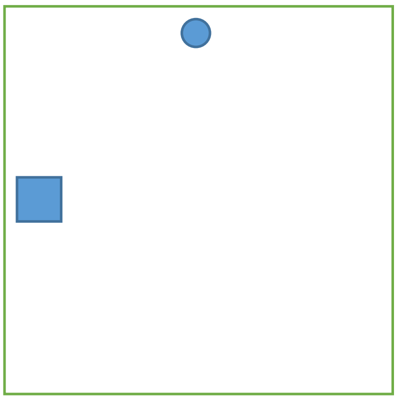
3.2.3 Scenarios Set 3

In the scenario set three wireless access point is stationary in the middle corner of the room and data gathering computers position is changed around the access point.

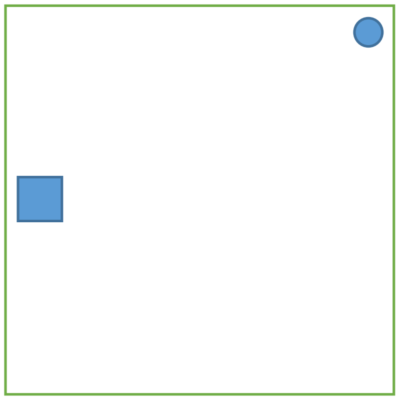
Scenario 13



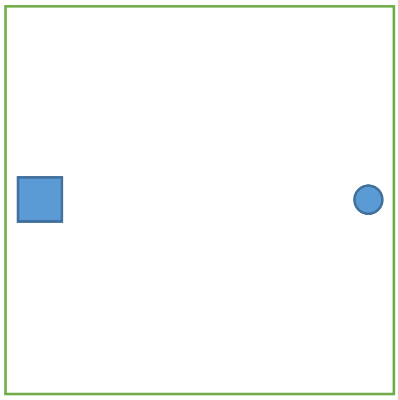
Scenario 14



Scenario 15



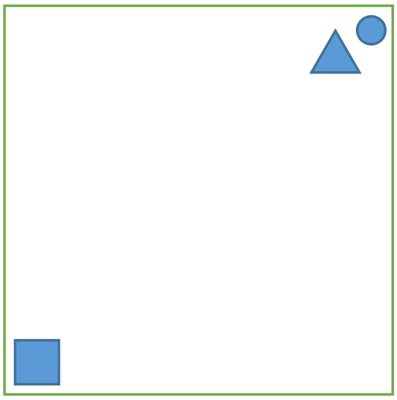
Scenario 16



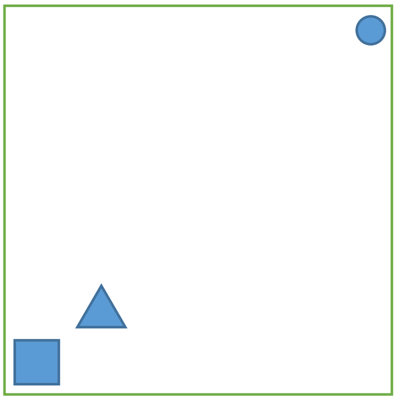
3.2.4 Scenarios Set 4

For the scenario set four we introduce a human as an obstructing object (human between the Wireless access point and Data gathering computer). From above scenarios I have randomly selected Scenario 12 for this test. Here the position of Wireless access point and Data gathering computer is constant and human position is changed in line-of-sight of the Wireless access point and Data gathering computer. Human subject will stay on three different positions when gathering the data. In scenario 12.1 human will be near the data gathering computer. In the 12.2 scenario human will be in the middle of the room same distance in the diagonal line from the data gathering computer and wireless access point. Finally in the scenario 12.3 human will be near the wireless access point obstructing the LOS of the wireless signal.

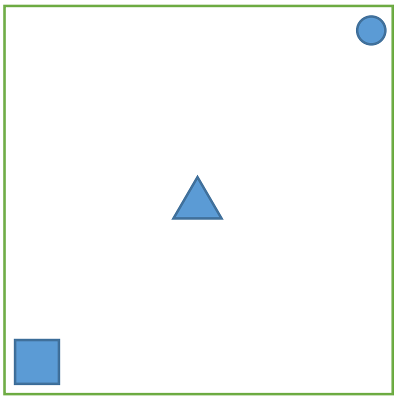
Scenario 12.1



Scenario 12.3



Scenario 12.2



3.3 Implementation

For this study setting up the designed scenario and gathering data is done in the implementation stage. Received Signal Strength is the value we are interested and it is gather as data. This is done using a python script (Please refer to appendix A) running on data gathering computer. Setup of the test bed environment is an ordinary household room with the dimensions length of 14 feet and a width 13 feet 10 inches, diagonal length of 19 feet and 8 inches. For the first three sets of scenarios data is gathered with line of sight by a wireless access point and the data gathering computer. These sets will act as control scenarios in the analysis stage to show the distribution when there is not human present in the room.

Figure 14 Data gathering for first three scenario sets

For the fourth set of scenario we place a human subject between the wireless access point and the data gathering computer with line of sight to both devices hence obstructing the direct line of sight of both devices.

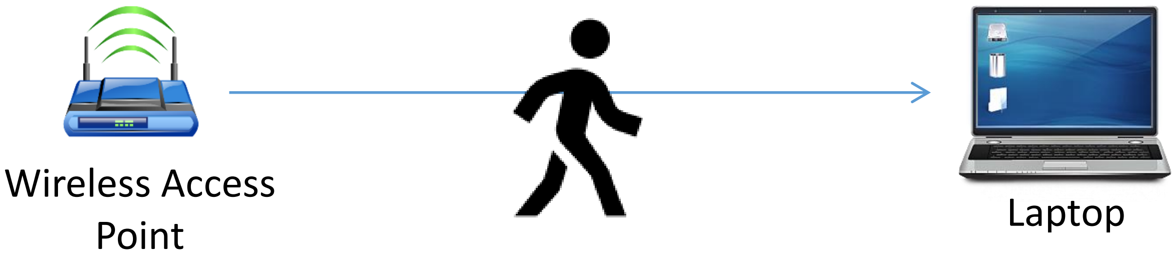


Figure 15 Data Gathering with human subject obstructing the line of sight path

Below figures show the arrangement inside the room with and without a human. Notice in the latter figure human is completely obstructing the line of sight of the wireless access point to the data gathering computer

