Node Localization using Fractal Signal Preprocessing and Artificial Neural Network

By

Tashniba Kaiser

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg

Abstract

This thesis proposes an integrated artificial neural network based approach to classify the position of a wireless device in an indoor protected area. Our experiments are conducted in two different types of interference affected indoor locations. We found that the environment greatly influences the received signal strength. We realized the need of incorporating a complexity measure of the Wi-Fi signal as additional information in our localization algorithm.

The inputs to the integrated artificial neural network were comprised of an integer dimension representation and a fractional dimension representation of the Wi-Fi signal. The integer dimension representation consisted of the raw signal strength, whereas the fractional dimension consisted of a variance fractal dimension of the Wi-Fi signal.

The results show that the proposed approach performed 8.7% better classification than the "one dimensional input" ANN approach, achieving an 86% correct classification rate. The conventional Trilateration method achieved only a 47.97% correct classification rate.

Acknowledgements

First and foremost, I would like to express my heartiest gratitude to my advisor and mentor, Dr. Ken Ferens for his continuous guidance, incredible support, invaluable time, flawless encouragement, and research grants for the completion of my graduate studies.

I sincerely acknowledge and appreciate Dr. W. Kinsner for his wonderful teaching, guidance and support in the development of this thesis. It has been a great privilege for me to take his course.

I would also like to thank the University of Manitoba and Seccuris Inc. for providing me funds to continue my higher education in Canada.

I would like to especially thank all the course instructors for their guidance and support. Special thanks to Amy Dario, the graduate student advisor of Electrical and Computer Engineering department, and administrative staffs of the Faculty of Graduate Studies for their help throughout these days.

I am also grateful to my friends and team in 'The Embedded System Laboratory' for the assistance and support in some experimental aspects of this work.

I am ever obliged to my Lord for giving me strength to accomplish my thesis. Last but not least, I am eternally grateful to my beloved parents and adorable sister for their unconditional love, sacrifices and support throughout my life.

I dedicate this work to my beloved parents and adorable sister who have sacrificed their entire life for my higher education.

Table of Contents

List of Acronyms

WSN	Wireless Sensor Network
AOA	Angle of Arrival
TOA	Time of Arrival
TDOA	Time Difference of Arrival
RSS	Received Signal Strength
AP	Access Point
BP	Back Propagation
BS	Base Station
BW	Bandwidth
ANN	Artificial Neural Network
FD	Fractal Dimension
WT	Wireless Transmitter
VFD	Variance Fractal Dimension
VFDT	Variance Fractal Dimension Trajectory
WLAN	Wireless Local Area Network
LAN	Local Area Network
Tx	Transmitter
Rx	Receiver
GPS	Global Positioning system
SA	Spectrum Analyzer
MLP	Multilayer Perceptron
FP	Fingerprint
MS	Mobile Station
BTS	Base Transceiver Station
RADAR	Radio Detection and Ranging
SONAR	Sound Navigation and Ranging
IrDA	Infrared Data Association
RFID	Radio Frequency Identification
GSM	Global system for mobile communications
FCC	Federal Communication Commission
IR	Infrared
CGA	Conjugate Gradient Algorithm
Rprop	Resilient BP
LM	Levenburg-Marquardt
SCGA	Scaled Conjugate Gradient Algorithm
GD	Gradient Decent

List of Figures

List of Tables

Chapter 1

1. Introduction and Problem Definition

Wireless sensor network node localization is a continuously expanding research field with innovative technologies being developed. Localization, tracking, and positioning of an object and people in an indoor or outdoor environment have received ample importance and security concerns in wireless sensor network studies. The Federal Communication Commission (FCC) mandated in 1996 that US cellular operators must have the means by October 2001 to accurately locate mobile phones when they call the emergency number 911 [1], [2]. Many interesting positioning dependent applications are being developed, e.g., environmental monitoring, target tracking in military applications, smart building failure detection, navigation, location based billing, vehicle fleet management, intelligent transport systems, location directory services for health system, and geographical routing. However, there remain problems of localization accuracy.

The next generation wireless communication systems aim to improve localization especially for indoor environments. The goal of this chapter is to outline the foundation and motivation for the research on indoor WSN localization techniques.

1.1 Wi-Fi Network Topology

There has been tremendous growth in wireless network technology in recent years. Wireless sensor networks are mostly infrastructure dependent and are a kind of distributed network. The active positioning technique for wireless local area networks involves a variety of devices and uplink connections with LAN and the internet to facilitate users (Fig.

1). Two anchors in 802.11 wireless standards can be connected with each other over the wireless network. The wireless devices are associated with the respected anchors within their wireless range. In a distributed system, wireless devices can connect with local area network and the internet.

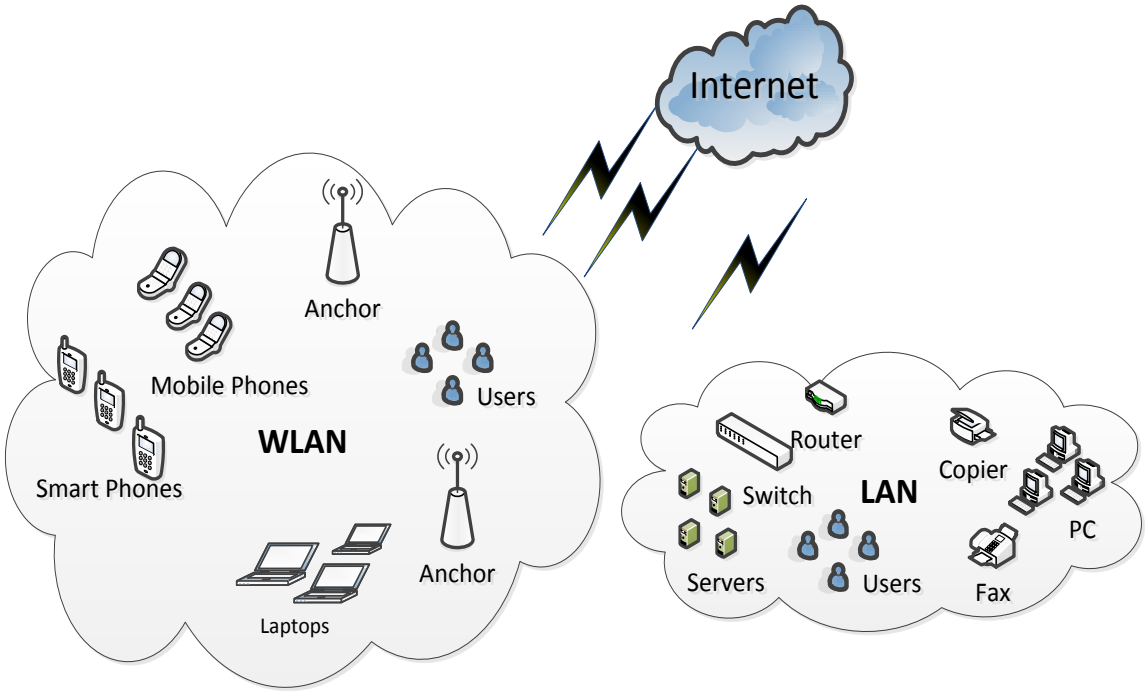


Fig. 1. An example of a wireless network topology in an enterprise.

1.1.1 Localization and Positioning

Localization of a wireless device involves extracting information from the environment. Localization of a wireless device defines the relative spatial place representation of a wireless device. Positioning defines the spatial co-ordinate of a wireless device.

1.2 Thesis Motivation

This section describes the motivation for doing the thesis. First the problem of node localization is described. This is followed by an explanation of why conventional techniques are not effective. Finally, this leads to the motivation for undertaking the research.

1.2.1 Problem Definition for Outdoor and Indoor Positioning System

An indoor localization measurement can be contaminated with channel errors, which include time varying impairments and environmental impairments. The time varying impairments result from interference by other deterministic narrowband wireless technologies (e.g., Wi-Fi, Bluetooth, ZigBee, cordless phones, and microwave ovens) and stochastic broadband electromagnetic (EM) noise signals.

The environmental impairments, such as fading, shadowing, multipath, non-line-of-sight (NLOS) propagations, reflections, and attenuation, result from obstacles in the environment (buildings, steel structures, and electronic equipment). Simple time averaging of the TOA, AOA, and RSS measurements would mitigate the effects of the time varying impairments to the extent that these impairments are purely random, but these measurements would be ineffective for the non-random components. Other more complex forms of time averaging or signal processing techniques would be required to take into account the not-so-quite random properties exhibited by these types of impairments, such as signals which contain constant slope and exponential decay power spectrum density [3].

1.2.2 Problem with Conventional Localization Techniques

The global positioning system (GPS) is a widely used technique for outdoor localization which has a localization precision of about 10 feet. Outdoor positioning uses the GPS receiver as a source [4] that calculates its position by a technique called satellite ranging, which involves measuring the distance between the GPS satellites and the receiver it is tracking. The GPS provides location and time information in an open space as long as the GPS receiver has a LOS with four or more GPS satellites. Each satellite position is pre-defined, and the GPS receiver uses this information to locate its own position. However, GPS cannot generally be used for indoor localization. There are various other technologies that have been widely used for indoor localization and positioning, such as the global system for mobile communication (GSM), infrared beacons, Bluetooth, Ultrasound, radio frequency ID, and wireless local area network (WLAN).

Due to the non-line-of-sight problem, global positioning system is not suitable for indoor positioning. Moreover, Wi-Fi based localization methods may not be suitable for a hospital environment as they can interfere with radio frequency signal based medical equipment. In contrast Wi-Fi-based localization techniques are the best choice for urban areas [3].

However, conventional localization techniques are not effective for indoor localization. Path loss mode parameters are changeable for different indoor locations. This is the reason why RSS readings greatly vary with environment. A new and innovative approach needs to be introduced, which can determine the complexity of the environment, and thus, mitigate the effects of environmental impairments.

In this thesis, the focus is on the challenges involved with localization algorithms in wireless sensor networks and the development of an integrated artificial learning scheme. The attenuated RSS measurements can cause significant errors in the estimated indoor position. The resulting error introduces inaccuracy in distance estimation. The same signal strength was received from multiple different distance positions. We want to detect all the positions that suffer from environmental impairments. A novel and innovative transformation scheme is analyzed and used as an input to represent the unique features of RSS captured from respective positions. In this thesis we use an ANN to predict the positions and fractal dimension to address the distinctive complexity of the system.

1.3 Thesis Statement

This thesis proposes to use a multi-dimensional (integrated) representation of the Wi-Fi signal as input to an artificial neural network for learning and discovering features for node localization in wireless sensor network applications. In this thesis the meaning of the term "integrated" is that the inputs to the artificial neural network were comprised of an integer dimension representation and a fractional dimension representation of the Wi-Fi signal. The integer dimension representation consisted of the raw signal strength, whereas the fractional dimension consisted of a variance fractal dimension of the Wi-Fi signal. This research discovers long range dependency of the Wi-Fi signal and extracts unique localization features using the variance fractal dimension trajectory.

1.4 Main Themes of the Thesis

The main ideas of the thesis are outlined as follows:

Interference Analysis: We observed different types of electromagnetic signals in several different test bed locations on Campus. Based on our observations, we set the center frequency and integrated bandwidth for the span of the Wi-Fi signal. The signals outside of the Wi-Fi span were removed from consideration. There were other signals still present in the Wi-Fi span, such as neighboring APs occupying an overlapping channel, video transmitter, and microwave oven. In the chosen three test bed locations, we observed signal impairments due to environmental structures, such as people, tables, and chairs. Our test bed locations were chosen specifically to vary between low to high competing signal interference and degradation due to structural impairments. In this way, we were able to emulate a typical scenario for node localization in "real" wireless sensor networks.

RSS Measurements: We obtained the RSS measurements in three different locations. First, we captured the RSS readings in a low interference indoor location known as a chamber room. Second, we captured the RSS readings in two indoor locations where RSS readings undergo several deformations and interferences. Third, we captured RSS readings in an outdoor open space where no other EM signals exist. The effect of contaminated channel errors in different indoor locations was observed.

Multi-Scale Fractal Analysis of a Wi-Fi signal: The wide variability in the RSS measurements causes significant errors in the estimated position. This thesis investigates the stationarity property of the signal that originates from a wireless transmitter. In this thesis, we have shown that the VFDT can extract features of a RSS signal, which are very similar to the features of a time domain representation of the signal.

Model Uniqueness: We obtained the variance fractal dimension trajectory for Wi-Fi signal, which is a unique approach in WSN localization research. The IANN approach

consists of a combination of variance fractal dimension and raw received signal strength inputs to an artificial neural network to learn, extract, and identify the unique features of the Wi-Fi signal for node localization in wireless sensor network applications.

Dependency and Cost Effectiveness: Our proposed integrated artificial neural network based (IANN) localization framework has no dependency on special hardware. Appending hardware raises the dependency on hardware and overall system costs.

Accuracy: We captured the received signal strength using Labview PC software application and a spectrum analyzer. Our research incorporated the use of a spectrum analyzer to allow very high precision of the RSS readings. We assume an access point used for localization will also employ a spectrum analyzer to provide highly precise RSS values.

1.5 Outline of the Thesis

This thesis is divided into the following chapters:

Chapter 2 presents a literature review on the different state-of-the-art approaches for localization techniques in wireless sensor networks. A qualitative comparison among the different localization approaches is provided in this chapter. We also discuss the functionalities, the work flow, and pros and cons of different localization techniques. We focus on the need and choice for our proposed research technique.

Chapter 3 introduces the proposed integrated artificial neural network framework as a range-free localization technique in a wireless sensor network. We discuss the design and implementation procedure for the multi-scale fractal analysis of Wi-Fi signals.

Chapter 4 describes our experimental setup as well as our hardware and software requirements and configurations. We identify the interference of other signals and discussed the NLOS effect on localization.

Chapter 5 presents the system-level simulation results for the proposed integrated frame-work and provides guidelines for selecting the system design parameters to satisfy the actual indoor positioning requirements. We verified our system model before functioning with the real data set. We discussed the findings from the multi-scale fractal analysis of Wi-Fi signal.

Chapter 6 summarizes the study and draws the conclusions of the thesis. This chapter describes the findings from the multi-scale fractal analysis of Wi-Fi signals. We also explain the major features and benefits of our proposed integrated ANN approach for WSN node localization. We suggest possible future work that can possibly improve the performance of the system.

Chapter 2

2. Background on Node Localization for Indoor Wireless Sensor Networks

The localization function plays a vital role in services that require locating nodes in a wireless sensor network. Node location estimations are widely used in different applications, like-mobile phone tracking, location sensitive billing in telecommunication, intelligent transportation system, environmental monitoring, mapping, event tracking, tracking of fire fighters and miners, patient health monitoring, fleet management, and E911 wireless emergency services. Location aware billing can offer different rates for subscribers at different locations. Many organizations need to know if a malicious intruder device is attempting to gain access to their network and protected data, and to quickly take corrective action to thwart and deny access. At the same time, known trusted clients need to be given access in a secure and controlled way, abiding by the organization's security policies. For example, many authorized clients visiting an organization with a smart phone, laptop, or PDA, require limited access to the organization's internal information. These devices cannot be managed easily because people will non-deterministically visit and leave an organization with previous or possibly new devices at any time. Furthermore, an organization may limit access depending on the location of the device within the premises of the organization itself. For example, clients in the reception area would be granted stock quote data, while a client in a laboratory would be granted experimental results. An organization would like to automatically detect a wireless-device in the proximity of their wireless network, and based on the calculated location and subsequent identification of the

user of the device, grant access to a limited amount and type of information behind the protected wireless network.

The goal of this chapter is to present a literature survey of today's localization techniques and present the tradeoffs inherent in algorithm design.

2.1 Localization Parameters

The following describes several parameters that impact the ability to determine a node's physical location in a wireless sensor network.

Link Quality Estimator:

The Link Quality Estimator is a measurement of the received signal strength (RSS) in the presence of a noisy channel. A weak signal in the presence of noise will result in an estimated low link quality value [5], and consequently, a low RSS. An outlier detection scheme has been proposed to filter out links which are far away from the target, thus improving the overall RSS and observed link quality [5].

Hearability:

Hearability refers to the degree to which a receiver can "hear" the transmitter's radio signal, which depends on the radio range of the transmitting device and the relative distance between the transmitter and receiver. Hearability has a great impact in cellular communication systems. Multiple base stations (BSs) must be able to sufficiently "hear" the signal of a mobile station (MS) (i.e., node) to identify its respective location. Unfortunately, often a BS that is located far from the MS cannot "hear" the signal due to poor coverage [6], [7].

Accuracy:

Signal accuracy of a transmitting device varies due to channel impairments. Indoor localization is highly affected by channel errors. Also, accuracy of the signal for range based approaches is affected by the hardware used.

Line of Sight:

The precision of mobile station location estimation suffers from the non-line of sight (NLOS) problem. The NLOS problem causes a signal to be received at an incorrect angle and at a later time. Transmission mediums greatly affect node localization algorithm. NLOS introduces erroneous distance from the positive bias magnitude of time or angle that is dependent on the propagation environment [8].

Target Device Type:

Wireless technology supports multi-platform devices. For the transmitter, throughput and antenna architecture varies from device to device. Even if two different hardware devices are located 10 m away from a receiver, signal strength varies for the same distance. Different types of antennas have their own designs and radiation patterns.

Target Device Location and Orientation:

The carrier for the transmitting devices are movable and do not have any fixed location or orientation. Targeted devices must be in a location covered by wireless and has cleared LOS.

Environment Structure and Layout:

The trajectory of the signal often suffers from interference, multi-path propagation, attenuation, fading, shadowing, reflection, and the NLOS problem.

2.2 Wireless Sensor Network Node Localization Technologies

The node localization function plays a vital role in applications that require locating nodes in a wireless sensor network. Below are the technologies that are widely used for wireless node localization applications.

2.2.1 Cellular Communication System

Cellular communication systems use the global system for mobile communications (GSM) to estimate the location of a mobile device. The distance of the mobile station can be estimated from the respected neighbouring based stations. The precision of the location estimation highly depends on propagation conditions of wireless channel, and coverage of BTS.

2.2.2 Wireless Local Area Network

The detection, location determination, and tracking of wireless-devices in the proximity of an organization's premises are of growing concerns. The current positioning technique equips the target and realizes localization by utilizing the wireless measurements between the target and anchor nodes. Table 1 shows some major specifications used for wireless local area network.

Table 1 Major specifications of the 802.11 wireless standard.

Specifications	**Wireless Standard**
Standard	802.11 a/b/g
Frequency	2.4 GHz
Channel	1,6,11

Security	WPA
Network	SSID
SSID	Broadcast
IP	DHCP
Antenna	Omni directional

2.2.3 Radar

"Radio Detection and Ranging"- RADAR uses reflected radio waves to determine the location of an object. The application of RADAR includes detection of motor vehicles, aircraft, ships, military applications, weather forecasting, and mapping composition of the Earth's crust.