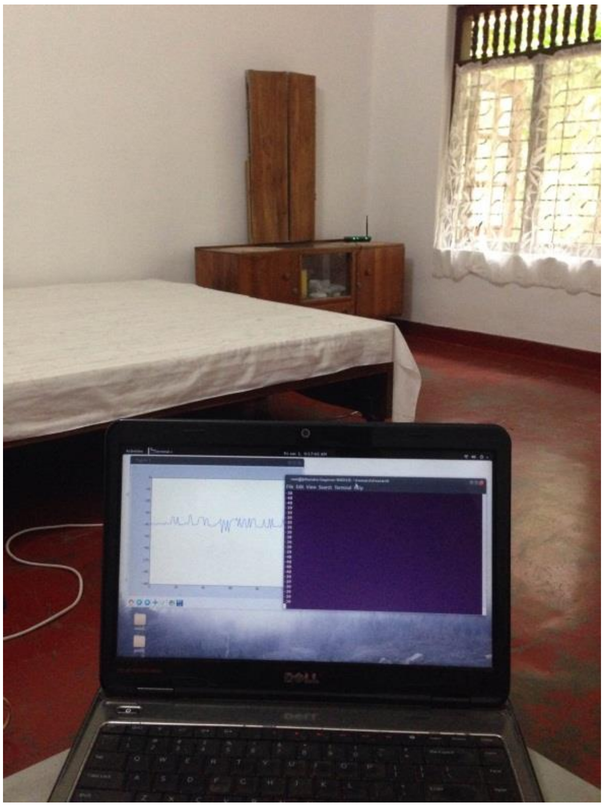


Figure 16- Data gathering without obstacles

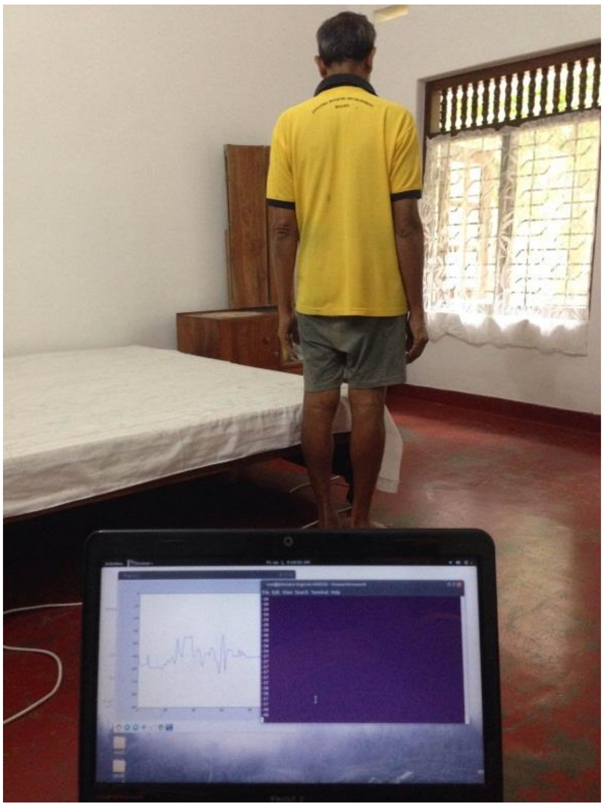


Figure 17- Data gathering with human obstructing the line of sight

As mentioned earlier data is gathering is done for a time period. For this study we are gathering the RSSI for 15 minutes for each scenario. Then these data is saved to a text file for future analysis. Python script and Linux I/O redirection is used for data gathering by taking the value from wireless card of the data gathering computer and saves the value roughly about 6 times a second. Root permission or super user privileges are needed by the script in order to extract the required data from the wireless card. That means RSSI value is taken every 0.166 MS or the sampling rate of 6 for my research.

Running the script as an example for the scenario 1

```
#python wifi_plot.py >> sc1.txt
```

Below screenshots shows the running stage of the script while gathering data and visualizing data in real-time.

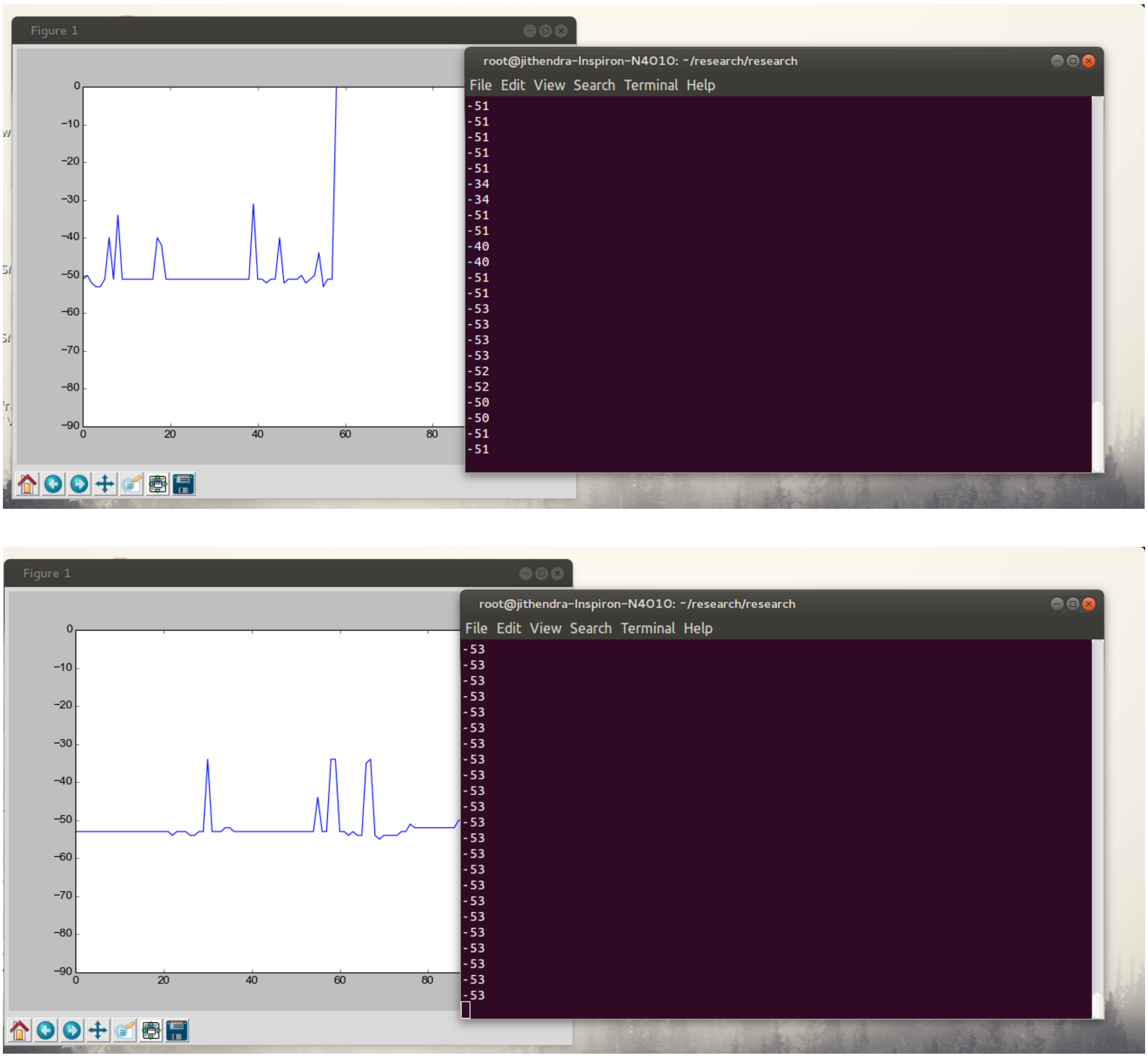


Figure 18- Running stage of the data gathering script

Chapter 4

4 Analysis and Evaluation

In the analysis stage data will be processed to derive a meaningful output from the raw data gathered in the previous stage.

4.1 Analyzing RSSI data

I’m going to use a statistical approach to analyze and evaluate the gathered data in hoping to find answers to the first research question. (Is there a significant amount of wireless signal strength drop when there is a human inside the line of sight of the wireless access point and data gathering computer?).

First all the data is plotted to normal distribution graphs of the RSSI values. As mentioned earlier RSSI is value is not a Constance every time it keeps on changing throughout the data gathering period. So to analyze we need to have an overview of how RSSI values are being distributed with respect to each scenario. Below are the normal distribution graphs with respect to their scenario. Axis of the graph will carry below information

- X axis = Signal Strength in dB
- Y axis = Density of probability (the chance of obtaining values near corresponding points on the X-axis)

4.1.1 Normal Distribution graphs for scenario set 1

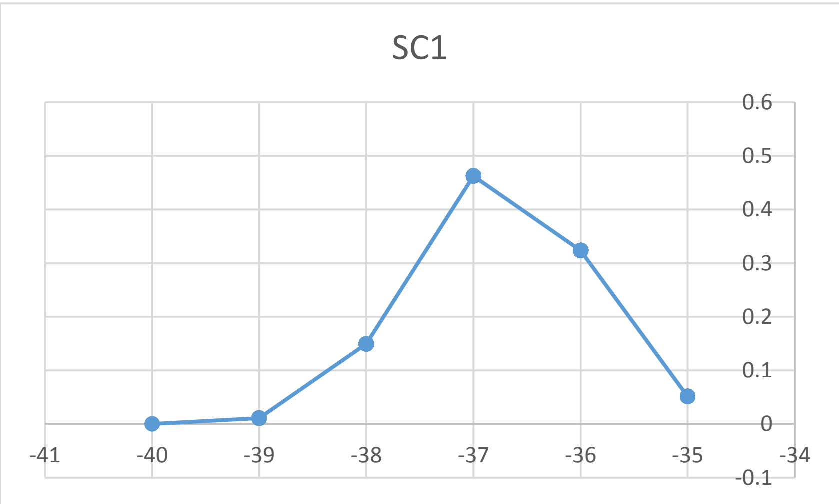


Figure 19 - RSSI value distribution graph for scenario 1

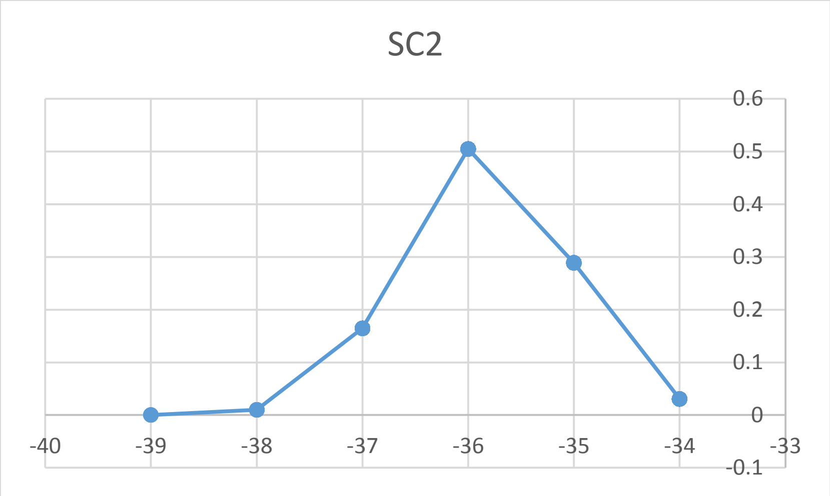


Figure 20 - RSSI value distribution graph for scenario 2

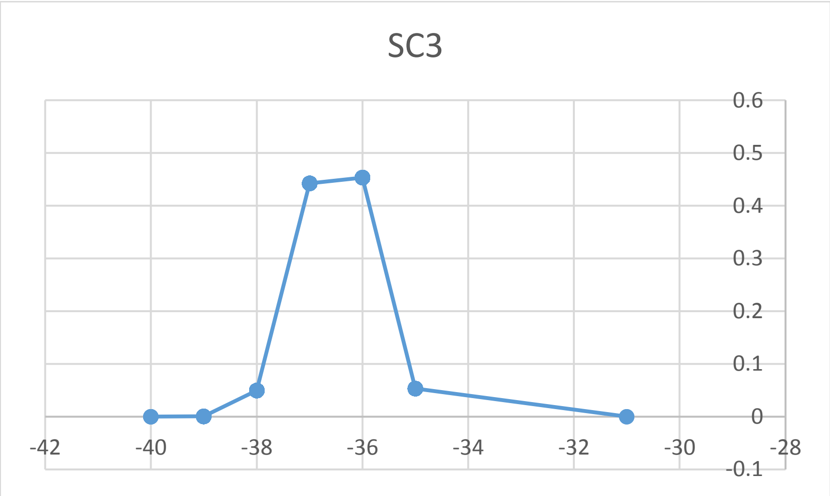


Figure 21 - RSSI value distribution graph for scenario 3

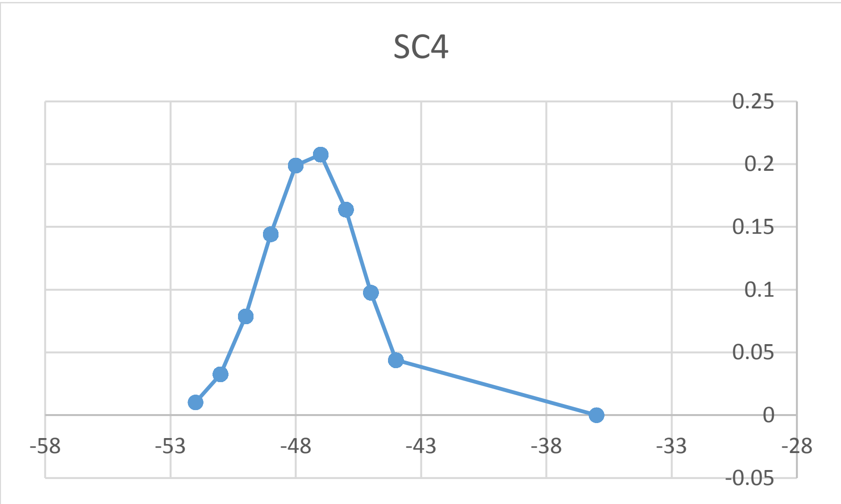


Figure 22 - RSSI value distribution graph for scenario 4

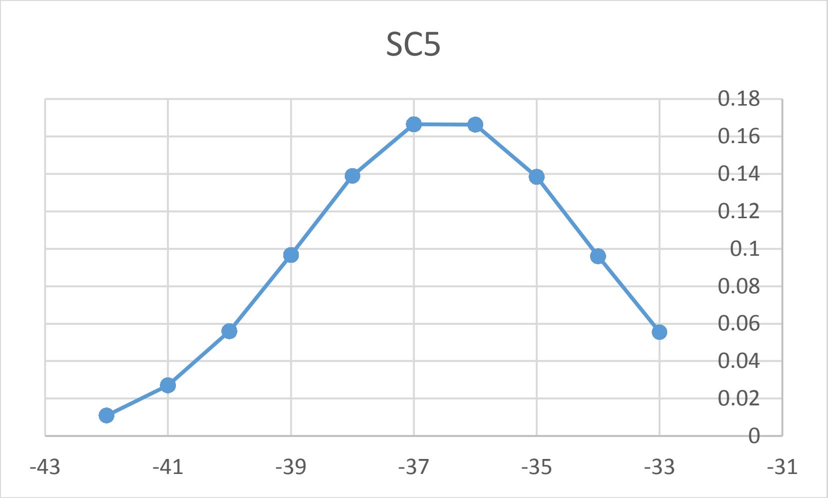


Figure 23 - RSSI value distribution graph for scenario 5

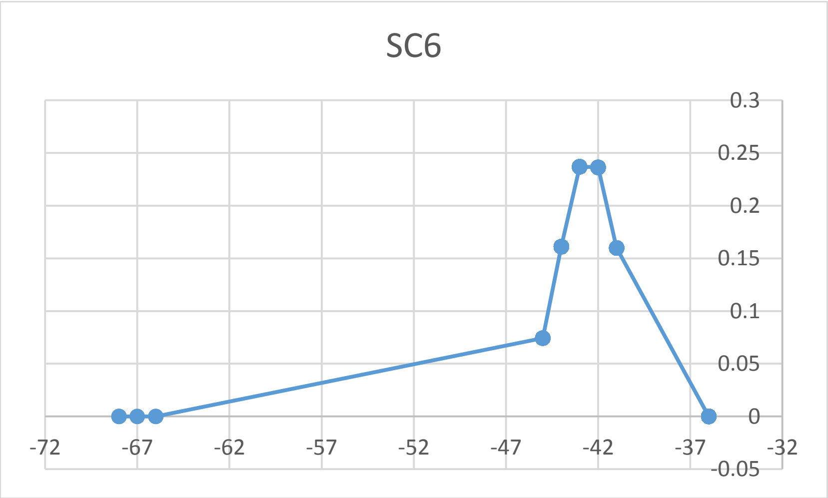


Figure 24 - RSSI value distribution graph for scenario 6

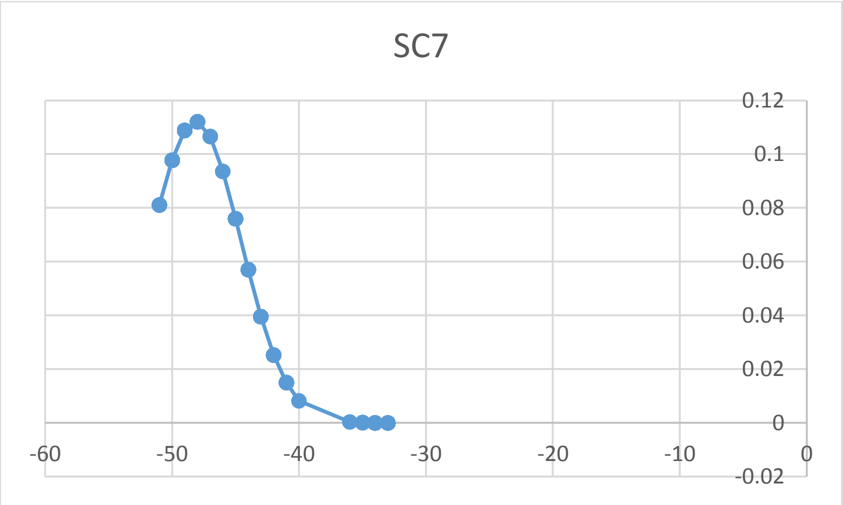


Figure 25 - RSSI value distribution graph for scenario 7

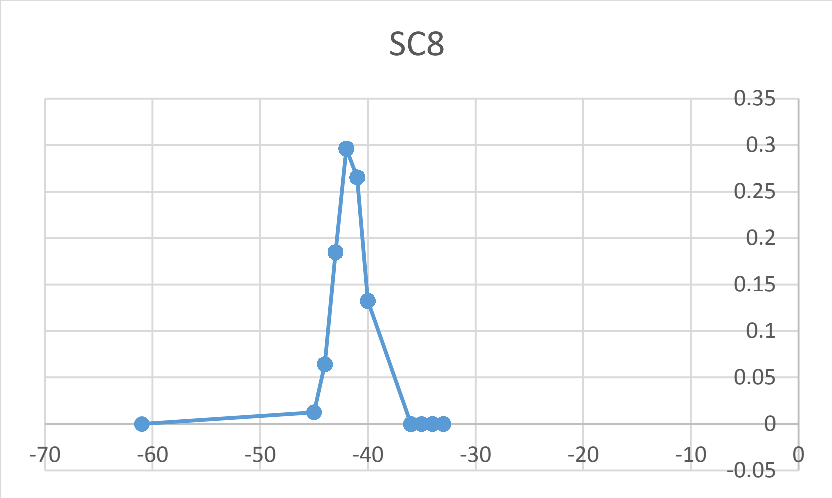


Figure 26 - RSSI value distribution graph for scenario 8

4.1.2 Normal Distribution graphs for scenario set 2

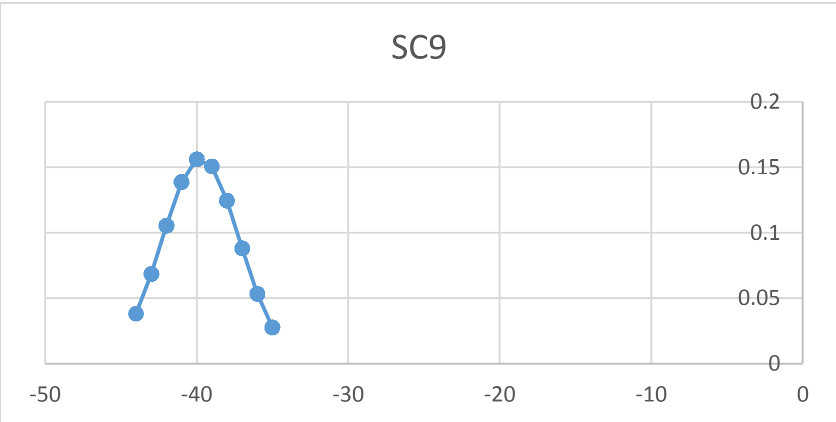


Figure 27 - RSSI value distribution graph for scenario 9

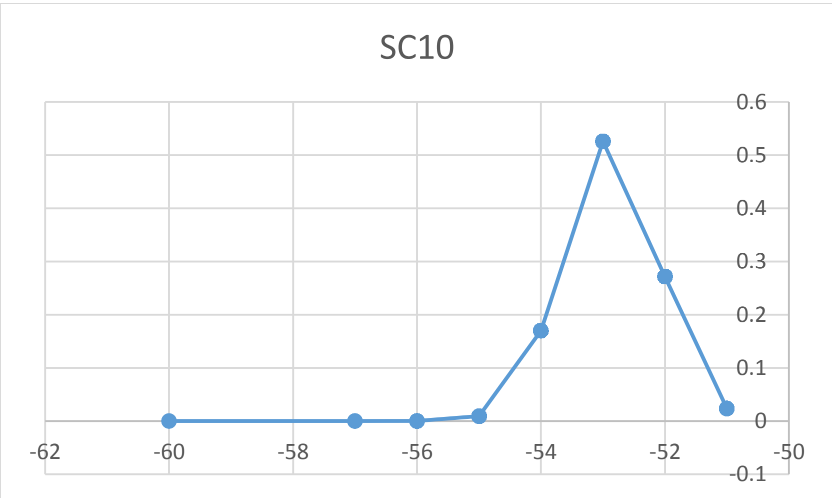


Figure 28 - RSSI value distribution graph for scenario 10