2.2.4 Sonar

"Sound Navigation and Ranging"- SONAR uses reflected signal pulses to determine the location of an object. SONAR determines the time measurements between the pulse emission and the reflected pulse. Respected distance can be calculated from this measured time and the propagation velocity.

2.2.5 Global Positioning System (GPS)

Global positioning system is a widely used localization technique for wireless node localization. GPS is a costly and energy consuming methodology. GPS is suitable for outdoor positioning where line of sight is available.

2.2.6 RFID

Radio Frequency Identification uses RFID tags to locate the position of a wireless device. Radio signals are transferred as the identification code of an electronic tag to a receiver. Signal strengths are obtained from the response messages from the tags and onwards RSS based localization technique is used to calculate the position.

2.2.7 Bluetooth

Bluetooth operates in the same frequency range as Wi-Fi. Bluetooth technology is used in short range localization. It has lower power consumption. Link quality and bit error rate are the mostly used parameters for the Bluetooth based positioning method. These parameters are not clearly defined in basic Bluetooth specifications, and thus are dependent on the equipment manufacturer.

2.2.8 Ultrasound

Ultrasound uses high frequency sound waves to determine the location of a wireless device. The ultrasound positioning target wears an ultrasound discharge tag and receivers are located in each targeted location. Many biomedical applications, chemistry and military applications use the ultrasound positioning technique, which is sensitive to the line of sight (LOS) requirement.

2.2.9 IrDA

“Infrared Data Association”- IrDA positioning technique uses infrared light transceivers to determine the position of a wireless device. IrDA is sensitive to the line of

sight (LOS) requirement. The active badge technique in IrDA requires a device or badge which periodically emits its ID code through IR transmitter for location detection.

2.3 Different Types of Localization Techniques in WSN

There are many different methods used for determining the location of a wireless device. There are also many different types of measurements which can be taken and used in node localization, i.e., to determine the physical location of a wireless device in a wireless network, such as client collaboration, time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), received signal strength indicator (RSS), fingerprinting, and pattern recognition.

2.3.1 Centralized verses Decentralized Algorithms

Centralized techniques are reliant on a central workstation which collects transmitted data from sensor nodes. Doherty, Pister and Ghaousi [9] developed a centralized technique using convex optimization to estimate positions. This technique is not feasible for mobile applications as it will draw a single dependency and it has a built in delay.

Decentralized approaches do not require centralized computation. These approaches can be further also categorized as range based or range free methods. The decentralized approach only requires few parameters with which it can communicate with the reference nodes.

2.3.2 Cooperative verses Autonomous Algorithms

There are generally two main approaches that have been used to detect and localize wireless devices in a network: cooperative and autonomous. The work done by [10] and [11] aimed to determine a fundamental limit of accuracy for a given localization algorithm. Unlike the TOA or RSS methods, these authors tried to use all of the available information of the captured signal. The method requires localization information from individual anchors and a priori knowledge of the agent's position in the network, which may not be known in a general scenario. Also, this method is based on the cooperative approach, which relies on (possibly untrusted and malicious) target devices to cooperate in the localization process.

Another work that is based on the cooperative approach is [12].This work describes how time-of-arrival (TOA), angle-of-arrival (AOA), and received-signal-strength (RSS) measurements can be used cooperatively to detect a device's location in wireless sensor networks; but this work requires a higher infrastructure cost. Along the same lines, several off-the-shelf technologies for wireless device location detection have been proposed, such as GPS, cell phone, and infrared (IR). These cooperative based approaches require the installation of custom hardware and/or software on the target device or establishment of anchors and agents in the network, and they must work in conjunction with the remote detection system to identify the target's physical location. However, the required subsystems may not be installed, or they can be easily removed and defeated. Or, the data may be altered on the target device before transmission to the detection system. Consequently, even a novice hacker may avoid detection or misrepresent of their location, leaving the system ineffective. To avoid these problems, this thesis uses an autonomous

system, which does not need any special hardware or software installed on the target device or any cooperation from the target device.

2.3.3 Taxonomy of Node Localization Techniques

The node localization technique is implemented within the reference network to locate the position of a wireless device. Node localization techniques use energy, distance or angle based estimation methods. The node localization technique can be further classified into two methods: Range Based Localization and Range-Free Localization.

2.3.3.1 Range-Based Localization Techniques

The ranged-based approaches consider the absolute distance between the target and the receiver devices. Range-based techniques are hardware dependent and also costly in most cases. They need a large number of devices to achieve high accuracy. Range-based methods suffer from the NLOS problem. Most popular range based localization techniques are TOA, TDOA, AOA, RSS etc. Some of the more recent work includes popular multidimensional scaling algorithm [13], belief propagation based algorithms [14], iterative algorithms [15] etc.

2.3.3.1.1 Time of Arrival (TOA)

Time of arrival (TOA) considers the signal propagation time between the transmitter and receiver device when speed is known. TOA varies along with the propagation medium. TOA needs clock synchronization that often requires expensive hardware [16]. GPS is one

example of TOA. The TOA based technique has been used in [17] [18]. The measurements and test bed experiments demonstrate 1-m RMS location errors using TOA in [17].

2.3.3.1.2 Time Difference of Arrival (TDOA)

Time difference of arrival (TDOA) gives better accuracy but requires two types of transmitter and receiver devices. To get better synchronization, TDOA provides the start of the transmission information to the receiver. TDOA needs prior ranging information and more complex hardware support as shown in [19], [20].

2.3.3.1.3 Angle of Arrival (AOA)

The systems based on AOA increases the overall cost of the system as it requires a larger number of antennas [21]. Location information is obtained by estimating and mapping angles between the transmitter and receivers in [22] [19].

2.3.3.1.4 Received Signal Strength (RSS)

The RSS technique involves the path loss model assuming known transmission power, received power and path loss coefficient. RSS based techniques alone are not sufficient for the accurate localization technique. Fig. 2 shows the three reference anchors that are located r1, r2, and r3 distance apart from the wireless transmitter. Position of the target can be estimated from the intersection of the three circles around Anchor1, Anchor2, and Anchor3. The RSS based technique has been used in [17], [18]. The measurements and test bed experiments demonstrate 1- to 2-m RMS location errors using RSS in [17].

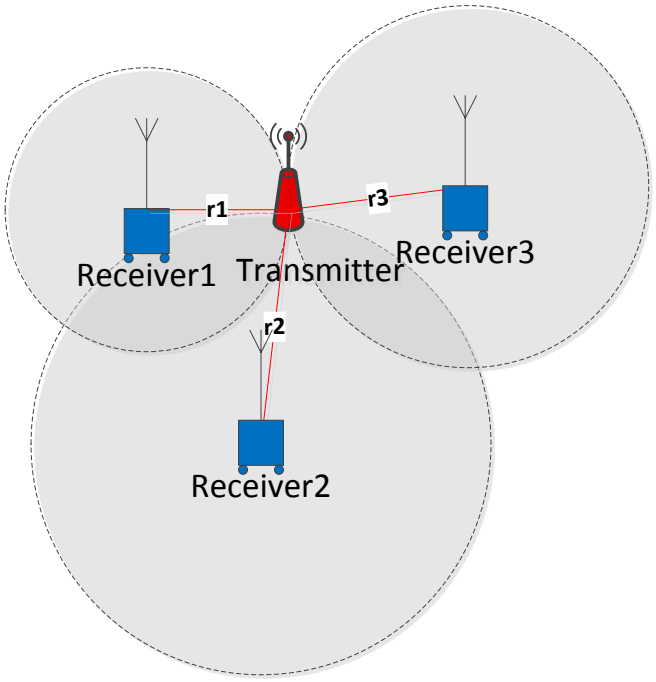


Fig. 2. Trilateration technique.

2.3.3.1.5 Trilateration

Trilateration uses three reference access points (AP) to calculate the position of a targeted device. Trilateration can be used for real time positioning systems, but it suffers from the LOS problem [23]. Lateration based on more than three AP can improve the localization result but will increase the dependency and complexity of the system.

2.3.3.1.6 Fingerprinting

Radio frequency (RF) fingerprinting collects and gathers on-site data and develops a customized model that helps to determine the distance between the sender and the receiver. Chan and Sohn [23] shown that probability of the fingerprint matching algorithm using RSS is not sufficient for realistic localization applications. Therefore, there is a need for range free techniques for this research area.

Drawbacks:

1. There is a significant trade-off between ranging accuracy and complexity of the system.
2. RSS based approaches are simpler but ranging errors are significantly higher.
3. The percentage of correctness decreases with the increase of distances for the Trilateration approach.
4. The TOA requires accurate clock synchronization and expensive hardware.

2.3.3.2 Range-Free Localization Techniques

Range free localization techniques are very appealing and popular today. Range free techniques are cost effective and reduce hardware dependency. Range free techniques do not employ the range based parameters directly to compute the distance of a wireless device. Range-free localization techniques do not require the actual distance measurements between unknown targets and the known receivers. The differential RSS measurement based robust device free wireless localization technique is applied in [5] to compensate for the negative effects from the environment. But, still this RSS based differential technique in [5] could not classify the variation of received signals.

2.3.3.2.1 HOP Count Based Localization Techniques

The HOP count localization technique propagates packets further on direction to determine the position of the anchor nodes. The HOP count technique requires fewer anchor nodes. Different types of HOP-count localization techniques are elaborately described in the literatures. The combined DV-HOP count and lateration technique is used

in [24]. They describe the information exchange procedure of an anchor node containing its own location along with the HOP count information. HOP-Terrain method is introduced in [25] at intermediate phase of the transmission. After obtaining the information using DV-HOP count method, transitional nodes send a broadcast message to a specific anchor if the current received message is less than the previous one. In the next stage, the DV_HOP uses the lateration technique to find the location. A multi-lateration HOP count technique is proposed in [26]. The targeted node calculates the shortest distance path to the anchor nodes and builds a bounding box for each anchor. The number of HOP count increases the overall range error in this system. Traditional HOP count localization techniques are not suitable for practical implementation in real world applications. The system complexity, transmission overhead, and precision are shown in [24], [25], [26], [27]. HOP count ratio based localization reduces the transmission overhead by 50% as shown in [28]. System accuracy could not be achieved due to environment dependency.

2.3.3.2.2 Heuristic Based Localization Techniques

Heuristic techniques are experience-based problem solving methods that are greatly used in wireless sensor network localization research. These techniques do not require the priori distance information between the target and the receivers. The localization problem has been known to be NP-Hard [29] for its complexity. Heuristics quickly produce a good enough solution for solving such a problem. There are different types of heuristic approaches (Fig. 3) that are widely used for WSN localization applications.

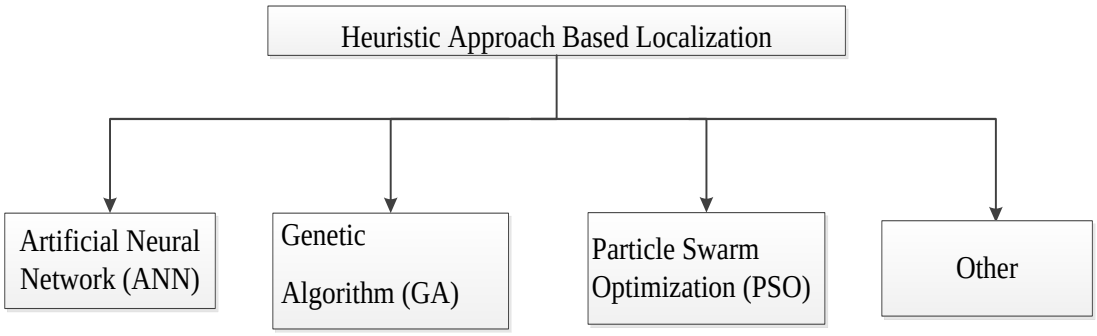


Fig. 3. Heuristic approach based localization method taxonomy.

Artificial Neural Network (ANN):

Different types of neural network techniques are widely used in WSN localization applications. Artificial neural networks are used to solve the nonlinear relationship between localization parameters [3]. Much of the literature describes different neural network techniques to solve the WSN localization problem [30], [31]. Section 2.4 explains in detail the current ANN approaches used in WSN localization applications.

Genetic Algorithm (GA):

The genetic algorithm was used in WSN localization to search the solution space. A few GA based algorithms are presented in [32] [33]. GA predicted results with an error less than 0.01% compared to the actual results after 33 iterations in [34]. GA converges very slowly in their search to detect an exact solution within polynomial time. Moreover, stopping criteria selection is not straightforward in GA.

Particle Swarm Optimization (PSO):

The PSO based algorithms for WSN localization is described in [35], [36]. But the PSO model provides saturated but slow convergence. H-BEST PSO introduced in [37] is more

effective in terms of faster and matured localization as compared with global best PSO. Still for a large environment HPSO, PSO is not suitable and yields less accuracy than biogeography-based optimization.

2.4 Artificial Neural Network for Location Estimation

Neural networks can be defined as a nonlinear function approximation. Neural networks are commonly used for several applications like classification, input-output mapping and prediction of future results. Different types of neural networks are used in medical studies, speech recognition, localization, WSN research, stock market, human behavior studies, robotics, etc. A neural network has the ability to learn, adjust, simplify, classify, and decide.

2.4.1 Introduction

Wireless sensor network localization is a nonlinear problem. The advantage of using artificial neural networks is that there is no need of advanced knowledge of the hardware and software components, system architecture, or environmental map. The algorithm is an automated and intelligent learning technique which is self-sufficient to calculate the position of the targeted device. Different types of artificial neural network approaches are discussed in the literature for WSN localization. Self-organizing maps [38], neural network techniques, and different manifold learning algorithms [39], multilayer perceptron (MLP), radial basis function (RBF), and recurrent neural network (RNN) are explained in [40].

2.4.2 Supervised vs. Unsupervised Networks

Unsupervised learning does not provided targeted output during training. Unsupervised learning is often used to cluster the input data into classes on the basis of their statistical properties. Only the input data set is provided for unsupervised learning. Supervised learning consists of a set of targeted output with respected input. The learning algorithm needs to achieve the targeted output. We consider supervised learning techniques in our research for WSN localization.

2.4.3 Multi-Layer Perceptron (MLP)

Multilayer perceptron (MLP) is a supervised learning method. MLP is a feed-forward artificial neural network model that maps sets of input data onto a set of appropriate outputs [41]. The Back-propagation neural network is the simplest learning method for ANN. Other frequently used learning techniques for ANN are the gradient decent(GD), conjugate gradient algorithm (CGA), Levenburg-Marquardt (LM) and resilient BP (Rprop) which are robust enough to run training parameters used for WSN localization applications [3], [31].

Supervised and unsupervised learning algorithms are described in literature for WSN localization applications. Conjugate gradient supervised learning algorithms are used in literature to solve the large problem of linearity [42]. The scaled conjugate gradient algorithm (SCGA) was designed to avoid the time-consuming linear search. The SCGA algorithm is too complex to explain in a few lines, but the basic idea is to combine the model-trust region approach, with the conjugate gradient approach. SCGA used in the WLAN location determination model [30] which shows 57% of the training samples

produce a distance error of 30cm. In [31], two different learning algorithms have compared. The LM provides a smaller average location error than the Rprop. The LM method was used in the indoor WSN localization application for LOS and NLOS sites. Simulation results show that RMSE increases from 0.54m to 5.1m for the LOS and 0.69m to 6.1m for the NLOS condition for 10 times greater area [30]. GD with momentum helps to recover from local minima. It works as a low pass filter to avoid diminutive features in the error surface.

Few research articles depicted the cascade connected neural network structure. Positioning in WLAN is described using the single MLP neural network and cascade-connected ANN [43]. The cascade-connected ANN approach provides additional preprocessed information which is fed to the network using the BP learning technique. The cascade-connected ANN method performed well in the WLAN environment. MLP with error back propagation has considerable benefits when compared with other learning techniques [44]. BPNN can be used to train data with a given set of input and corresponding set of output to overcome the problem of nonlinear relationship. Gholami and Brennan presented an ANN approach and signified that noise intensity, location of anchors, and their interaction have significant impact on neural network architecture [45]. The literature review clearly states a need of some more information of indoor environments which can help to find the position of the targeted device.

2.4.4 Advantage of Artificial Neural Network

The artificial neural networks are able to resolve any exclusive problem. We can test the system with an exclusive data set in Fig. 4.

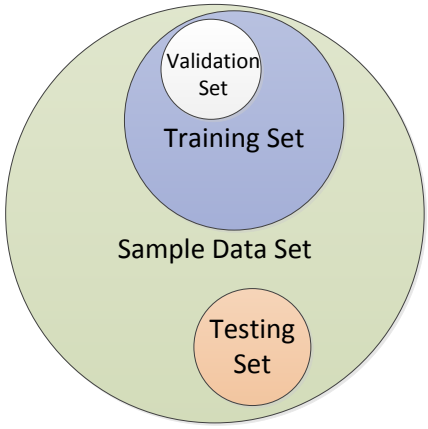


Fig. 4. Partitioning data set.

Neural networks are able to adapt the characteristics of the system. Neural networks can classify, approximate, and solve the problem of nonlinear relationships.

2.5 Fractal Theory

Indoor RSS is a mixture of many other signals that are generated from reflection, attenuation, dispersion, multipath fading, etc. The propagated signal shows a stochastic, non-linear pattern. Fractal theory can be used to measure the complex and stochastic nature of a radio signal.

2.5.1 Introduction

There is a trade-off between the complexity of the algorithm along with the cost and distance ranging function. The location patterning approach using RSS as the input set requires the non-linear activation function for the classification of a signal for a fixed position. Signal classification is needed for indoor localization. Complexity analysis of a signal can extract the underlying features. The signals produced from non-linear systems are too complex in nature to analyze with the traditional techniques such as the Fourier

transform, which often depends on the stationarity of the a signal and linearity of the analyzed system. The radio signals of interest in this thesis are not only stochastic in nature, but are also contaminated with different channel errors and consist of multiple signals. Stochastic broadband electromagnetic noise signals, interference by other signals, fading, shadowing, reflection, and attenuation need to be considered for defining a complete description of the physical process from which a Wi-Fi signal is generated.

Traditional techniques are not sufficient enough to completely define the overall signal and classify its features. A different transformation technique is needed. This thesis uses the variance fractal dimension to measure a signal's complexity and character. The multi-fractal characterization of the signal using variance fractal dimension trajectory (VFDT) can be observed and used as a tool to classifying the RSS parameter for indoor localization.

2.5.2 Variance Fractal Dimension

Different types of features can be obtained in different ways to measure fractal dimension. The transformed fractal dimensions rely on altering the original fractal into a different domain in which its properties can be observed [46]. The variance fractal dimension transforms the signal from the time domain to the variance domain. But VFD is not same as the variance of process. The variance fractal dimension trajectory is used in the literature for the detection of boundaries of a noise utterance [47]. The variance fractal dimension algorithm is used to produce trajectories as features of speech. These trajectories are used for classification and recognition of the speech [47]. Moreover, the variance fractal dimension trajectory is successfully used for heart sound localization at

low and medium flows [48]. The multivariate signal obtained from the nonlinear system is used for classification of the signal through its VFDT in [49].

2.5.3 Benefit of using Variance Fractal Dimension

Fractal and chaos theory based models have proven to be more suitable than the classical stochastic models in describing the dynamic principle of multipath fading channels [50]. This work shows that the multipath fading channel has a fractal property. The reconstructed signal through fractal interpolation has a very refined and rational construction due to the self-similarity property.

The RSS consists of a multipath fading signal which is nonlinear in nature and hard to describe alone with traditional localization techniques. Fractal theory is suitable to resolve such non-linear problems and reconstruct the trajectory.

Fractal analysis based on variance is unaffected by noise. The simplicity of the variance fractal dimension algorithm is appropriate for real time fractal analysis. Moreover, the variance fractal dimension requires a window in the time series, which excludes windows artifacts obtained from the Fourier domain [51].

2.6 Summary

Our proposed integrated artificial neural network (IANN) based localization algorithm uses wireless local area network (WLAN) architecture. A change in the WLAN architecture may introduce the need to retrain the system. Long term dependency varies with the change of complexity of the system. The variance fractal dimension measures the complexity of RSS which is contaminated with channel errors. Usually range-based

techniques are more accurate, but are expensive and complex for practical implication. Range-free techniques would be better if they could provide greater node localization accuracy. Our proposed range-free localization algorithm uses additional parameters that would not only help, but obviously would outperform general range-free techniques.

Chapter 3

3. Integrating Fractal Dimension with ANNs for Node Localization

3.1 Trilateration Approach

We placed three wireless receivers in a 2D grid, as shown in Fig. 5. The locations of the receivers were known. The relative position from each of these three receivers to the unknown transmitter (aka, client) was calculated using the Euclidian distance. We solved the relative distance between the receiver and transmitter. We obtained three circles around each of the receivers using this distance as radius. The Trilateration technique determined the intersect point of the three circles to be the actual position of the unknown client.

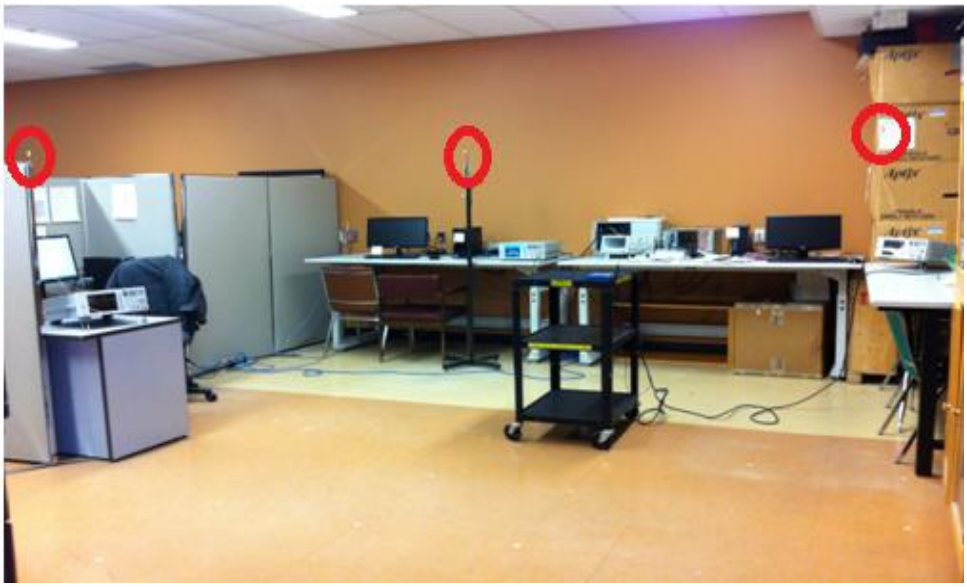


Fig. 5. Trilateration for 2D coordinate system (E1-527).

3.1.1 Received Signal Strength:

We determined the average RSS value for each of the positions by calculating the statistical mean of total N samples, as given by (1).

$$RSS = \frac{\sum_{i=1}^{N} RSS_i}{N} \quad (1)$$

In (1) the *RSS* represents received signal strength of *ith* samples, *i* is the current sample number, and *RSS* is the mean of total N samples.

3.1.2 Mathematical Analysis for Indoor Path Loss Model for the Wireless Client

The indoor signal propagation suffers from signal attenuation. The path loss model for indoor propagation is defined as (2) [52]:

$$PL = PL_{1M} + 10 * log\ (d^n) + s \quad (2)$$

In (2) *PL* is the total path loss experienced between the receiver and transmitter in dB; PL_{1M} represents the reference path loss in dB when the receiver-to-sender distance is 1 meter; *d* represents the distance between the transmitter and receiver in meter; *n* indicates the rate at which path loss increases with the distance; and *s* represents the standard deviation of the signal.

Path loss is the difference between the transmitted signal and the received signal. *PL* shows the signal attenuation due to environmental impairments. Here antenna gain and cable losses are not taken into consideration. The rate of path loss is also known as the path loss exponent. The path loss exponent for densely deployed space is greater than free

space. The path loss exponent *n* consists of deterministic parameters of the environment, such as the location and height of the antennas, antenna radiation pattern, and reflections or multi-path signal propagation, which cause significant constructive or destructive interference of signals at the receiving antenna. A typical path loss exponent for an indoor office environment may be 3.5, while for a dense home, commercial or industrial environment it is 3.7 to 4.5; Hence, we considered the path loss exponent to be 4.5. In many indoor installations using diversity antennas, the standard deviation of shadow fading is often between 3 and 8.5 dB. The standard deviation represents the variation of signal strength with a Gaussian random variable with zero mean and standard deviation of σ dB. This number could vary according to environmental impairments.

The general method to calculate received signal strength, given known quantities for transmit power, path loss, antenna gain, and cable losses is as follows [52] :

$$RX_{PWR} = TX_{PWR} - Loss_{TX} + Gain_{TX} - PL + Gain_{Rx} - Loss_{RX} \quad (3)$$

Here, RX_{PWR} is the detected received signal strength in dB; TX_{PWR} is the detected transmitted signal strength in dB; $Loss_{TX}$ and $Loss_{RX}$ are the transmit-side and receive side cable and connector loss in dB; and $Gain_{TX}$ and $Gain_{Rx}$ are the transmit-side and receive side antenna gain in dBi.

3.1.3 Distance Calculation with Trilateration Approach

The Pythagoras formula is used to represent circles as follows:

$$(x_1 - x_c)^2 + (y_1 - y_c)^2 = r_1^2 \quad (4)$$

$$(x_2 - x_c)^2 + (y_2 - y_c)^2 = r_2^2 \quad (5)$$

$$(x_3 - x_c)^2 + (y_3 - y_c)^2 = r_3^2 \quad (6)$$

Assuming that there are three anchors with known positions (x_1, y_1), (x_2, y_2) and (x_3, y_3), and the unknown client position is (x_c, y_c), then the relative distances from the unknown client to the three anchors are r_1, r_2 and r_3.

Now subtracting (6) from (4) and (5) we find the below equations,

$$(x_1 - x_c)^2 - (x_3 - x_c)^2 + (y_1 - y_c)^2 + (y_3 - y_c)^2 = r_1^2 - r_3^2 \quad (7)$$

$$(x_2 - x_c)^2 - (x_3 - x_c)^2 + (y_2 - y_c)^2 + (y_3 - y_c)^2 = r_2^2 - r_3^2 \quad (8)$$

Now rearranging (7) and (8), we find the below equations,

$$2(x_3 - x_1)x_c + 2(y_3 - y_1)y_c = (r_1^2 - r_3^2) - (x_1^2 - x_3^2) - (y_1^2 - y_3^2) \quad (9)$$

$$2(x_3 - x_2)x_c + 2(y_3 - y_2)y_c = (r_2^2 - r_3^2) - (x_2^2 - x_3^2) - (y_2^2 - y_3^2) \quad (10)$$

Converting (9) and (10) as a linear equation of A*x=B we find,

$$2 * \begin{bmatrix} (x3 - x1) & (y3 - y1) \\ (x3 - x2) & (y3 - y2) \end{bmatrix} * \begin{bmatrix} xc \\ yc \end{bmatrix} = \begin{bmatrix} (r12 - r32) - (x12 - x32) - (y12 - y32) \\ (r22 - r32) - (x22 - x32) - (y22 - y32) \end{bmatrix} \quad (11)$$

Solving (11), the unknown client's position (x_c, y_c) was determined.

3.2 Fractal Theory

Fractals are geometrical phenomena that have a normal or irregularity behaviour repeated over all scales of measurement, and they have a fractional dimension representing complexity. In this thesis, fractal analysis is used as a tool to measure the complexity of the received signal strength. The significance of this study is to discover the extended range dependencies of the received signal strength, and to extract possible features from the Wi-Fi signals.

The variance fractal dimension (VFD) has been utilized to approximate the complexity of a stationary and self-affine (monofractal) time series signal in the literature [51]. However, a single scalar value cannot represent the complexity of a multi-fractal time series signal. The variance fractal dimension measure is calculated repeatedly over a piecewise stationary sliding window, and this process produces a trajectory of variance fractal dimensions. In this thesis, the variance fractal dimension trajectory (VFDT) has been used to measure the complexity of the received signal strength. The experiments show that the variance fractal dimension along with RSS yields better classification than simply using captured RSS triplet set as input for artificial neural network based localization algorithm.

3.2.1 Statistical Analysis of the Signal

Statistical analysis of the signal is necessary to perform multi-scale fractal analysis. We captured the histogram and performed a stationarity analysis of the signal.

3.2.1.1 Histogram

We observed the distribution of the signal from the histogram. By normalizing the histogram, we obtained the probability mass function. The histogram is sub-Gaussian which reveals the underlying properties of the signal. The histogram can extract many features of the signal.

3.2.1.2 Stationarity Analysis

A signal is said to be stationary if the statistical measurements of the signal do not change with time. We observed the 1^{st} moment (mean), 2^{nd} moment (variance), 3^{rd} moment (skewness), and 4^{th} moment (kurtosis) of the signal for different time intervals. If the first two moments do not change by ±5% with time, we could consider the signal as weak sense stationary. If none of the moments change with time, the signal is strong sense stationary. When the statistical properties of the signal do change with time, the signal is known as non-stationary. If the process is stationary, the statistical properties of the entire signal may be obtained from any portion of the signal.

3.2.2 Feasibility Test to Perform Fractal Dimension

The RSS is captured in the frequency domain using a spectrum analyzer. We converted the signal into a time series and obtained the histogram from the signal. We analyzed the histogram and the distribution of the signal. If the histogram was satisfactory, we proceeded to do stationarity analysis shown in Fig. 6. Otherwise we recaptured the signal. Stationarity analysis of the signal tells how the statistical properties of the signal vary over

time. If we achieve a 95% confidence level for different moments, the signal will be eligible to perform fractal dimension analysis.

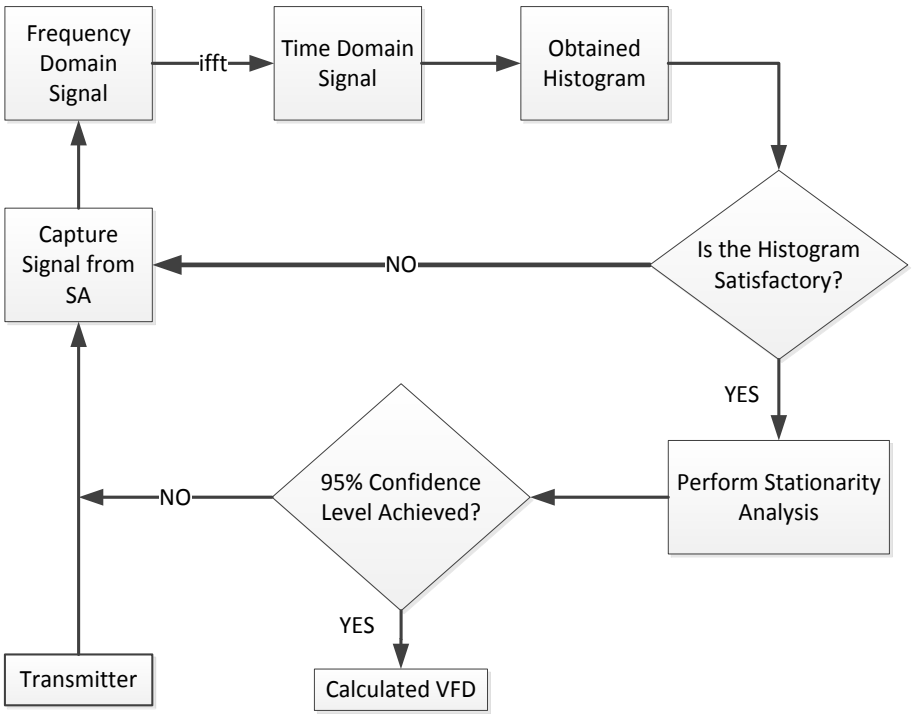


Fig. 6. Flowchart for feasibility test to perform fractal dimension.

3.2.3 The Reason Behind Choosing VFD

Our analysis showed that the Wi-Fi signal is non-stationary as a whole. Other FDs are not suitable to perform multi-fractal analysis for the entire signal. However, the signal is piece-wise stationary in some of the windows. The VFD allows to deal with piecewise stationary signal in a real and practical environment. Though it may not be as complete as Renyi dimension, but is suitable for piecewise stationary signal. Moreover, as discussed in [51], VFD does not require a window in the Fourier sense. Data to be transformed could be

broken down into frames, usually overlapped with each other. The VFD avoids the windowing artifact and could be computed in two different ways: batch and real-time.

3.2.4 Variance Fractal Dimension

Variance fractal dimension is not the same as the variance of the process. Let us consider the received signal strength indicator as *B(t).* The variance σ^2 of its amplitude increments over a time increment is related to the time increments over the below power law proposed in [51],

$$Var[B(t_2) - B(t_1)] \sim |t_2 - t_1|^{2H} \tag{12}$$

From (12), we understand that the variance increases with the increase of the time interval. Here *H* is called the Hurst exponent. The time displacement over which the amplitude increment is being measured is,

$$\Delta t := |t_2 - t_1| \tag{13}$$

$$(\Delta B)_{\Delta t} := B(t_2) - B(t_1) \tag{14}$$

The slope can be calculated by using least squares regression on a set of ordered pairs, as proposed in [51].

$$s = \frac{k\sum_{i=1}^{k} X_i Y_i - \sum_{i=1}^{k} X_i \sum_{i=1}^{k} Y_i}{k\sum_{i=1}^{k} X_i^2 - \left(\sum_{i=1}^{k} X_i\right)^2} \tag{15}$$

In (15) N is the number of ordered pairs used.

$$H = \frac{1}{2}s \tag{16}$$

Finally, the variance fractal dimension can be obtained from (17),

$$D_\sigma = E + 1 - H \tag{17}$$

In (17) E=1 for a single variable time series and range of H is limited to 0<H<1. Thus D_σ must be limited for $1 < D_\sigma < 2$. Representation of the symbols for VFDT algorithm is shown in Table 2.

Table 2 Parameters for the VFDT algorithm [51]

Representation of Symbols	
Symbol	**Description**
ΔB	Change in magnitude of the signal
Δt	Time displacement
B(t)	The signal
σ^2	Variance
E	Given no of independent variables
D_σ	Variance fractal dimension
i	Iteration
δt	Unit time interval
n_k	Size of the window
K_{hi}	Maximum allowable separation
K_{low}	Minimum allowable separation
H	Hurst exponent
T	Entire time interval
b	adic sequence
k	A variable refers to no of windows in current stage
s	Slope
d	Displacement of interval
j	Variable
N	The piecewise window of the signal

3.2.4.1 Variance Fractal Dimension Algorithm

The variance fractal dimension was computed in a similar manner to the algorithm as described in [47].

Table 3 The variance fractal dimension algorithm [51].

Step	Procedure
1.	Compute the non-overlapping intervals from the stationarity analysis of the entire signal.
2.	Obtained a reasonable window displacement for each interval.
3.	Find the number of windows and size of the window for this interval.
4.	For maximum allowable separation of the signal find the time displacement. We choose the diadic sequence such that, $\Delta t_{max} = n_{Kmax} . \delta t \leq T$ and $n_k = 2,4,8 \ldots$
5.	Set $N_1 > 30$, so that the variance is statistically justified.
6.	Calculate K_{max} =int(log(N_T)/log(b)) and K_{buf} = ceil(log30/log(b) where, $b_{Kmax} \leq N_T$
7.	Obtain maximum allowable separation of two points of Δt, $K_{hi} = K_{max}=K_{buf}$
8.	Obtain minimum allowable separation of two points of Δt , $K_{low} \geq 1$
9.	For k= K_{hi} to K_{low}, calculate the variance for all the windows (k) in the current stage. $$\mathrm{var}(\Delta B)_k = \frac{1}{Nk-1}\left[\sum_{j=1}^{Nk}(\Delta B)_{jk}{}^2 - \frac{1}{Nk}\sum_{j=1}^{Nk}(\Delta B_{jk})^2\right]$$
10.	Calculate two values, $X_k = \log[n_k]$, $Y_k = \log[\mathrm{var}(\Delta B)_k]$the point on the log-log plot for current stage.
11.	Repeat the steps(4-10) for other stages for this interval and obtain slope s from the log-log plot.
12.	Compute the Hurst exponent from the slope.
13.	Compute the Variance fractal dimension, Dσ.
14.	Move forward towards the end of the displacement of interval. We continue steps (3-13) for the next sequence of interval windows and obtain corresponding value of VFD for the respective interval.
15.	This sequence of dimension will create the unique VFDT for the entire signal.
16.	Final interval can be omitted for the VFD calculation.

The VFD is calculated at different scales over and over. The measurement is no longer a vector or scalar; it is a poly-scale measure. The minimum number of windows is considered to be 32 to statistically justify the variance [51]. Magnitude of the slope indicates the degree of complexity.

3.2.5 Variance Fractal Dimension Trajectory

The VFDT for a RSS signal is obtained from the calculation of the VFD within a sequence of signals, as given in Table 3. The VFDT was calculated to characterize the unique features of each signal. We observed that the VFDT was reflecting the similar but compressed trend as the time series plot of the RSS. The multi-fractal analysis of RSS was used as a tool to analyze the complexity and classify the unique property of the received signal strength for our proposed localization algorithm.

When the dimension is obtained in the first complete interval of N, it is recalculated on the next intervals of the same size, until the end of the time series is reached and the VFDT is obtained. The VFDT was computed using a window size of 512 samples to avoid erratic behavior of dimension trajectory and prevent loss of features, as suggested in [47]. The window displacement of 32 samples was used; as it produced a good resolution of the VFDT. The advantage of the VFDT, as explained in [53], is that it uniquely identifies the signal based on the fundamental complexity. One dimensional signal will have a complexity value in between the range1~2, after transformation.