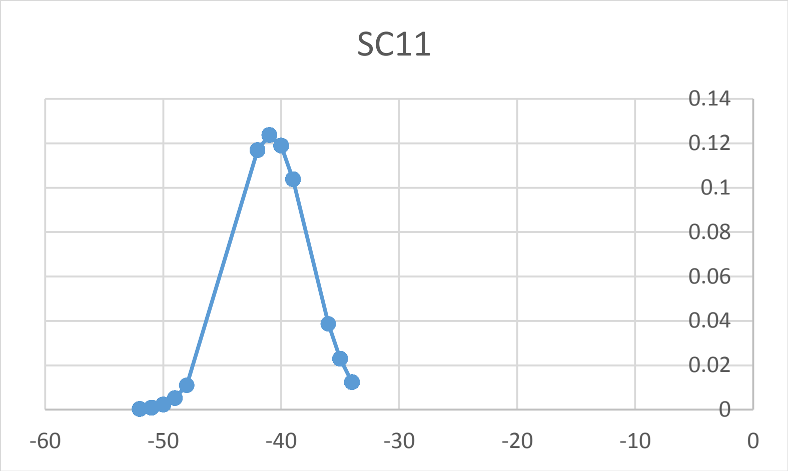


Figure 29 - RSSI value distribution graph for scenario 11

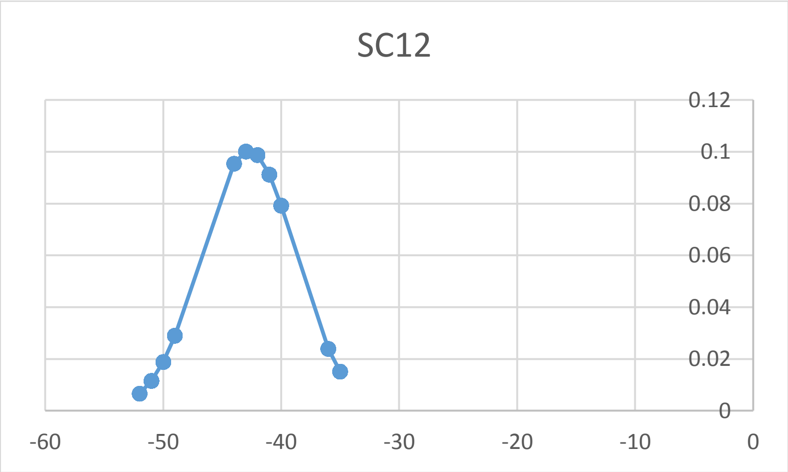


Figure 30 - RSSI value distribution graph for scenario 12

4.1.3 Normal Distribution graphs for scenario set 3

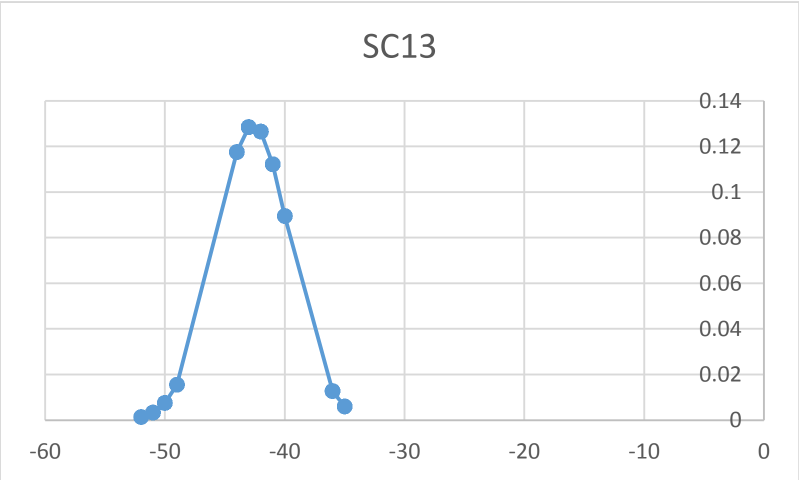


Figure 31 - RSSI value distribution graph for scenario 13

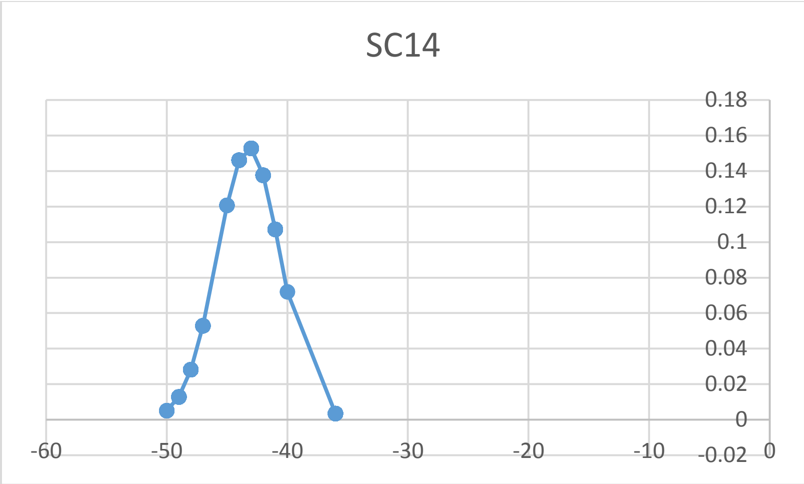


Figure 32 - RSSI value distribution graph for scenario 14

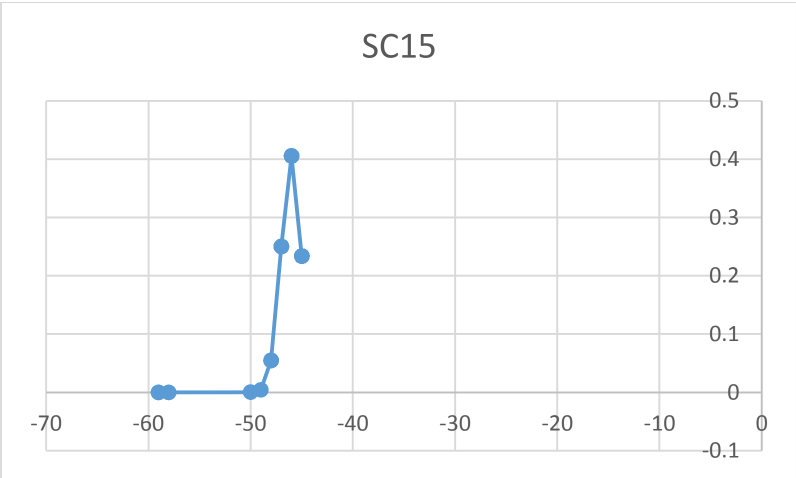


Figure 33 - RSSI value distribution graph for scenario 15

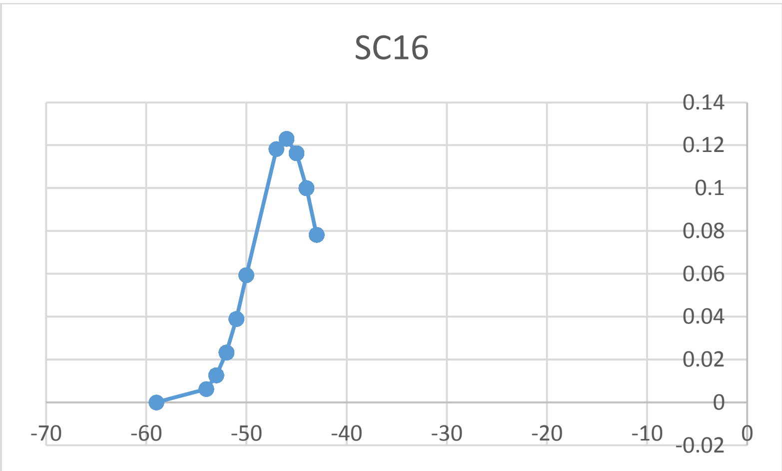


Figure 34 - RSSI value distribution graph for scenario 16

4.1.4 Normal Distribution graphs for scenario set 4

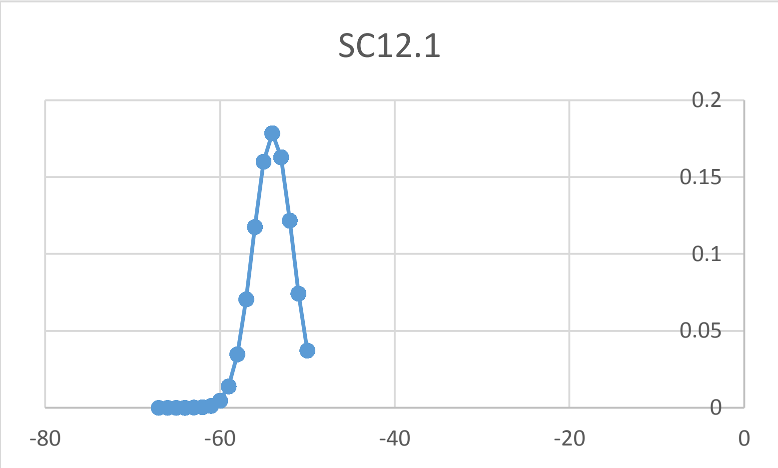


Figure 35 - RSSI value distribution graph for scenario 12.1

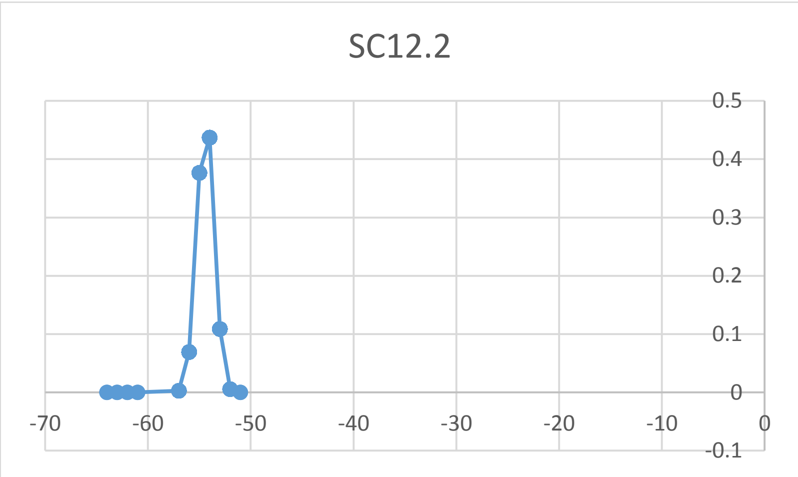


Figure 36 - RSSI value distribution graph for scenario 12.2

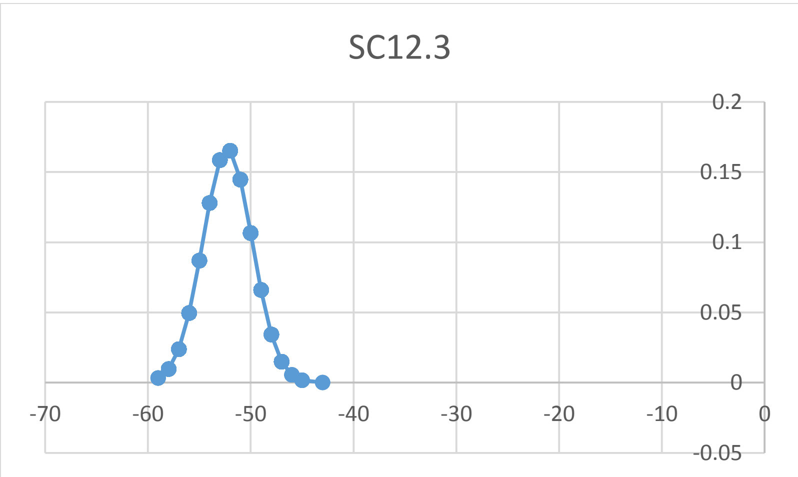


Figure 37 - RSSI value distribution graph for scenario 12.3

At the same time average value of the RSSI data and standard deviation value is calculated. Below is the summarized table of average value and standard deviation value of the RSSI data with respect to each scenario. Standard deviation is calculated using below formula.

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Where,

x = RSSI values

$\bar{x}$ = Mean value

n = Total number of values.

Scenario	Average(mean) of RSSI values	Standard deviation value
scenario 1	-36.72923077	0.816114501
scenario 2	-35.82	0.772871436
scenario 3	-36.48846154	0.680182433
scenario 4	-47.34346154	1.885783548
scenario 5	-36.54230769	2.360910279