3.3 Artificial Neural Network

Machine learning has been used widely in the literature to solve the WSN localization problem. Single layer Perceptron is not able to solve the nonlinearly separable problem. The localization technique described in this research involves the use of an integrated artificial neural network and fractal dimension approach to determine the spatial position of a wireless device. Diversified and large numbers of samples were collected for our study. No well-defined relationship can be obtained from the complex data. It is difficult to describe the system with conventional techniques. Artificial neural network helps to find the complex association among the parameters within a set of patterns. The ANN has the capability to analyze multiple stochastic relationships among the patterns and obtain a quick solution for the system.

3.3.1 System Architecture

In this thesis we use an ANN as an input-output mapping, using relationship between inputs- Received Signal Strength and targeted outputs - absolute coordinate position in the grid. We used the multilayer feed-forward neural network architecture to solve the WSN localization problem. The proposed ANN architecture has three layers: one input layer, one hidden layer, and one output layer, as shown in Fig. 7. By varying the number of processing neurons from 3 to 6 in the hidden layer, and applying gradient decent as a training algorithm for the system, performance was improved.

The training data set is fed initially to the ANN with a specific learning rate, and it begins to adjust the weights to minimize the errors. A non-linear transfer function with sufficient number of neurons guarantees the capability of the network for the relationship

among the parameters of learning data set. The error information is fed back to the system which formulates all tuning to their parameters in an orderly mode. The training of the network continues many times until the network achieves its targeted output. After the training phase, we setup the ANN for the validation phase with partial input of data sets and testing phase with totally unknown data sets, which were excluded from the training phase. Simplification occurs for a realistic output. The breakdown of the partitions of the input datasets is given in Table 4:

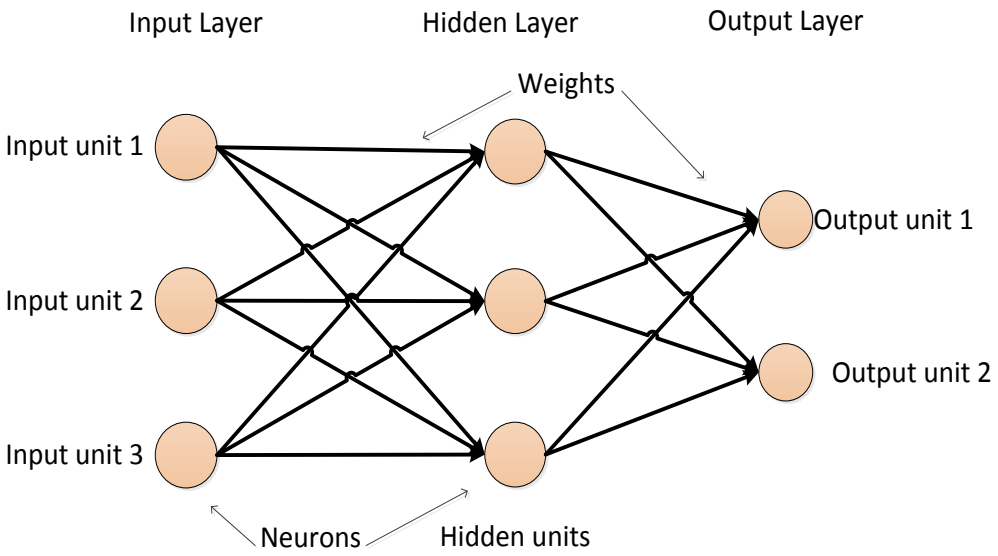


Fig. 7. A standard multi-layer perceptron architecture.

Table 4 Data set partition.

Data Sets	**Percentage of Data Set**
Training Data	85%
Validation Data	5%
Testing Data	10%

Table 5 ANN architecture for localization.

ANN Architecture	
Data division	Random
Free Parameters	Appropriate random initial values, 0-1
Training	Back Propagation
Performance	Mean Square Error
Samples	12500
Number of times repeated	10
Learning rate	0.01~ 0.2
Momentum	0.5
Layers	3
MSE	0.001
Epochs	3000

The ANN architecture is shown in

Table 5. We fed the captured received signal strength indicator values as input data set and corresponding spatial coordinates as output data set, as shown in Table 6. Here M is the total number of samples provided to the system.

Table 6 ANN training samples.

Triplet Input Sets	**Targeted Output**
P1= { RSS1,F1, RSS2,F1,RSS3,F1}	F1(X1, Y1)
P2= { RSS1,F2, RSS2,F2…..,RSS3,F2}	F2(X2, Y2)
……..	……..
PM= { RSS1,FM, RSS2,FM…..,RSS3,FM}	FM(XM, YM)

3.3.2 Forward Propagation

In the forward propagation stage, the inputs are associated with respected weights; the inputs pass forward to a non-linear activation function, and the outputs of the activation

function are used as inputs for the next layer to learn a training set of input output pairs { $P_{M,}$ F_M }. Table 7 shows the detail representation of symbols.

Three layers are connected with small value of weights, w_{pq}. The connection w_{ij} represents the weight from hidden units to input units, w_{jk} represents the weight from output units to hidden units. Initially we generated the weights randomly from 0-1. Weights connected with individual nodes in the hidden and output layers of a network are also known as a bias value. Practically, the bias value is constant, usually equal to 1, and are calculated in a manner as the same as the weight matrix. Hidden unit j receives net input for a given pattern M as,

$$h_j^M = \sum_i w_{ji} X_i^M \tag{18}$$

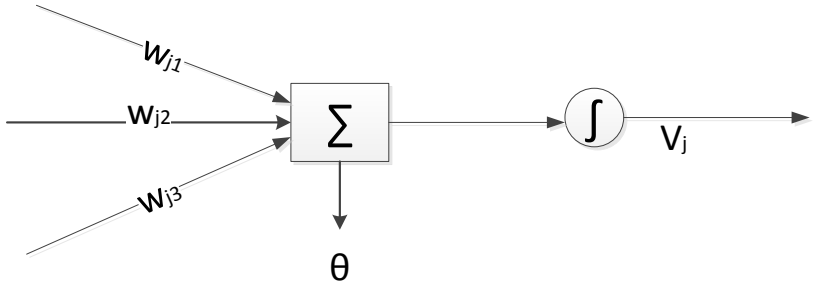


Fig. 8. Single neuron action.

The single neuron operation for our model is shown in Fig. 8.

The output produced from the *jth* unit of a hidden layer is,

$$(V_j^M) = g(h_j^M) = g(\sum_i w_{ji} X_i^M) \tag{19}$$

The output unit *k* receives a net input,

$$h_k^M = \sum_j w_{kj} V_j^M = \sum_j w_{kj}\, g(\sum_i w_{ji} X_i^M) \tag{20}$$

Final output from *k* unit of output layer is,

$$O_k^M = g\,(h_k^M) = g\sum_j w_{kj} V_j^M = g(\sum_j w_{kj}\, g(\sum_k w_{ji} X_i^M) \tag{21}$$

The cost function,

$$E(M) = \frac{1}{2}\sum_{Mk}[\,D_k^M - O_k^M]^2 = \frac{1}{2}\sum_{Mk}[\,D_k^M - g(\sum_j w_{kj}\, g(\sum_k w_{ji} X_i^M)]^2 \tag{22}$$

In (22) *M* is the number of samples with which the network has trained, and *k* includes all the neurons in the output layer.

Therefore the average error E_{av} was obtained by summing *E(M)* over all *M*, which is given by(23),

$$E_{av} = \frac{1}{M}\sum_{N\in k}^{M} E(M) \tag{23}$$

The purpose of the transfer function is to calculate a layer's output from its net input. The activation function applied in the system was non-linear, sigmoid (Fig. 9), and essential for gradient-descent learning. Another advantage of the sigmoid function is that these are smooth and differentiable functions, which is a requirement for the proposed ANN.

$$\sigma = g\left(h_k^M\right) = \frac{1}{1+e^{-h}} \tag{24}$$

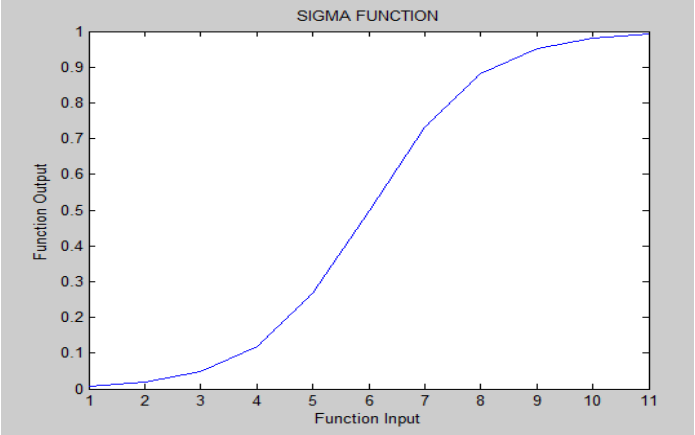


Fig. 9. Sigmoid function.

Table 7 Representation of symbols.

Representation of Symbols	
Symbol	**Description**
L	Layer
i	No of input layer neurons
j	No of hidden layer neurons
k	No of output layer neurons
M	No of patterns
m	No of hidden neuron
η	Learning rate
α	Momentum
Σ	Summation
X_i	Input Units
D_k	Targeted Output Units
w_{kj}	Weight from output units to hidden units
w_{ji}	Weight from hidden units to input units
Δw	Weight change
E	Error
C	Includes all the neurons in the output layer
O_k	Obtained Output Units
E(M)	MSE
E_{av}	Average MSE
h_k^M	Output unit k receives a net input

h_j^M	Hidden unit j receives net input for a given pattern M
V_j^M	Output produced from j unit of hidden layer
Δw_{kj}	Weight change from output units to hidden units
Δw_{ji}	Weight change from hidden units to input units
δ	Propagation error
σ	Non-linear sigmoid function
Θ	Bias

3.3.3 Back Propagation

Back propagation is a learning algorithm for multi-layered feed forward networks that uses obtained error from forward pass to propagate, calculate the local gradient and propagate the local gradient forward to the next layer. BP algorithm is described in Table 8.

Table 8 Back Propagation algorithm.

Steps	Procedure
1.	Data Structure: ANNmain: This function contains Input, Output, Weight and Bias matrix. BackPropagation: This function performs back propagation algorithm.
2.	Initialization: Input: Create a matrix to store the entire input pattern, initialize with training data set. Output: Create a matrix to store the entire input pattern, initialize with training data set. Weight: Generate small random values of weight and store in a matrix. Bias: Create a bias matrix with all values equal to 1. Network: Create a network consist of 3 layers. Create a matrix with 3 input layer nodes, 3 hidden layer nodes and 2 output layer nodes. Stopping Condition: Epoch=3000 or MSE=0.001.
3.	While stopping condition are false, do steps 4-10.
4.	For each training pair, do steps 5-9.

5.	for n=1:M Propagate inputs with associated weight shown in Fig. 10. End.
6.	Calculate activation output for all hidden layer nodes.
7.	Propagate the net input along with associated weight towards output layer.
8.	Perform summing function and calculate final output for each output node.
9.	Calculate the error, change in weight and gradient decent for the output layer.
10.	Backward propagate the gradient decent to update the weight.
11.	Calculate MSE for each epoch, update epoch and go to step 3 for next epoch.
12.	Calculate the average MSE, plot epochs Vs MSE.

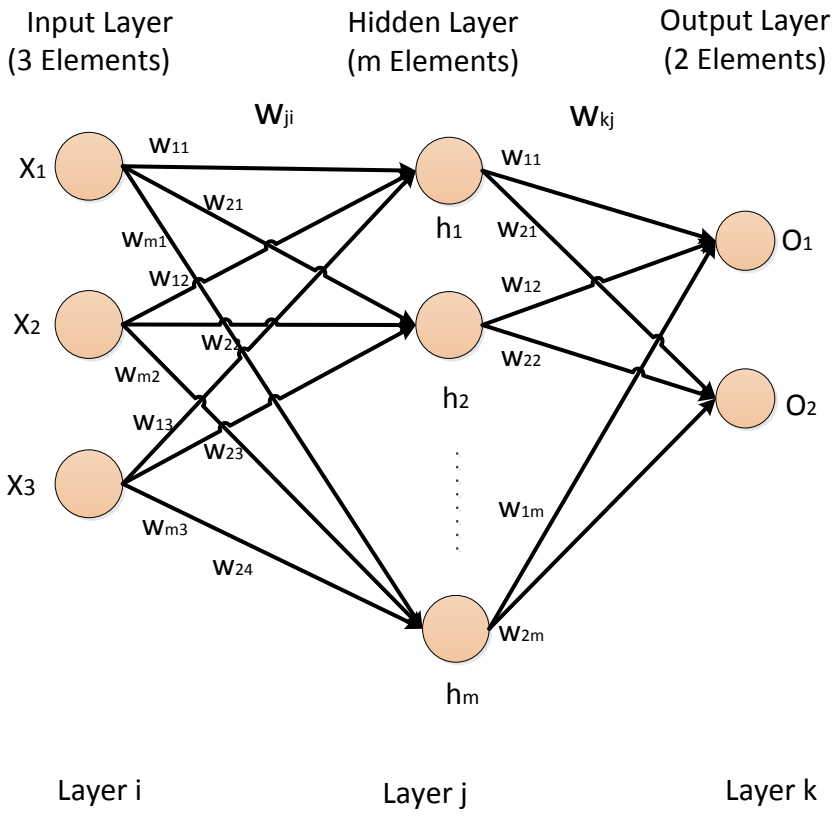


Fig. 10. ANN architecture.

The rate of change of error in output layer is,

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial}{\partial w_{kj}}\left(\frac{1}{2}\sum_{Mk}[D_k^M - O_k^M]^2\right) \tag{25}$$

$$\frac{\partial E}{\partial w_{kj}} = \sum_{Mk} \left(\left(D_k^M - O_k^M\right) g'\left(h_k^M\right) \left(V_j^M\right) \right)$$

$$\frac{\partial E}{\partial w_{kj}} = \sum_{Mk} \left(\left(V_j^M\right) \delta_k^M \right)$$

The propagation error or local gradient in output layer can be written as,

$$\delta_k^M = \left(D_k^M - O_k^M\right) g'\left(h_k^M\right) \tag{26}$$

The rate of change of error in hidden layer is,

$$\frac{\partial E}{\partial w_{ji}} = \sum_{Mk} \left(\frac{\partial E}{\partial V_j^M} \frac{\partial V_j^M}{\partial w_{ji}} \right)$$

$$\frac{\partial E}{\partial w_{ji}} = \sum_{Mk} \left(D_k^M - O_k^M\right) g'\left(h_k^M\right) \left(V_j^M\right) w_{kj} X_i^M \; g'\left(h_j^M\right) \tag{27}$$

$$\frac{\partial E}{\partial w_{ji}} = \delta_k^M w_{kj} X_i^M \; g'\left(h_j^M\right)$$

$$\frac{\partial E}{\partial w_{ji}} = \delta_j^M X_i^M$$

The propagation error or local gradient in hidden layer can be written as,

$$\delta_j^M = \sum_k \left(w_{kj} \delta_k^M\right) \; g'\left(h_j^M\right) \tag{28}$$

For a single epoch w_k is a vector of weights and biases, and $\Delta E(w_k)$ is the current gradient. Thus, the relationship of these parameters for an epoch can be written as (29),

$$w_{k+1} = \propto \; w_k - \eta \Delta E(w_k) \tag{29}$$

η is a significant parameter in artificial neural networks and is called the learning rate. It affects the learning capability of the ANN and an appropriate value is required to perform a satisfactory training. Learning rate η could be time dependent and adaptive. The learning rate is chosen based on the architecture of our system.

The α is the momentum term in (29), which is added to improve the speed of convergence of the network. We set the stopping criteria based on the epoch or network performance parameter and the mean square error.

The gradient decent is used to speed up the learning per epoch. The training process of a neural network must be tolerable so that problems like poor fitting or over-fitting will not arise. Due to over-fitting, ANN collapses all the generalization aptitude. In contrast, poor training results in scarce learning.

Then, the Euclidian distance can be calculated from (30),

$$D = \sqrt{((x_r - x_c)^2 + (y_r - y_c)^2)} \tag{30}$$

In (30), the x_r,y_r are the coordinates of each receiver and the x_c,y_c are coordinates of the wireless client.The system terminates the training phase when MSE value is equal to the threshold or the maximum epoch has been reached.

The cost function surface is such that steepest decent turns into local minima. Local minima can be affected by the learning method and network constitution. The gradient decent technique could get trapped into local minima of the cost function. The initial weight should be set to a small value; otherwise the sigmoid will saturate at the starting phase of learning.

Ultimately we have to consider the cost function for each of the training sets, so we choose batch mode for training. Batch mode of training goes through all sets of patterns and updates weight with completion of first epoch. Sequential mode of training update weights are updated pattern by pattern. Once all samples are done, it starts the next epoch. Sequential mode of training is a stochastic process with less possibility to trap in local minima. Batch mode trainings are more suitable to estimate the gradient decent. One of the great advantages of neural network is the parallelization.

When a primary try did not generate good results, minor improvements were sought by retraining the data. If retraining did not help, then hidden layer neuron size was increased to get better performance. Once the system is trained, the ANN is ready to work with testing data set.

3.3.4 Integrated Artificial Neural Network Architecture for WSN Localization

The complex, nonlinear relationship between input and targeted output is obtained through a unique transformation of the signal, and the fractional dimension is used as an input for the proposed integrated artificial neural network approach, as shown in Fig. 11. The input units consisted of an integer dimension representation and a fractional dimension representation of the Wi-Fi signal. There were two output units and 3~6 hidden units in the IANN architecture.

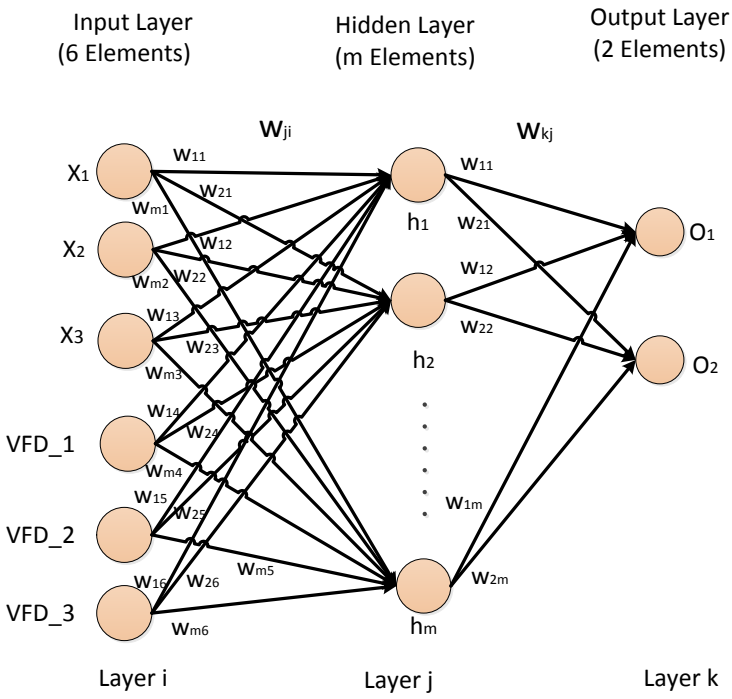


Fig. 11. IANN architecture.

The captured raw signal is preprocessed first. We removed the outliers and performed the necessary statistical analysis. When we found the captured signal is workable for the practical experiment, we initiated a multi-fractal analysis on the time domain signal. The nature of the signal suggests that the variance fractal dimension can best describe the signal. We obtained the distinctive trajectory for the signal. The trajectory extracts the features of the signal, which reveal the underlying complexity of the system. The VFDT obtained for a signal from specific co-ordinate position has a unique feature. The proposed IANN architecture utilized the captured RSS values from three of the receivers, along with the VFD for the signal as an input set to obtain the corresponding location of a wireless device (Fig. 12).

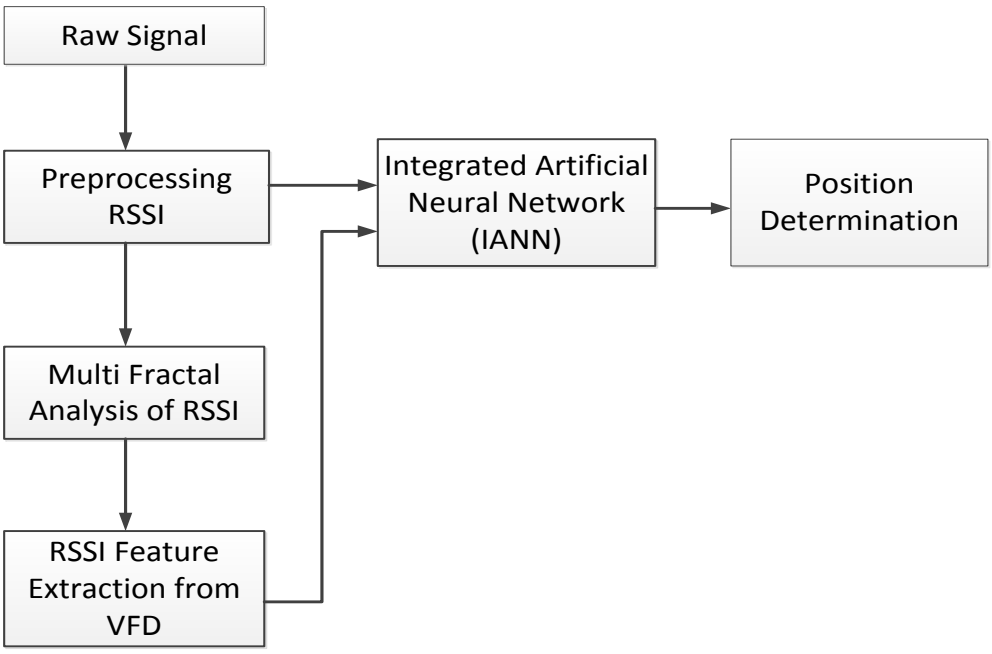


Fig. 12. Step by step procedure for proposed location determination algorithm.

Chapter 4

4. System Setup

4.1 System Setup Consideration

A cost effective, practical and feasible setup has been created and implemented. We ensured the availability of all required equipment. We carefully measured all the grid points. The minimum spacing (1m) between the wireless terminal and an anchor has been verified such that the receiver records unique RSS values from different grid points. Moreover minimum spacing (1m) between transmitter and receiver with respect to the ground has been ensured. Once the minimum spacing was determined, then the constant of proportionality for the mapping function which maps the RSS value to distance will be determined.

4.2 Data Collection

We collected data from both indoor and outdoor locations. Three different indoor locations were chosen to capture data for our experiment. We performed the indoor experiment using a subset of the deployed wireless infrastructure in the EITC building at University of Manitoba. The outdoor experiment was performed in the Smartpark area near the University of Manitoba.

4.2.1 Software Setup

The Comview application for Wi-Fi version 6.3 was installed separately in three laptops to provide the inputs for the conventional Trilateration and our ANN algorithms. Initially

we used Comview to capture wireless packets at the receiver at Indoor location 1. Comview filters the respective wireless channel from the transmitting access point and decodes packets. The Trilateration and ANN algorithms were run to determine the respective position of the wireless transmitter. The main disadvantage was that Comview does not provide precise RSS values. Comview provides dBm values, but they appear to be rounded to an integer, so it is hard to get differentiable and precise readings. There is a huge fluctuation and gap in captured RSS readings with respect to distance. It was difficult to classify and differentiate the captured RSS readings using Comview. Therefore, we needed a receiver that would provide high precision values of the received signal strength, in order to facilitate more accurate node localization.

We used three spectrum analyzers (SA) as receivers to capture RSS readings from the transmitter device. Spectrum analyzers capture a wide frequency range signal. The SA has good frequency stability and resolution. The SA supports both linear and logarithmic modes to display the amplitude of the signal in dB scale. Most importantly, SA provides precise values of received signal strength. We installed LABVIEW software in a laptop to capture data from a SA. We installed the National Instrument driver for Agilent N9320B Spectrum Analyzer to calibrate with LABVIEW. Fig. 13 shows the step by step data capturing procedure. RSS readings were obtained in the Fourier domain. We pulled the captured data using LABVIEW software.

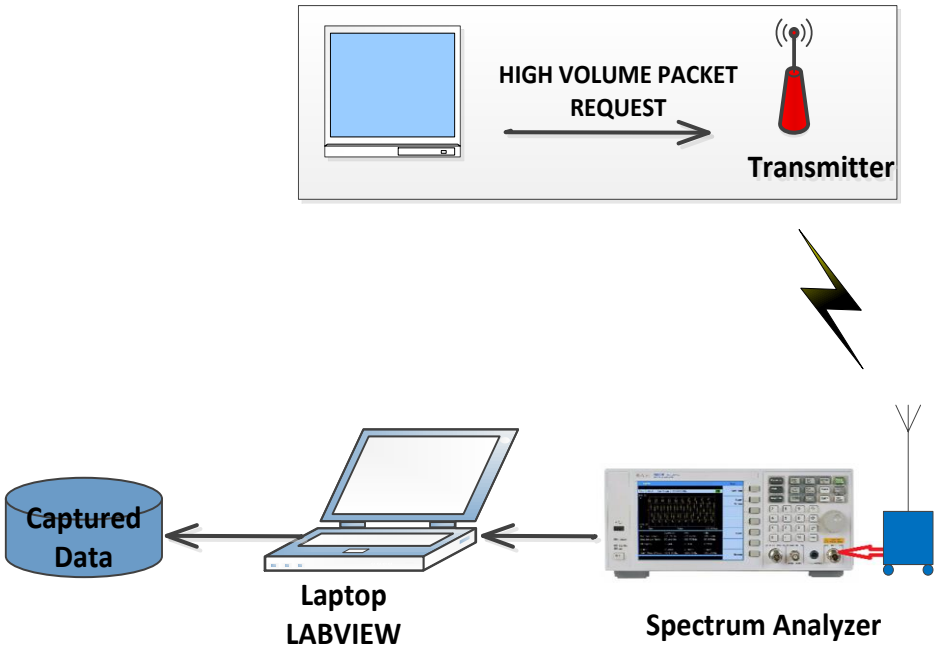


Fig. 13. Steps to capture the RSS for outdoor and indoor locations with a SA.

4.2.2 Hardware Setup

We used three SAs as receiver devices to capture RSS from a transmitter device. Our goal is to reduce the NLOS effect after analyzing the data and performing the IANN technique on three different sets of data to determine the actual location of the wireless device. All the required equipment is listed in Table 9.

4.2.2.1 Basic Hardware Setup at Transmitter End

Initially we used a movable desktop with a TRENDnet TEW-424UB USB wireless adapter as a transmitter. Then we used a D-Link wireless router as a transmitter to get better

results. We used the movable desktop transmitter only in the indoor location 1 (E2-229). For the rest of the experiments we used the D-Link wireless router as a transmitter device.

Table 9 Required equipment.

Required Equipment	
1.	3 sets of Agilent Spectrum Analyzer (Receiver device)-N9320B
2.	3 pcs of 2.4 GHz wireless antenna along with RF cables and SMA female-female adapters.
3.	1 set of wireless access point(Transmitter device)- D-Link, 2.4GHz Access Points
4.	3 laptops
5.	1 desktop
6.	TRENDnet TEW-424UB USB wireless adapter
7.	1 or 2 extended power cords
8.	3 fixed lab stools
9.	2 movable lab stools
10.	3 stands for hanging the antenna
11.	Measurement tape
12.	Pencil
13.	Grid point marker/ Tent pegs (25)
14.	Hammer
15.	2 Tables
16.	Power Supply-CP1000AVRLCD

Movable Desktop Transmitter:

First we used a movable desktop as our transmitter and sent continuous traffic from a command prompt: ping –l 6500 192.168.1.1. We connected the TRENDnet TEW-424UB USB wireless adapter to the transmitter device (desktop).

Wireless D-Link Router:

We generated high volume UDP traffic at the transmitter end to ensure acceptable SNR (See Fig. 20). We used a Java program (mentioned in Table 10) to simulate the continuous

UDP traffic. We configured the laptop with IP: 192.168.1.2 and sent continuous UDP traffic from the wireless socket.

Table 10 Java based simulator to generate continuous UDP packet.

UDP Packet Generation Algorithm	
1.	Initialize parameters:
	hostname: hostname to which packet will be sent; this will be given during program execution.
	buf: byte array of 512 bytes
	address: instance of the class InetAddress, which represents IP address of the host, the host's name has been given during program execution
2.	if (hostname=null)
	Return;
3.	Create a DatagramSocket socket for sending and receiving datagram packets.
4.	Create a 512 byte long DatagramPacket packet for sending packets to the port number 4445 on the specified host which has been given during program execution.
5.	while(true)
	socket.send(packet);

Fig. 14. Transmitter configuration.

After sending continuous UDP traffic to the transmitter, we observed a noticeable traffic change at the receiver end, as shown in Fig. 14. Then we configured the wireless

transmitter with IP: 192.168.1.1 and channel 6. We configured the SSID as Oceanview and broadcast the SSID to easily send the high volume UDP traffic towards it.

4.2.2.2 Basic Hardware Setup at Receiver End

We used the Comview software and a spectrum analyzer to capture RSS for separate experiments.

Comview:

Initially Comview for Wi-Fi version 6.3 was used as the network analyzer for capturing data with 802.11n standard at the receiver end. We configured specific rules to filter service port, traffic type, device IP, time period, time interval, units, etc. We scanned the existing wireless network and filtered the respected transmitter device with the required channel. The data capturing procedure is described in Fig. 15.

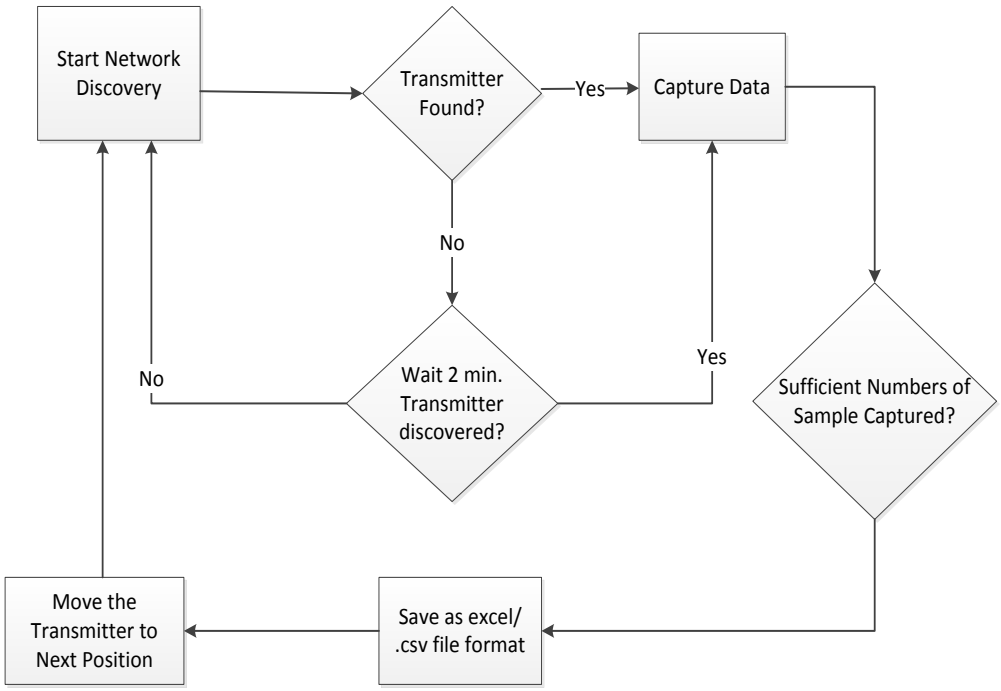


Fig. 15. Data capture procedure with Comview.

Then we opened the dashboard and ran the software to capture the wireless signal. We saved the file as .csv format to analyze the captured signal using Microsoft Excel. We captured the RSS readings from the Comview software, which were given in low precision integer dBm values.

Spectrum Analyzer:

We connected a 2.4 GHz wireless antenna in the RF IN 50 Ohm port of the Agilent N9320B spectrum analyzer, which is marked with a red arrow in Fig. 16.

Fig. 16. Agilent N9320B spectrum analyzer.

Fig. 17. 2.4 GHz wireless antenna.

RF cables and SMA female-female adapters were used to connect the antenna (Fig. 17) with the spectrum analyzer. We used the USB cable (Fig. 18) to connect the SA with the LABVIEW software, which was installed on a PC.

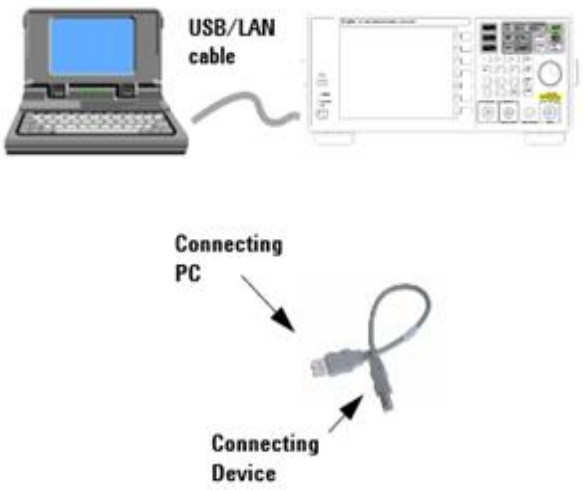


Fig. 18. SA connecting with LABVIEW installed on a PC.

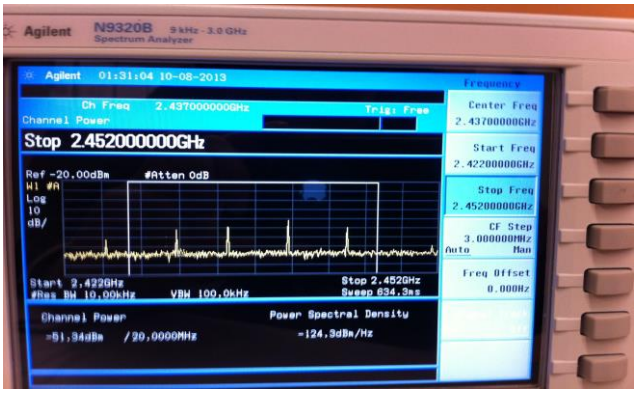


Fig. 19. SA capturing normal traffic from the transmitter.

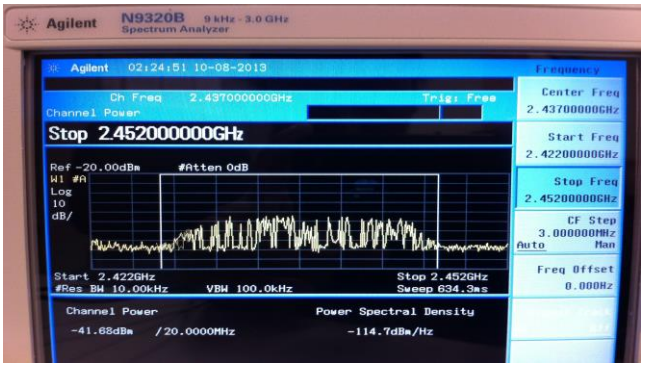


Fig. 20. SA capturing high volume traffic from the transmitter.

LABVIEW:

LABVIEW software has different VIs. We chose Agilent N9320 Acquire Trace VI to capture the signal. We configured the Agilent N9320Acquire Trace VI, as per our requirements shown in Fig. 21. We captured 500~15000 samples at a time using Labview software. We configured the center frequency for channel 6, 2.4GHz signal as 2.37GHz, and start frequency and stop frequency as 2.422GHz and 2.522GHz. We selected the respective USB port to establish the connection with SA.

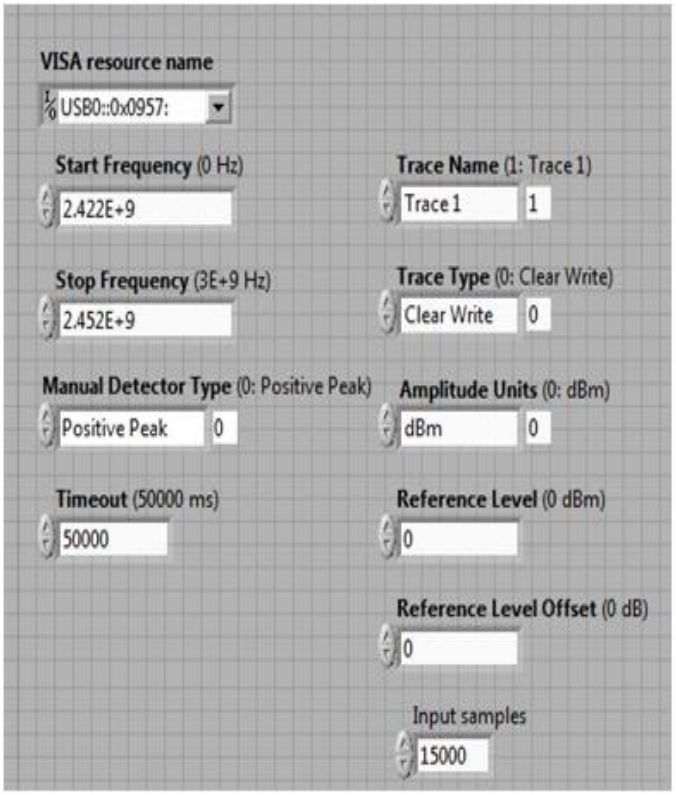


Fig. 21. Setting parameters in LABVIEW.

Table 11 The specifications of the spectrum analyzer.

Spectrum Analyzer Settings	
Mode	
Type:	Spectrum Analyzer

Frequency for Channel	
Center Frequency:	2.437 GHz
Start Frequency:	2.422 GHz
Stop Frequency:	2.522 GHz
Span	
Span:	30 MHz
Amplitude	
Attenuation:	0 dB
Scale Type:	log/linear
Scale/Div:	10
Ref Level(dBm):	-20
Measurement	
Integration BW:	20MHz
Channel Power Span:	30Mhz
Sweep	
Sweep:	Continuous
Sweep Time:	10ms(Auto)
BW/Avg	
Average:	100(ON)

Table 11 shows the SA parameter settings for our experiment.