scenario 6	-42.50269231	1.609141652
scenario 7	-48.21019231	3.515832894
scenario 8	-41.68923077	1.303803373
scenario 9	-39.72076923	2.540934613
scenario 10	-52.87115385	0.745786791
scenario 11	-40.97269231	3.305387441
scenario 12	-42.78019231	3.994749386
scenario 12.1	-53.95653846	2.232675289

scenario 12.2	-54.40423077	0.804429328
scenario 12.3	-52.2975	2.401623461
scenario 13	-42.64692308	3.083992847
scenario 14	-43.20269231	2.604845615
scenario 15	-46.03230769	0.983668557
scenario 16	-46.08615385	3.241317076

Table 3 - Average and Standard deviation values of RSSI data in all scenarios

4.2 Evaluation

In the evaluation stage we are focusing on the answer the research questions using the analyzed data. For the first research question we are using a statistical approach. I'm going to use the statistical hypothesis testing methodologies.

4.2.1 Statistical Hypothesis Testing

Statistical hypothesis testing is an assumption about a population parameter. This assumption may or may not be true. Hypothesis testing refers to the formal procedures used by statisticians to accept or reject statistical hypotheses.

First step of this methodology is to build a hypothesis by defining the null hypothesis and the alternative hypothesis. With respect to this study I have built below hypothesis and with the distribution graphs we can do the two-tailed test.

- Null hypothesis – H0 = there is no change to amount of wireless signal strength when there is a human and no human inside the line of sight of the wireless access point and data gathering computer
- Alternative hypothesis – H1 = there is a significant amount of wireless signal strength drop when there is a human inside the line of sight of the wireless access point and data gathering computer