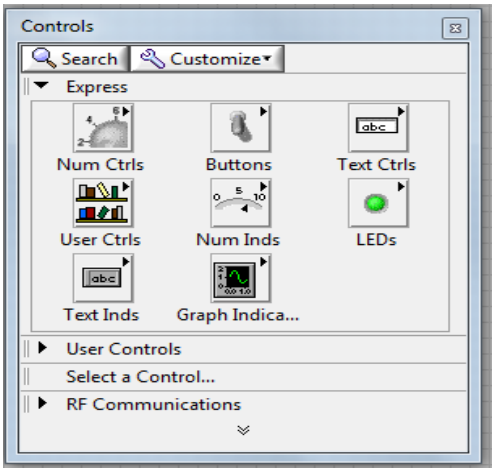


Fig. 22. Control button.

Using the control buttons (Fig. 22), we modified the block diagram. We captured a large number of samples (500~15000) for our experiment and saved the data in a folder (Fig. 23).

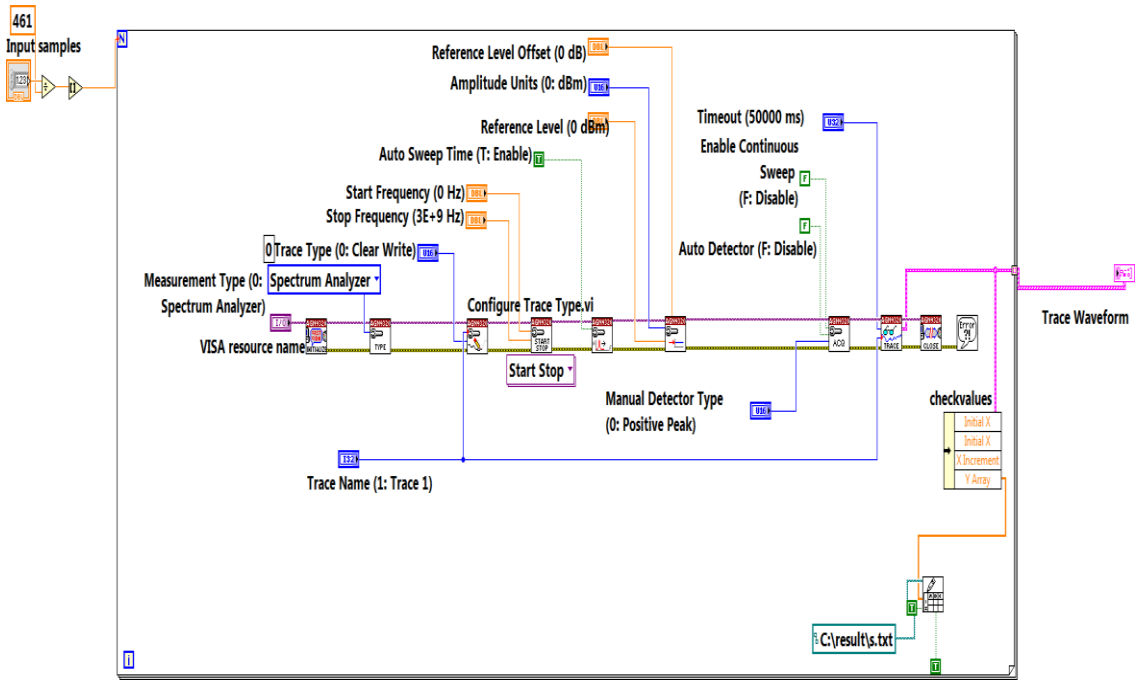


Fig. 23. Block diagram of the VI.

The trace waveform (Fig. 24) shows the amplitude vs. frequency plot for a captured signal.

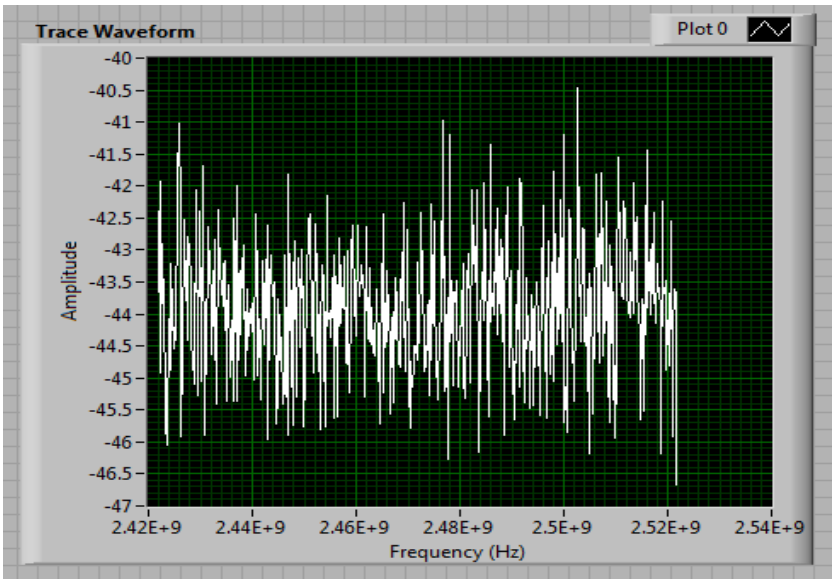


Fig. 24. Trace waveform.

4.2.2.3 Test Bed Setup in Indoor Location 1 (E2 229)

We chose an indoor location area which had average interference and obstacles to do the experiment. We considered 25 positions in a 12x12 grid in Room E2-229 at the EITC building. In order to determine the received signal strength between transmitters and receivers, we employed the log-distance path loss model. We analyzed the captured data and ran the Trilateration and ANN algorithms.

Software Support

Comview for Wi-Fi version 6.3 was used as the network analyzer for capturing data from 802.11n standard devices at the receiver end. We used the TRENDnet TEW-424UB USB wireless adapter in the transmitter device (desktop). It can easily support the 802.11g wireless networks with speeds of up to 54Mbps and transmit power of 15 dBm.

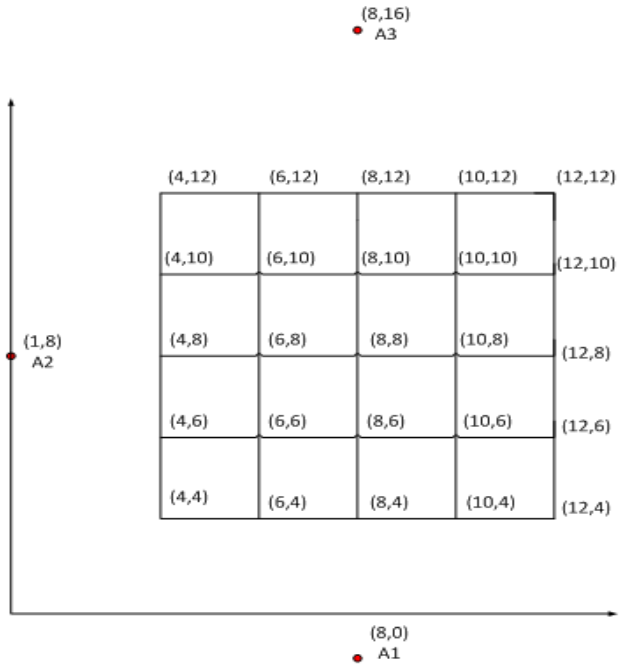


Fig. 25. 12m X 12m grid locations with three wireless receivers.

Hardware Setup

We installed Comview software on three laptops. These laptops were used as three wireless receiver devices to run the Trilateration and ANN algorithms in the protected area. We considered the minimum spacing between the wireless transmitter and receiver as at least 3m. For each possible grid point in the area 12m x 12m grid (Fig. 25), three anchors were strategically placed at A1(8,0), A2(1,8) and A3(8,16) positions around the grid area, and RSS values were recorded from all the 25 specified grid locations. Every grid point was located at least 1m apart from each other. Also, all wireless receivers were placed at about 1m above ground.

Data Collection

The data was collected using three receivers, specifically 10 times for all experimental cases from the EITC building, room E2-229. The wireless receivers were connected to the U of M Wi-Fi network, which operates at 2.4 Ghz range and RSS were captured in the units of dB using open source software ”Comview for Wi-Fi”. We configured the Comview software to filter the traffic by transmitter IP and Mac address.

After acceptable RSS values were obtained for 25 possible positions in the 12mx12m grid area, data was recorded from all determined grid locations. We placed the wireless transmitter in each of the grid positions, generated 50 ICMP pings with low delay response at each instance and captured the RSS values at the three receivers.

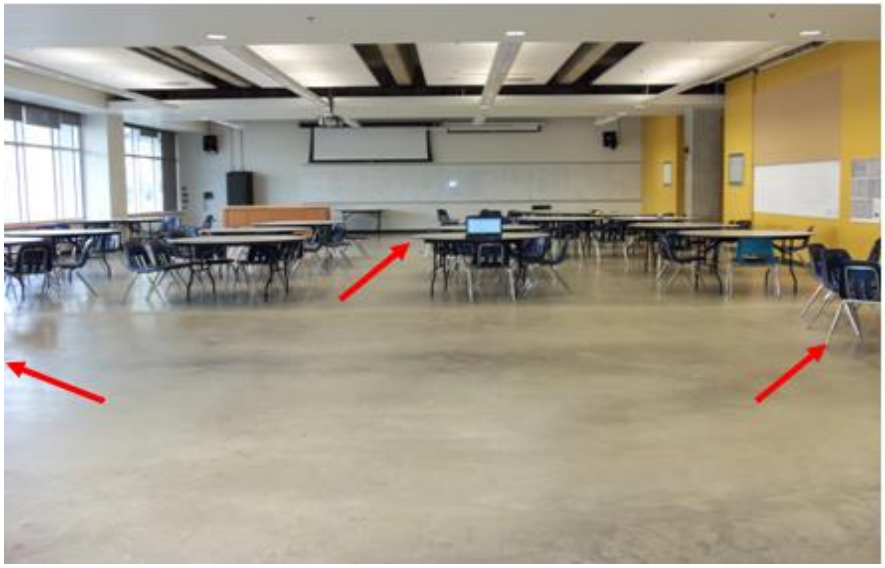


Fig. 26. Measurement location [Room E2-229] inside the EITC building.

For each position we initiated this command from command prompt:

```
ping -l 6500 192.168.1.1 -n 50
```

To locate a wireless transmitter using 2D trilateration, the sensing client must know the locations of three wireless receiver points, which are marked with a red arrow in Fig. 26, within their own coordinate system, and be able to measure distances from these three points. Trilateration calculation was performed against these receivers with known positions. We vacated the protected area and waited until capturing was completed for that position. Once we captured data for all positions, we saved the data and started the analysis. The attenuation rate for different indoor materials is shown in Table 12.

4.2.2.4 Test Bed Setup in Indoor Location 2 (E1 527)

We chose a typical indoor area which has average interference and obstacles (Fig. 27). We considered 5m x 5m grids of 25 positions in Room E1-527 at EITC Building. We also generated the synthetic data for this setup and tested ANN and IANN algorithms for both the synthetic and real data set.

Step by Step Procedure of Preparation:

1. Ensure a minimum distance of (1m) between the grid points.
2. Ensure a minimum distance of (1m) between the transmitter and the receiver.
3. Ensure a minimum distance of (1m) above ground for the transmitter and the receiver.
4. Create a 5mx5m grid, and place stickers at each grid location for identifying the position of the wireless transmitter (WT).
5. Perform the experiment in three phases: Receiver needs to be placed on a fixed lab stool. Receiver 1, Receiver 2 and Receiver 3 need to be placed on (5,0), (5,5) and (0,5) coordinates at 3 different phases of the experiment.

6. The transmitter was kept on a movable lab stool. The transmitter was configured with a unique SSID: Oceanview and IP: 192.168.1.1.
7. A laptop needs to be configured in the same subnet as the transmitter. Run the Java code to send continuous UDP traffic. The laptop needs to be placed on a movable lab stool such that it can move with the transmitter.
8. Another laptop needs to be placed near the receiver devices, so that it can pull data from the receivers.

Step by Step Procedure of Calibration:

1. Configure the transmitter with IP: 192.168.1.1 and SSID: Oceanview
2. Install the LABVIEW software in the laptop. Install required driver for Agilent N9320B device.
3. Configure the laptop with IP: 192.168.1.2. Check the connectivity from the laptop to the transmitter AP using command prompt: ping 192.168.1.1 and ensure traffic if shown on the SA screen (Fig. 19).
4. Run the java program to generate high volume UDP packets at the transmitter end.
   command prompt: cd C:\Users\kaisert\Desktop\UDP
   command prompt: java UDPClient 192.168.1.1
5. Place the spectrum analyzer at the (5,0) coordinate.
6. Connect the analyzer and the laptop for remote control using a USB Host connector**.**
7. Specify the required parameters in the Agilent spectrum analyzer, as per Table 11.
8. Place the wireless transmitter at point 1(0,0).

9. Leave the testing area and wait at least 2 min to allow human subjects to leave the testing area.
10. Start the Labview software; configure the front panel as per Fig. 21; and capture the readings.
11. Save data from the labview software for the respective point, e.g., point 1(0,0).
12. When Labview resumes, extract 15000 samples and paste them into one excel sheet. Save the file (e.g., for point 1 and receiver 1, save the file as AP1P1.xlsx.)
13. Place the transmitter on the next point and repeat step 8-12. Continue for other points {1, 2, 3, …, 25}.
14. Calculate the average RSS value for each of the 25 files captured.
    1. If the average RSS value for the files which are 1m apart distance points in the set {1, 2, 3, …, 25} differs by at least 1 dBm from the RSS value of all other files in the set {1, 2, 3, …, 25}, then note this set and end. Otherwise, go to Step 14.2
    2. Otherwise, continue in a similar fashion until the 1 dBm requirement has been met. Phase 1 is done here.
15. For phase 2, place receiver 2 at (5,5) co-ordinate position. Repeat steps 6-13.
16. For phase 3, place receiver 3 at at (0,5) co-ordinate position. Repeat steps 6-13.

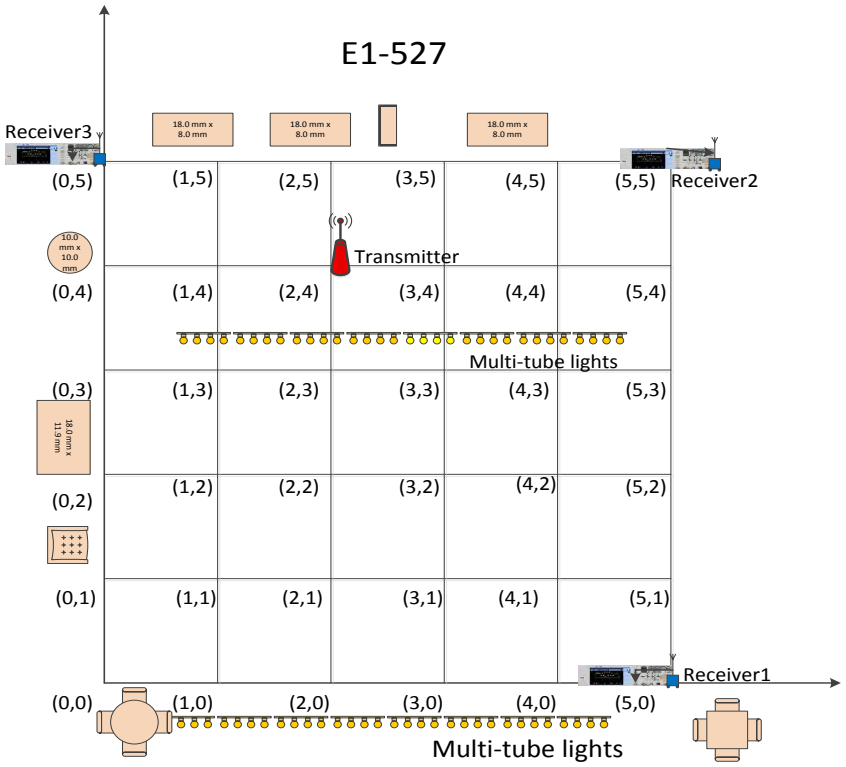


Fig. 27. Test grid location 2- E1 527, EITC building.

4.2.2.5 Test Bed Setup in Indoor Location 3 (E1 566A)

We choose a chamber room as our experiment location 3, which has minimum interference with other Wi-Fi networks (Fig. 28). We considered 5mx5m grids of 25 positions in Room E1-566A at the EITC Building. This room is surrounded by foam material to shield other EM signals from propagating into the room.

We also generated synthetic data for this setup, tested ANN and IANN algorithms for both the synthetic and real data set. The preparation and calibration procedure for location 3 is the same as location 2.

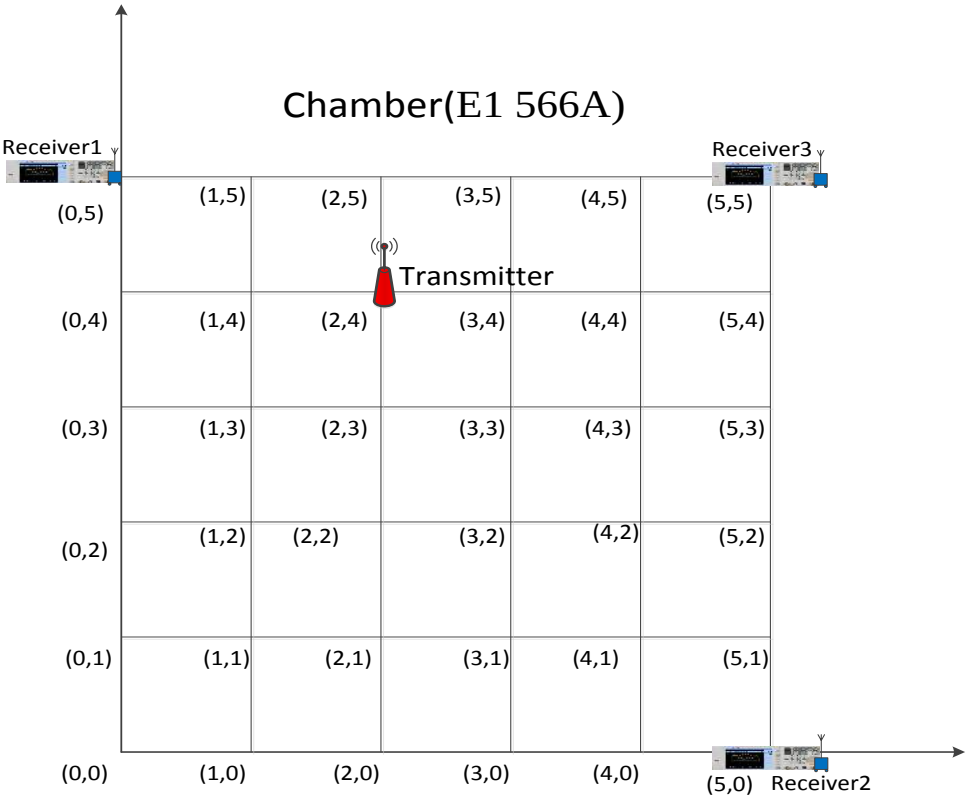


Fig. 28. Test grid location 3- E1 566A, EITC building.

4.2.2.6 Test Bed Setup in Outdoor Location

We located an outdoor field which is at least 100m away from any Wi-Fi infrastructure, at Smartpark, near the University of Manitoba. We placed a transmitter at positions 1-25 m away from the receiver. We checked the weather forecast before starting the test. We used an uninterruptable power supply for this setup. We placed tent pegs at each grid location for the wireless terminal (WT), as shown in Fig. 29. We placed a receiver (spectrum analyzer) at location (0,0), as shown in Fig. 30 and moved the wireless transmitter to the 25 different positions. Then we followed the same steps for preparation (1-3), (6-8) and calibration (1-4), (6-14) procedure, similar to what was done for location 2.

Fig. 29. Outdoor measurement setup at transmitter end.

Fig. 30. Outdoor measurement setup at receiver end.

4.3 Identifying Interference

We used Chanalyzer-4 network scanning to detect the presence of other signals in the experiment locations. The dashboard represents all existing signals observed within the respective frequency band 2.4 GHz. The intensity graph Fig. 31 shown the presence of high intensity signals in the experiment location.

The signals below were detected in the indoor experiment:

- Neighboring APs occupying an overlapping channel
- Cordless phones
- Microwave ovens
- Wireless audio transmitters
- Video transmitters (security cameras, IP cameras)
- PIR motion detectors
- WLAN downloads
- WLAN uploads

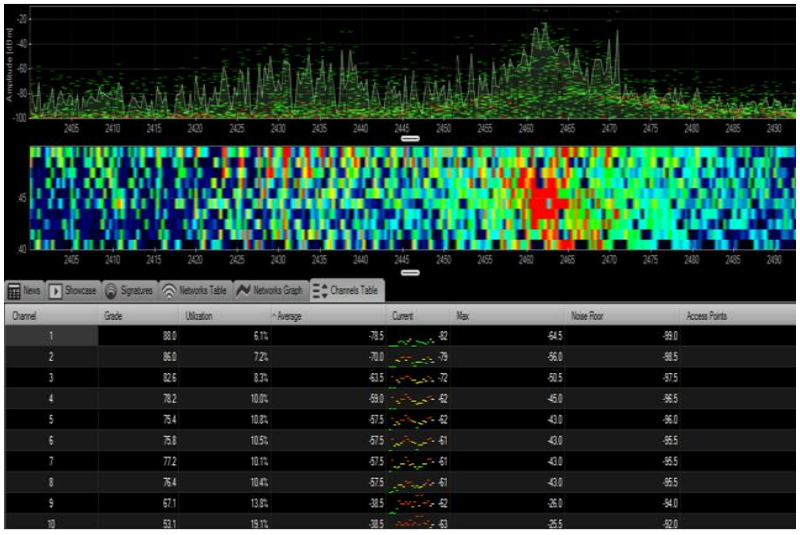


Fig. 31. Detection of other signals in E1-527, indoor location.

We observed that there are other networks operating in the frequency channel (1-12).

Fig. 32. Matching signatures of the captured signal at the experiment location.

4.4 Non Line of Sight Effect

The non-line-of-sight effect occurs when a Wi-Fi signal propagates through air and encounters natural or manmade obstacles. The signal loses strength due to the obstacles in the LOS path. Indoor propagation greatly suffers from the NLOS effect. Table 12 shows the attenuation rate of different indoor materials.

Table 12 Attenuation rate of different materials [54].

Material	Amount of Signal Attenuation
Concrete Wall	10-15 dBm
Metal Door	6 dBm
Brick Wall	8 dBm
Glass Wall with Metal Frame	6 dBm
Cinder Block Wall	4 dBm
Cubicle	3-5 dBm
Office Window	3 dBm
Wooden Door	3 dBm
Plasterboard Wall	3 dBm
Office Window	3 dBm
Glass window	2 dBm

We compared the synthetic data with our captured RSS readings. There was noticeable variation in captured RSS readings due to NLOS effect. Fig. 33 shows some of the environmental structures that contributed to the NLOS and interference problems in the indoor location 2 (Room E1-527).

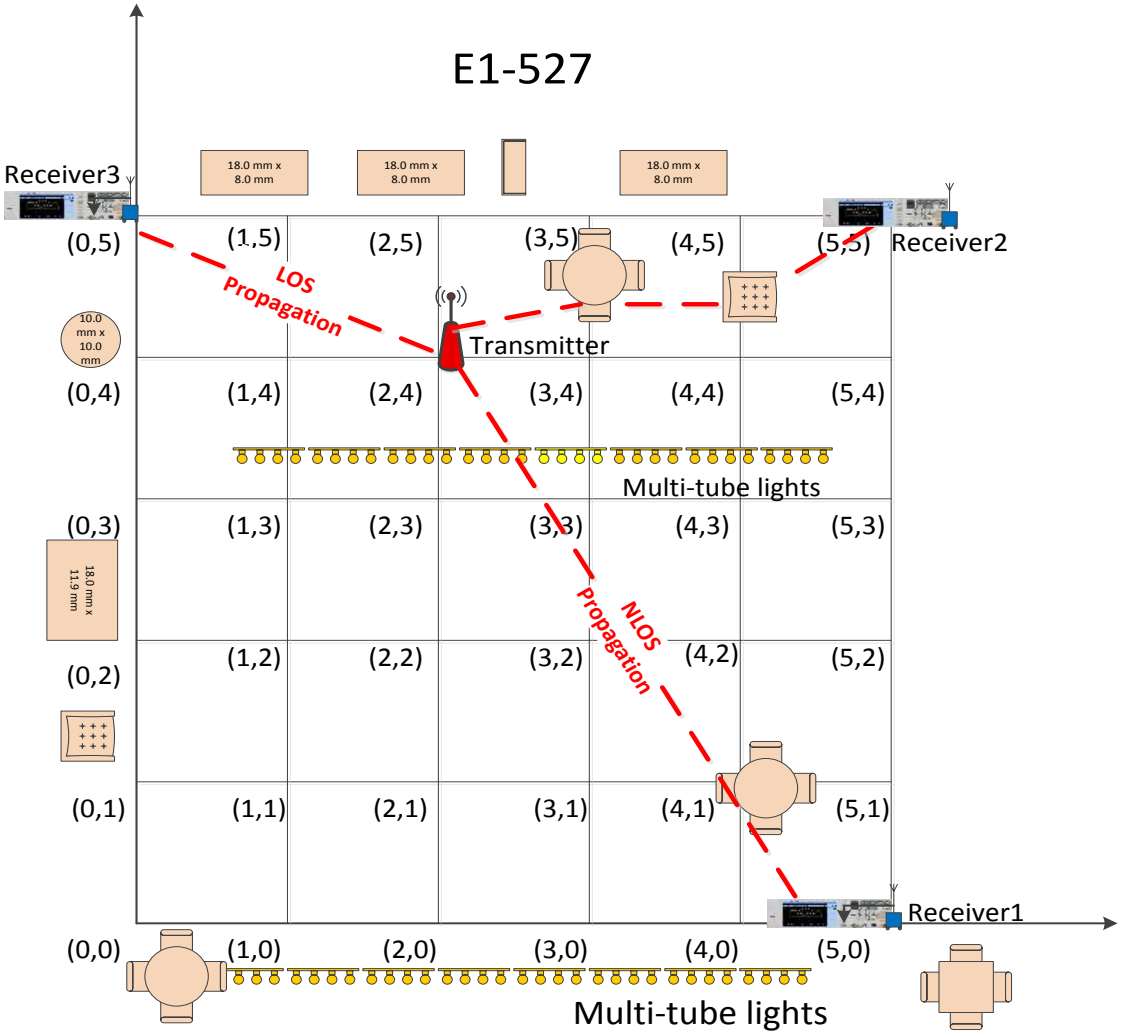


Fig. 33. NLOS and interference problems in the indoor location 2 (Room E1-527).

Chapter 5

5. Experiments and Results

Software was written in MATLAB to analyze the complexity of the captured receive signal strength and to simulate the proposed WSN localization algorithm to obtain their localization performance. The proposed algorithm, an ANN integrated with fractal preprocessing of the input signal, was simulated in MATLAB for WSN node localization. Three receivers measured their received signal strength of a signal transmitted by a stationary wireless node. The sequence of captured RSS triplets was transformed to a variance fractal dimension trajectory for analysis and presentation to the inputs of an ANN for localization. System verification and testing techniques were designed and implemented to verify correctness of the MATLAB simulations. This chapter describes each experiment and presents and analyzes the results.

5.1 System Verification and Testing

It is very important to ensure that the software developed to analyze and measure signal properties is properly designed and performs as intended. Verification testing of the software was done to verify the software was working as intended.

5.1.1 Received Signal Strength Indicator

A Java based simulator was developed to generate continuous high volume traffic at the transmitter end, so that a continuous signal may be captured (Fig. 20). Minimum acceptable distance between transmitter and receiver needs to ensure for the testing.

Minimum distance from the ground needs to be maintained. All measurements need to be taken carefully. Proper frequency channel needs to be established to avoid interference with an existing Wi-Fi network. We validated the captured readings using calibration Step 14 on Page 68.

5.1.2 Spectrum Analyzer

We generated a sine wave from a signal generator and validated the signal from an oscilloscope. We connected the signal generator to a spectrum analyzer (see Fig. 34) and observed the Fourier domain signal of the generated sine wave. We observed a spike at the generated frequency of the sine wave, and so verified the equipment setting and configuration. Since the spectrum analyzer could not be configured to output Fourier data, we could not perform an inverse transform to verify the waveform.

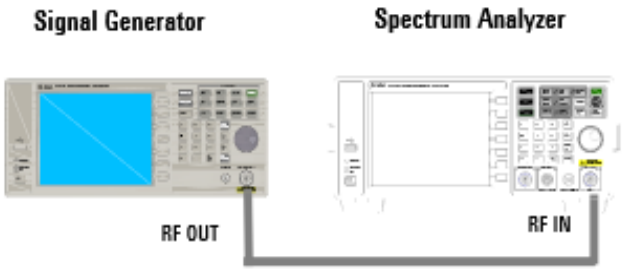


Fig. 34. Sine wave generation for the spectrum analyzer.

5.1.3 Artificial Neural Network

We verified our ANN model with synthetic data, which was generated using an indoor path loss model in [52] . The model yielded perfect classification rate for the synthetic data

set, i.e., there were zero classification errors. We calculated RSS values based on the indoor path loss model. Twenty-five samples of input-output pairs were fed to the ANN. We generated the RSS values for a 5mx5m grid and fed the ANN with the synthetic data set. We used 20% of the data set as testing samples. We randomly partitioned 20% of data as testing data set for each time of training. The remaining 80% of the data was used as the training set. The MSE obtained in the training phase was 0.0138. Table 13 shows the parameters and the results of the test.

Table 13 Results for synthetic data sets.

Performance Evaluation for Synthetic Data Set	
Architecture(I-H-O)	3-3-2
Epochs	1000
Iteration	60000
No of Samples	25
Momentum	0.5
Learning Rate	0.2
Training Data set	80%
Testing Data Set	20%
Data Selection per Iteration	Random
Error count for Testing	0
Training MMSE	0.0138
Testing MMSE	0.0281
Variance Training MSE	$9.25X10^{-6}$
Variance Testing MSE	$6.36X10^{-4}$

5.1.4 Variance Fractal Dimension Trajectory (VFDT)

First, the software implementation of the VFD algorithm was tested on white Gaussian noise. We know theoretically that white Gaussian noise should have a fractal dimension of 2. The software implementation of the VFD algorithm was 2.00047 (see Fig. 35).

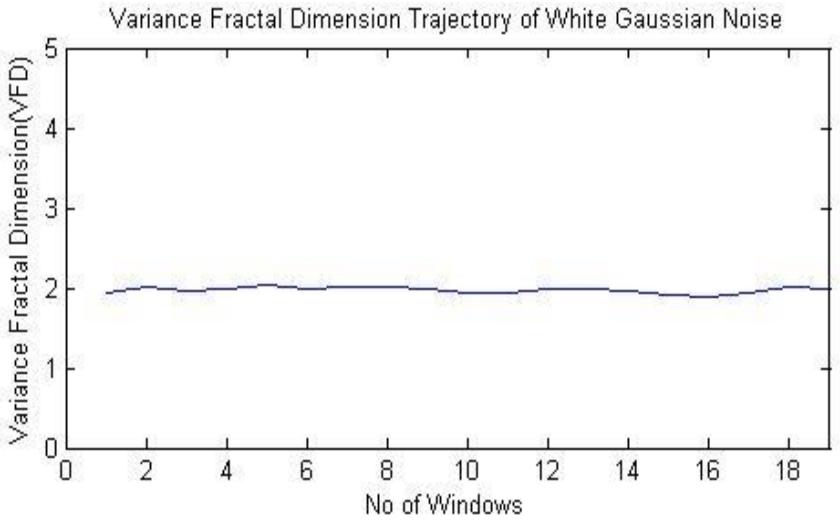


Fig. 35. Variance fractal dimension trajectory for white Gaussian noise.

We also observed the distribution of generated white Gaussian noise. As shown in Fig. 36, the histogram indicates a Gaussian noise with 0 mean.

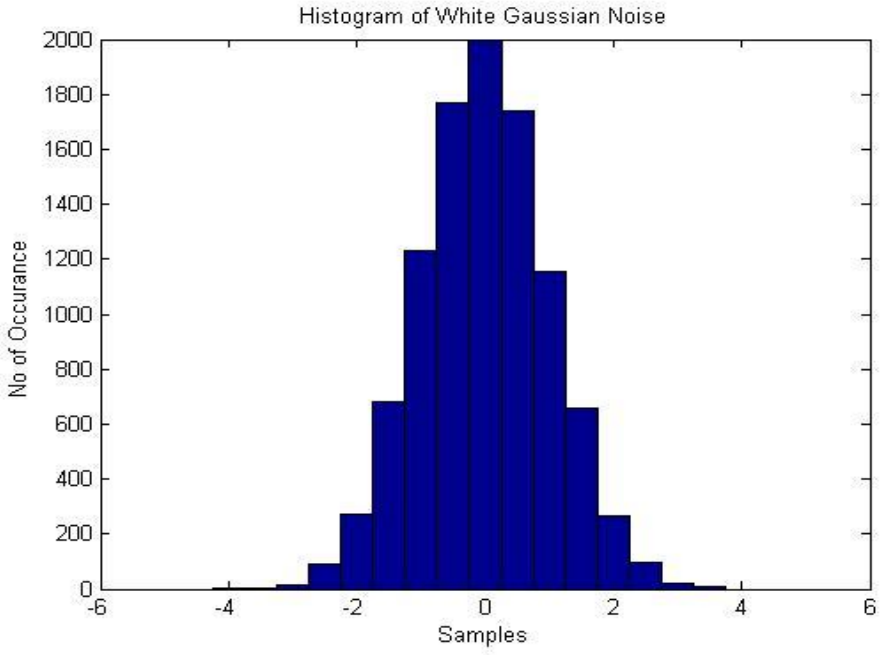


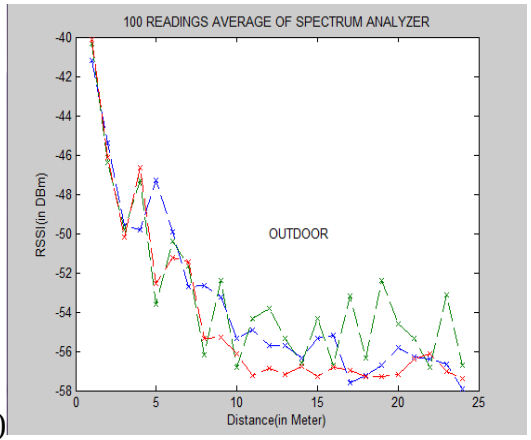
Fig. 36. Histogram of white Gaussian noise.

5.2 RSS Measurements with Different Data Sets

5.2.1 Outdoor Experiments

We chose to perform outdoor experiments so that we can capture the signal without an NLOS effect. The outdoor readings were obtained using a SA. Average values of 100 samples are plotted in Fig. 37. Average values obtained from Labview for 500 samples are plotted in Fig. 37(b). As shown, the average readings obtained from Labview are smoother than those obtained from the SA. We took a higher number of samples using Labview. Using the SA, we manually took the readings and reset the device for the next grid point.

(a)



(b)

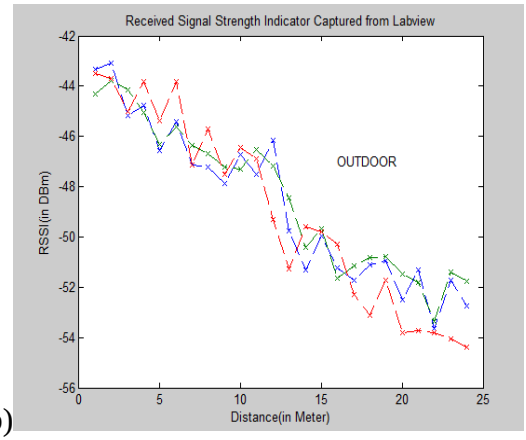


Fig. 37. Plot of received signal strength vs. distance.

Observations:

We can see that the RSS captured from an outdoor location are smoother than those from an indoor location. The signals obtained from an indoor location suffer from refection, attenuation and NLOS effects. However, outdoor RSS readings still fluctuated, and this may be due to antenna orientation. We obtained RSS readings for 500 samples

using Labview and obtained an average. Fig. 38 shows that the average RSS obtained from larger samples gives better results. In Table 14, we show the captured readings from Labview and SA.

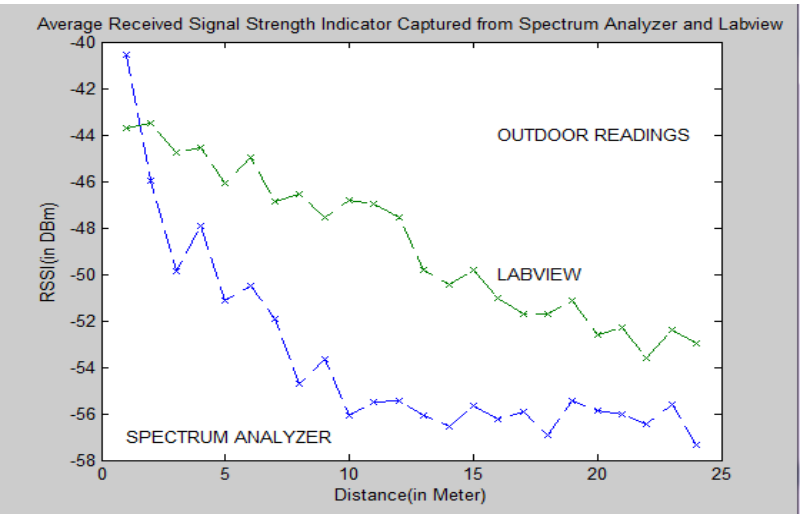


Fig. 38. Average received signal strength vs. distance in outdoor location.