Hypothesis testing uses mean value (μ) of the distribution to determine if we need to ignore the null hypothesis and go to test the alternative hypothesis. If the mean values of two scenarios are defined by $\mu 1$ and $\mu 2$. Then for the null hypothesis to become true $\mu 1$ should be equal to μ2. If the null hypothesis becomes true we can say there is no significant difference between the scenarios. But if μ1 > μ2 or μ1 < μ2 then we can test the alternative hypothesis to determine whether there is a significant difference between the scenarios. To test this hypothesis we have randomly selected scenario 12 and derived scenario 12.1, 12.2 and 12.3.

Secondly we are to specify a significance level. The significance level is the highest value of a probability value for which the null hypothesis is rejected. Common significance levels are 0.05 and 0.01. If the 0.05 level is used, then the null hypothesis is rejected if the probability value is less than or equal to 0.05. Here we are taking 0.05 as the significance level.

Third step is to calculate the values. Here we are calculating the t-test values using the below formula.

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{{S_1}^2}{N_1} + \frac{{S_2}^2}{N_2}}}$$

Where,

$\bar{x}_1$ is the mean of first data set

$\bar{x}_2$ is the mean of first data set

S_1 is the standard deviation of first data set

S_2 is the standard deviation of first data set

N_1 is the number of elements in the first data set

N_2 is the number of elements in the first data set

Below are the results of the calculations.

Scenario	Mean (dB)	T-Test value
Scenario 12	-42.78019231	-
Scenario 12.1	-53.95653846	0
Scenario 12.2	-54.40423077	0
Scenario 12.3	-52.2975	0

Table 4 - Calculated values for hypothesis testing

From above results we can discard the Null hypothesis because all the μ values are different and the value of the T-test is 0 for all three scenarios. This is below the 0.05 standard, so the result is statistically significant for all scenarios.

This proves the alternative hypothesis is true which in turn mean that there is a significant amount of wireless signal strength drop when there is a human inside the line of sight of the wireless access point and data gathering computer and answering the first research question with respect to selected scenarios.

4.2.2 Normal Probability Density Function

To answer the second research question (How accurately we can identify human presence by observing received signal strength indicator (RSSI) values) we are using a probability based approach. Here we are making use of the normal probability density function which is defined using below formula

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where

x = value we are looking for probability

μ = mean value

σ = standard deviation.

When normal probability density function is applied for a population and for an observed value (denoted by X) from the population, this function returns the probability of values less than the observed value. This can also be taken as the “percentile” for the observed value. Below picture shows if the observed value x is 35 the probability is calculated by the values below 35 as shaded.

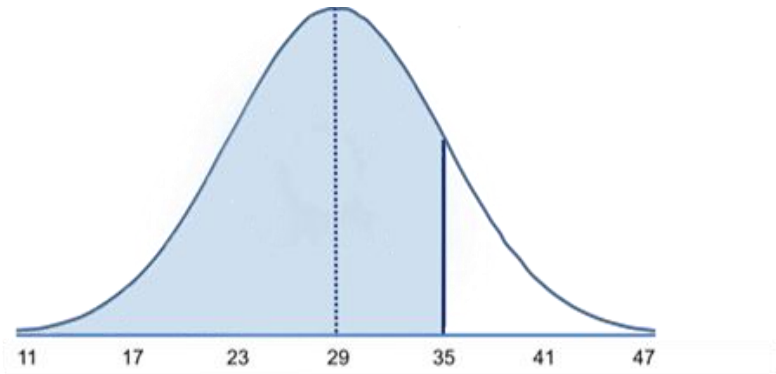


Figure 38 - Normal probability density function applied for a distribution

For this test also we are taking scenario 12 and its sub scenarios, scenario 12.1, 12.2 and 12.3. First we calculate the RSSI values and their occurrence frequencies from the gathered data for above scenarios.

RSSI Value	Occurrence Frequency
-52	20
-51	416
-50	484
-49	74
-44	10
-43	32
-42	1831
-41	1629
-40	430
-36	78
-35	196

Table 5 - RSSI value frequencies of scenario 12

RSSI Value	Occurrence Frequency
-67	6
-66	6
-65	6
-64	20
-63	12
-62	20
-61	40
-60	88
-59	88
-58	94
-57	114
-56	256
-55	568
-54	1350
-53	1476
-52	728
-51	288
-50	40

Table 6 - RSSI value frequencies of scenario 12.1

RSSI Value	Occurrence Frequency
-64	2
-63	2
-62	2
-61	2
-57	38
-56	262
-55	1796
-54	2714
-53	366
-52	14
-51	2

Table 7- RSSI value frequencies of scenario 12.2

RSSI Value	Occurrence Frequency
-59	4
-58	48
-57	119
-56	332
-55	406
-54	794
-52	660
-51	826

-50	530
-49	294
-48	249
-47	96
-46	4
-45	6
-43	2

Table 8- RSSI value frequencies of scenario 12.3

Now we take the RSSI value with highest frequency from these scenarios and calculate the occurrence probability using the normal probability density function.

Scenario	RSSI value with highest frequency (dBm)	Occurrence probability
Scenario 12	-42	0.577422691
Scenario 12.1	-53	0.665829879
Scenario 12.2	-54	0.692344273
Scenario 12.3	-51	0.705490974

Table 9 - Cumulative probability of the highest occurring RSSI value

These results show that the occurrence probability and the highest frequency value differ when the human is moved in line-of-sight of the wireless access point and the data gathering computer. With these occurrence probabilities we can get an idea about the location of the human subject. Probabilities are not as high as other researches due to not using specialized high end equipment's but with more than 50% confidence level we can derive the position by analyzing data. Hence we are able to localizing the obstructing object, answering the second research question. Because of this it proves that by using radio tomographic methodologies it is possible to violate privacy of an individual even without him knowing that he is being tracked.

Chapter 5

5 Conclusion

Conclusion will be the last chapter of this dissertation. It will discuss about a summary of the research and its results. Also this contains possible future works that can be done using this research.

5.1 Contributions

This study opened a new area in the field of radio tomography which is privacy and security implications of radio tomography. This dissertation contains design, analysis and evaluation of human identification with changes in the RSSI values of the Wi-Fi signal by using radio tomographic methodologies which in turn can be used to breach privacy of a person. Significant of this research is privacy can be breached without the knowledge of the person who is being watched. Other importance is that this exploitation can be done using simple devices without the need of highly expensive or specialized devices.

At the beginning of the study there were two research questions; is there a significant drop in the RSSI levels when there is a human present obstructing the wireless access point and the data gathering computer(in line of sight). The next questing if there is a human presence with how much accuracy we can tell his location. In hoping to answer these two questions I have designed, implemented and evaluate the gathered data.

Gathered data is a distribution of data, it is not a simple value but a collection of values over the period of data gathering time. To answer the first research question I have used the statistical hypothesis testing. Statistical hypothesis testing uses mean and standard deviation to prove or disprove the concepts. In hypothesis testing there are two types of hypothesis; null hypothesis and alternative hypothesis. When applied to this study null hypothesis defines there is no significant change in RSSI when there is a human obstructing the wireless signal and alternative hypothesis defines that there is significant change in RSSI when there is a human obstructing the wireless signal. When the mean and standard deviation are calculated we get following values for scenarios 12, 12.1, 12.2 and 12.3 [mean, standard deviation] [-42.78019231 dBm, 3.994749386] , [-53.95653846 dBm, 2.232675289] , [-54.40423077 dBm, 0.804429328], [-52.2975 dBm. 2.401623461]. As you can see from these results mean is different for all the scenarios. Then we calculate the T-Test value with a pre-defined significant level of 0.05. T-Test value obtained is 0 for all the scenarios. Because of that it is proved that there is indeed a significant level of RSSI level drop when there is a human obstructing the line of sight and it can be easily identified.

The second part of the project is to localize the human obstacle to identify his whereabouts. To do this I have taken a probability based approach. By using the normal probability density function to calculate the occurrence probability of the highest occurrence of RSSI values to show that occurrence probability of a RSSI value is unique to a specific scenario. By applying this techniques for scenario 12 the highest occurrence RSSI value is -42dBm with an occurrence probability of 0.577422691, with respect to scenario 12.1 values are -53dBm and 0.665829879, for scenario 12.2 -54dBm and 0.692344273, for scenario 12.3 -51dBm and 0.705490974. By looking at these results gained we can determine a person's location with more than 50% probability. This is the answer for the second research question. With this answer we can prove that it is possible to track a human within an indoor environment using ordinary hardware by analyzing the RSSI values of a wireless signal. In turn we can also state that someone's privacy can be breached by using above proposed methodology.

5.2 Future Work

The study can be extended for future research in the area of privacy and security aspects of radio tomography. Below are some areas of interest that one researcher can do by extending this project.

1. Expansion of the researching environment

   Present day Wi-Fi networks are used in both indoor and outdoor environments to provide wireless connectivity. Here we have considered only an indoor environment. As a future expansion one can study on identification of the human presence in an outdoor environment as well as in hybrid situations by converging both indoor and outdoor environments.

2. Error correction for environmental factors

   Jose Otero, Pavana Yalamanchili and Hans-Werner Braun have done research on the effects of weather on the wireless signals [19]. For this study we have not consider the weather factors like wind, temperature and humidity. All these will affect the signal strength of the wireless signal. Results of this research may improve by factor in the errors introduced by the weather to the gathered data and doing an error correction on the analysis and evaluation stage.

3. Errors due to radio noise levels

   There are wireless radio noises in the environment around us. This value is also not a constant it keeps on changing due to many reasons. In high density wireless areas the noise may be high. For this research we have not taken this factor into account. As a future project one can analyze the effects of wireless noise on the human tractability shown in this project.

4. Human fingerprinting to uniquely identify a person

   This study only shows that we can see a significant drop in the RSSI level when there is a human obstructing the wireless signal and we can predict his location with a reasonable accuracy by using probability. Different people might give different levels of RSSI level drops. By using this one can do research on possibility of fingerprinting different people using signal level drops that are unique to individuals. If this is a success by analyzing the RSSI levels we can uniquely identify a person.

5. Real-time framework to identify human presence

   This study uses offline methods for analyzing and detecting human presence using RSSI levels. This means we gather data and analyze it later and obtain results. As improvement to this project one can develop a framework to analyze data in real time to get the outputs.

6. Auditing framework to identify privacy breach points in a typical environment

   Radio tomography is a relatively new and immerging area. Researchers have shown that by using radio tomography it is possible to track single human, groups of human or human gestures. But privacy and security aspects are not being taken into account. In information technology there are many auditing frameworks available to secure individuals. But unfortunately there is not auditing framework to identify privacy breaches happen using radio tomography. Wireless networks are exploding everywhere on the planet so need of such a framework in absolutely necessary.

References

[1] University of Utah. (2015 October, 01). *Radio Tomographic Imaging* [Online]. Available: http://span.ece.utah.edu/radio-tomographic-imaging [Accessed 3 Oct. 2015].

[2] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," IEEE Transactions on Mobile Computing, vol. 9, no. 5, pp. 621–632, May 2010. Published, 01/08/2010.

[3] Michael McFarland, SJ. *What is Privacy?* [online] Scu.edu. Available at: http://www.scu.edu/ethics/practicing/focusareas/technology/internet/privacy/what-is-privacy.html [Accessed 3 Oct. 2015].

[4] Samuel D. Warren and Louis D. Brandeis, "The Right to Privacy," Harvard Law Review, 4 (5), (1890): 193-220, p. 195, citing Judge Cooley in Cooley on Torts, 2nd Ed.

[5] Wikipedia, (2015). *Wi-Fi*. [online] Available at: https://en.wikipedia.org/wiki/Wi-Fi [Accessed 5 Oct. 2015].

[6] Wikipedia, (2015). *IEEE 802.11*. [online] Available at: https://en.wikipedia.org/wiki/IEEE_802.11 [Accessed 5 Oct. 2015].

[7] Wikipedia, (2015). *Decibel*. [online] Available at: https://en.wikipedia.org/wiki/Decibel [Accessed 5 Oct. 2015].

[8] Wikipedia, (2015). *Wireless access point*. [online] Available at: https://en.wikipedia.org/wiki/Wireless_access_point [Accessed 5 Oct. 2015].

[9] Dlink.co.in, (2015). *D-Link Wireless N ADSL2+ 4-Port Wi-Fi Router*. [online] Available at: http://www.dlink.co.in/products/?pid=452 [Accessed 5 Oct. 2015].

[10] Wikipedia, (2015). *Received signal strength indication*. [online] Available at: https://en.wikipedia.org/wiki/Received_signal_strength_indication [Accessed 5 Oct. 2015].

[11] En.community.dell.com, (2015). *DW1501 Wireless-N Wlan Half-Mini Card Specifications*. [online] Available at: http://en.community.dell.com/cfs-file/__key/communityserver-discussions-components-files/3324/4338.1501.1.jpg [Accessed 6 Oct. 2015].

[12] Goldsmith, A. (2005). *Wireless communications*. Cambridge: Cambridge University Press.

[13] Bahl, P. and Padmanabhan, V. (2000). *RADAR: An In-Building RF-based User Location and Tracking System*. [online] Available at: http://research.microsoft.com/en-us/people/padmanab/infocom2000.pdf [Accessed 12 Oct. 2015].

[14] Joey Wilson and Neal Patwari. A fade-level skew-laplace signal strength model for device-free localization with wireless networks. *IEEE Transactions on Mobile Computing*, 11(6):947-958, June 2012.

[15] Moustafa Seifeldin and Moustafa Youssef. A deterministic large-scale device-free passive localization system for wireless environments. 2010.

[16] Yang Zhao, Neal Patwari, Jeff M. Phillips, and Suresh Venkatasubramanian. Radio tomographic imaging and tracking of stationary and moving people via kernel distance. In Proceedings of the *12th International Conference on Information Processing in Sensor Networks*, IPSN '13, pages 229-240, New York, NY, USA, 2013. ACM.

[17] Fadel Adib and Dina Katabi. See through walls with wi-fi! *SIGCOMM Comput. Commun*. Rev., 43(4):75-86, August 2013.

[18] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C. Miller. 3d tracking via body radio reflections. In *Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation*, NSDI'14, pages 317-329, Berkeley, CA, USA, 2014. USENIX Association.

[19] Otero, J., Yalamanchili, P. and Braun, H. (2010). *High Performance Wireless Networking and Weather*. 1st ed. [ebook] Available at: http://hpwren.ucsd.edu/info/images/weather.pdf [Accessed 10 Dec. 2015].

[20] Pu, Q., Gupta, S., Patel, S. and Gollakota, S. (2013). Whole-Home Gesture Recognition Using Wireless Signals. In: *The 19th Annual International Conference on Mobile Computing and Networking (Mobicom'13)*. [online] Available at: http://wisee.cs.washington.edu/wisee_paper.pdf [Accessed 27 Dec. 2015].

Appendices

Appendix A

Data Gathering Code

Data gathering code is developed by Mr. Asanka P. Sayakkara of UCSC. (Source: http://recolog.blogspot.com/2015/03/plotting-wifi-signal-strength-variation.html)

```
import matplotlib.pyplot as plt
import time
import random
from collections import deque
import numpy as np
import os


def read_wifi():
	while True:
		f=os.popen('sudo iwconfig wlan0 | grep -e "Signal level"')
		line = f.read()
		splitted_line = line.split()
		level = splitted_line[3].split('=')
		print level[1]
		val = level[1]
		yield val
		time.sleep(0.1)


a1 = deque([0]*100)
ax = plt.axes(xlim=(0, 100), ylim=(0, 10))
d = read_wifi()
line, = plt.plot(a1)
plt.ion()
plt.ylim([-90,0])
```

```
plt.show()

for i in range(0,10000):
	a1.appendleft(next(d))
	datatoplot = a1.pop()
	line.set_ydata(a1)
	plt.draw()
	print a1[0]
	i += 1
	time.sleep(0.1)
	plt.pause(0.0001)
```

Appendix B

Gathered Data

All gathered data is available at github repository which can be accesses using this URL. https://github.com/jithendra89/research/