Table 14 Outdoor RSS Readings.

Frequency = 2.437 GHz										
OUTDOOR RSS READINGS										
RECEIVER POSITIONS(0,0)			SPECTRUM ANALYZER				LABVIEW			
Transmitter Positions	Distance Between RX and TX (in Meters)=sqrt((x-x1)^2 +(y-y1)^2)	Average RSS (DBm)	RSS_1 (DBm)	RSS_2 (DBm)	RSS_3 (DBm)	Average RSS value from SA (DBm)	RSS_4 (DBm)	RSS_5 (DBm)	RSS_6 (DBm)	Average RSS value from Labview (DBm)
(0,1)	1	42.14	-41.2	-40.15	-40.36	-40.57	-44.33	-43.51	-43.31	-43.7167
(0,2)	2	-44.75	-45.4	-46.12	-46.42	-45.98	-43.8	-43.68	-43.08	-43.52
(0,3)	3	-47.33	-49.61	-50.2	-49.82	-49.87667	-44.15	-45.02	-45.16	-44.7767
(0,4)	4	-46.24	-49.8	-46.68	-47.31	-47.93	-45.07	-43.81	-44.76	-44.5467
(0,5)	5	-48.62	-47.3	-52.56	-53.61	-51.15667	-46.31	-45.38	-46.55	-46.08
(0,6)	6	-47.73	-49.9	-51.22	-50.38	-50.5	-45.62	-43.84	-45.43	-44.9633
(0,7)	7	-49.4	-52.7	-51.45	-51.66	-51.93667	-46.34	-47.14	-47.12	-46.8667
(0,8)	8	-50.63	-52.68	-55.32	-56.21	-54.73667	-46.67	-45.69	-47.23	-46.53
(0,9)	9	-50.59	-53.26	-55.3	-52.39	-53.65	-47.23	-47.5	-47.88	-47.5367

(0,10)	10	-51.46	-55.33	-56.11	-56.83	-56.09	-47.3	-46.45	-46.72	-46.8233
(0,11)	11	-51.24	-54.92	-57.25	-54.32	-55.49667	-46.52	-46.9	-47.52	-46.98
(0,12)	12	-51.51	-55.71	-56.88	-53.81	-55.46667	-47.18	-49.3	-46.15	-47.5433
(0,13)	13	-52.95	-55.69	-57.18	-55.36	-56.07667	-48.46	-51.27	-49.76	-49.83
(0,14)	14	-53.5	-56.32	-56.78	-56.6	-56.56667	-50.41	-49.58	-51.33	-50.44
(0,15)	15	-52.73	-55.36	-57.27	-54.32	-55.65	-49.67	-49.78	-49.97	-49.8067
(0,16)	16	-53.65	-55.21	-56.84	-56.71	-56.25333	-51.65	-50.27	-51.23	-51.05
(0,17)	17	-53.83	-57.63	-56.96	-53.21	-55.93333	-51.15	-52.3	-51.73	-51.7267
(0,18)	18	-54.32	-57.22	-57.27	-56.34	-56.94333	-50.84	-53.11	-51.12	-51.69
(0,19)	19	-53.31	-56.69	-57.28	-52.41	-55.46	-50.8	-51.72	-50.94	-51.1533
(0,20)	20	-54.23	-55.82	-57.18	-54.6	-55.86667	-51.47	-53.82	-52.5	-52.5967
(0,21)	21	-54.15	-56.31	-56.38	-55.32	-56.00333	-51.8	-53.74	-51.32	-52.2867
(0,22)	22	-55.03	-56.39	-56.13	-56.82	-56.44667	-53.38	-53.79	-53.64	-53.6033
(0,23)	23	-54.01	-56.68	-57.05	-53.11	-55.61333	-51.41	-54.07	-51.74	-52.4067
(0,24)	24	-55.15	-57.92	-57.41	-56.7	-57.34333	-51.77	-54.37	-52.73	-52.9567

5.2.2 Indoor Location 1: E2-229

We used Comview PC software in the emulated wireless receivers to capture the RSS. This location E2-229 had an average level of interference and NLOS effects. Table 15 shows the average RSS readings. There was fluctuation in the average RSS plots, as shown in Fig. 39. We ran trilateration and ANN algorithms on these readings. We were not able to obtain satisfactory results, probably because Comview does not provide sufficient resolution in the readings. We noted that the Comview readings were always in integer dBm format. We summarized that the given resolution was inadequate to achieve sufficient accuracy of the experiment.

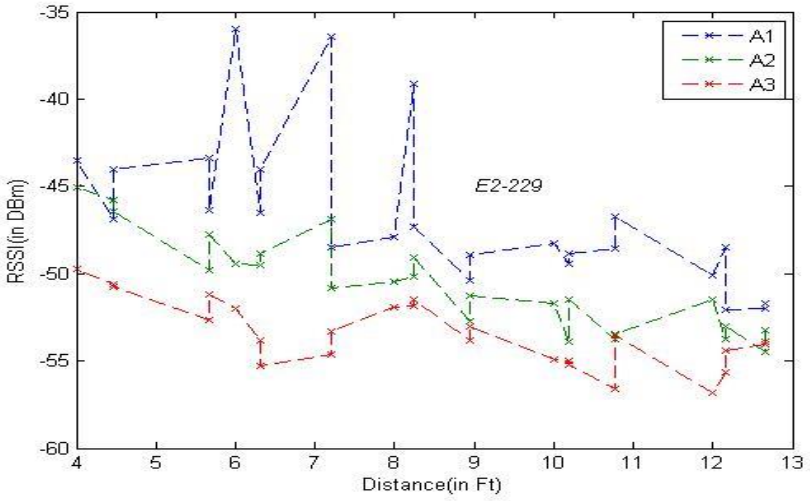


Fig. 39. Distance vs. average value of RSS captured with Comview in Rm. E2-229.

Table 15 Average RSS readings with corresponding distances from three anchors.

Receiver 1		Receiver 2		Receiver 3	
Average (RSS)	Distance (in meter)	Average (RSS)	Distance (in meter)	Average (RSS)	Distance (in meter)
-43.393	5.66	-47.753	5.67	-54.044	12.65
-46.86	4.47	-46.891	7.21	-55.662	12.17
-43.492	4	-52.71	8.94	-56.794	12
-44.062	4.47	-51.464	10.77	-54.409	12.17
-46.341	5.66	-54.495	12.65	-53.923	12.65
-36.419	7.21	-51.511	12.16	-56.611	10.77
-46.532	6.32	-53.899	10.2	-55.004	10.2
-35.991	6	-50.179	8.25	-54.952	10
-44.04	6.32	-49.534	6.32	-55.187	10.2
-48.476	7.21	-45.812	4.24	-53.522	10.77
-50.418	8.94	-45.019	4	-53.836	8.94
-39.156	8.25	-49.468	6	-51.842	8.25
-47.901	8	-50.464	8	-51.93	8
-47.343	8.25	-51.712	10	-51.496	8.25
-48.942	8.94	-53.436	12	-52.984	8.94
-48.582	10.77	-53.724	12.17	-54.657	7.21
-49.413	10.2	-53.025	12.33	-53.809	6.32
-48.246	10	-49.099	8.25	-52.011	6

-48.834	10.2	-48.855	6.32	-55.291	6.32
-46.707	10.77	-46.435	4.47	-53.305	7.21
-52.005	12.65	-49.827	5.66	-52.65	5.66
-48.502	12.16	-50.791	7.21	-50.572	4.47
-50.076	12	-51.237	8.94	-49.758	4
-52.05	12.17	-53.763	10.77	-50.75	4.47
-51.724	12.65	-53.239	12.65	-51.189	5.66

5.2.3 Indoor Location 2: E1-527

E1-527 is a densely furnished room with the presence of NLOS effects and other interfering signals. We used a SA as a receiver to capture the RSS in this location. We found that, although the SA provided higher precision than Comview, still the readings were fluctuating. We took an average of 500 samples and plotted the average RSS with respect to distance. We noticed that few of the points suffered from signal attenuation and reflection due to the furnished environment (Fig. 40).

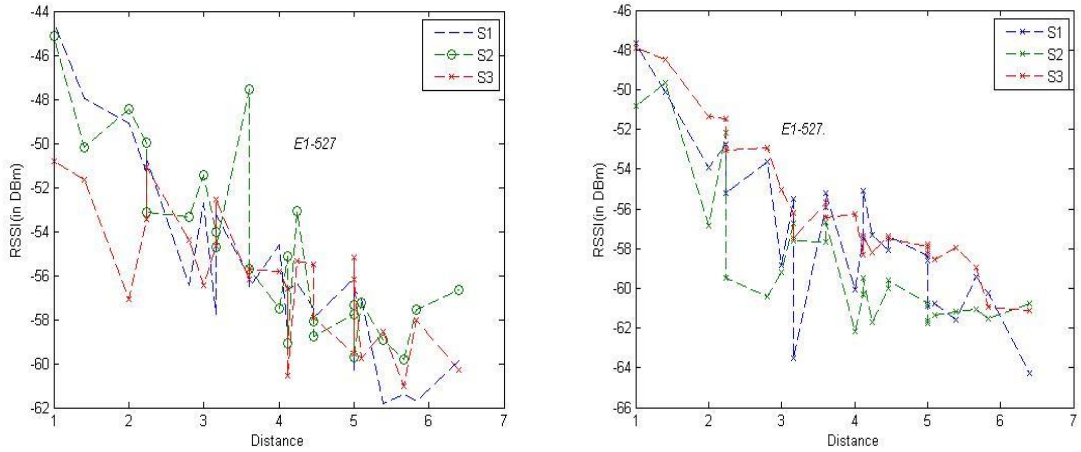


Fig. 40. Distance vs. average value of RSS captured in Rm. E1-527.

5.2.4 Indoor Location 3: E1-566A

The E1-566A room is internally surrounded by electromagnetic absorption material. This room has minimum interference with other signals. The RSS readings observed in this location appeared to be more stable in Fig. 41 than the other locations.

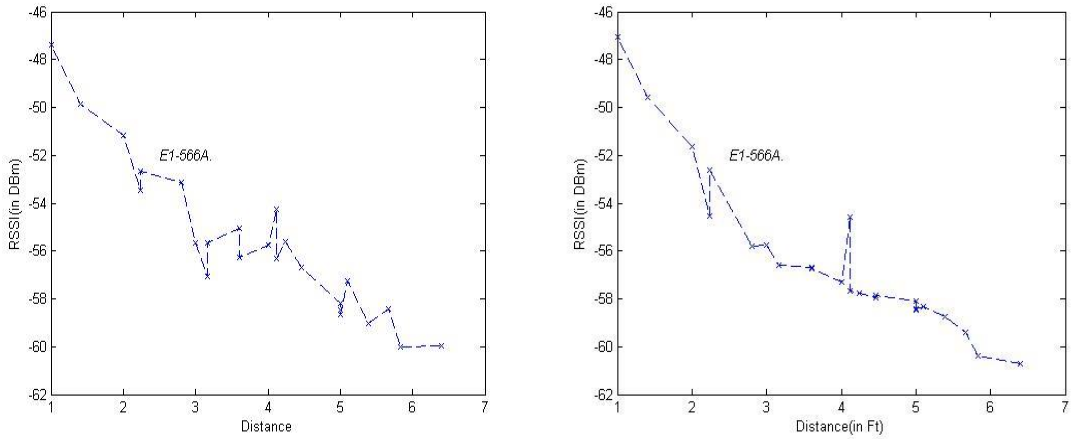


Fig. 41. Distance vs. average value of RSS captured in Rm. E1-566A.

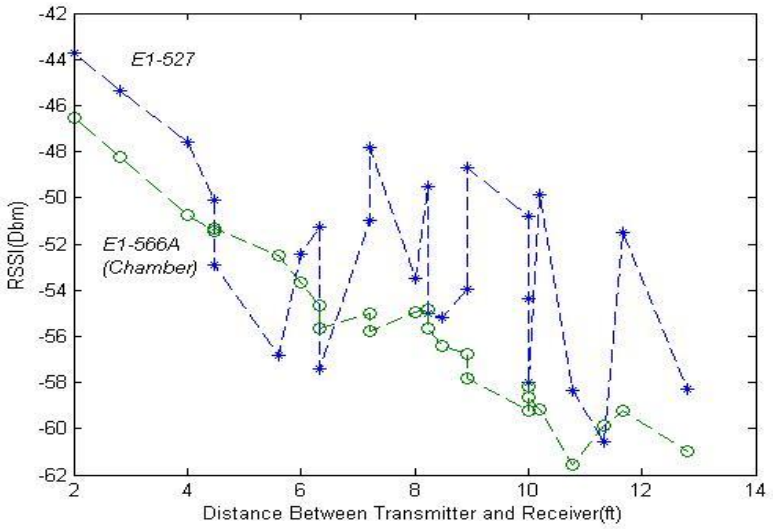


Fig. 42. Comparing RSS readings for different locations.

Observation:

We chose three locations such that two locations suffer from average interference: E2-229 was the least densely furnished environment and E1-527 was the most densely furnished environment. The third location (E1-566A) suffered from less or a minimum of interference. Fig. 42 shows the impact of interference on the RSS. The RSS readings were contaminated by inferences from other signals in E1-527. However, readings are more stable in the chamber (E1-566A).

5.3 Statistical Analysis

A statistical analysis of a signal may reveal underlying properties of a signal. We placed the transmitter and the receiver at a fixed position in a grid. We captured the received signal strength for 10,000 samples. Fig. 43 shows the time series representation of the received signal strength (RSS).

5.3.1 Distribution of Typical RSS

We plotted the captured signal with respect to time. We noticed the fluctuation in the captured signal. The possible reason of fluctuation could be any of refection, refraction or attenuation from obstacles. We computed the mean, variance, standard deviation, error bar, histogram and probability mass function of the signal.

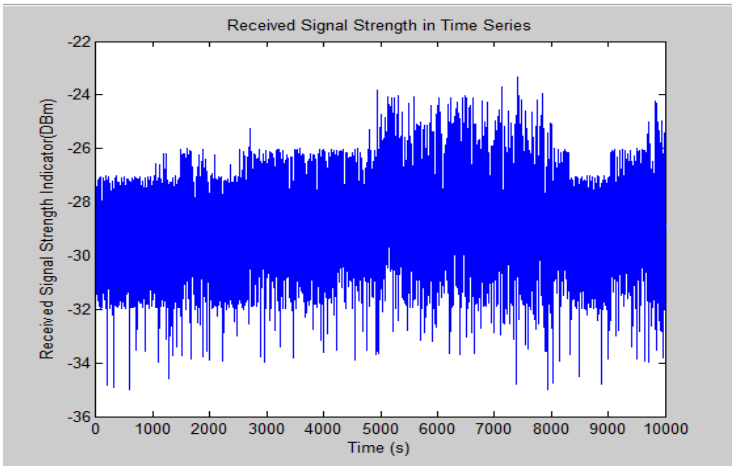


Fig. 43. Received signal strength in time series.

We obtained the histogram considering bin size of 100 to see the RSS distribution for number of occurrences shown in Fig. 44.

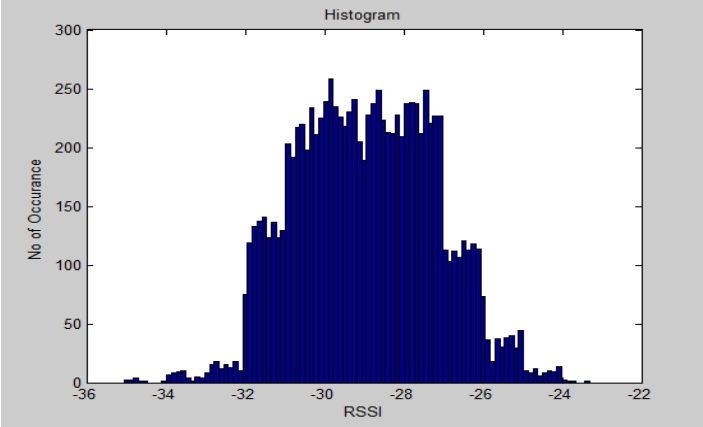


Fig. 44. Histogram of received signal strength indicator (10000 samples).

By normalizing the histogram, the probability mass function was obtained, which is a discrete distribution. Fig. 45 shows that there is a rich variation in the probability distribution. It is stochastic in nature, semi-bell shaped curve, and the signal is non periodic.

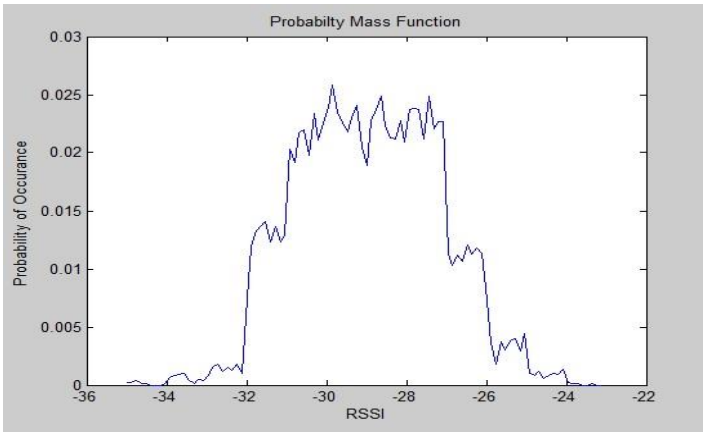


Fig. 45. Probability mass function of received signal strength indicator.

The error bars in Fig. 46 show the standard deviation of the mean. The mean varies noticeably with the number of samples. Thus, it is clear that the signal is non-Gaussian, non-stationary and in general has complex features.

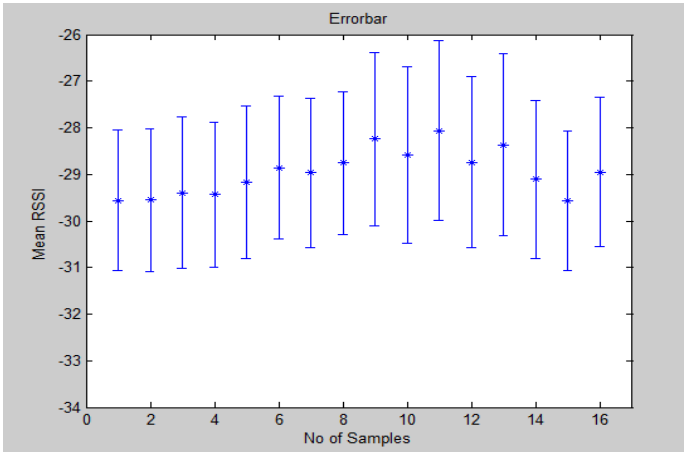


Fig. 46. Error bar for mean value of RSS.

5.3.2 Stationarity analysis

Stationarity analysis can be used to demonstrate how the statistical properties of a signal vary with time. We considered a sliding window of 400, 500 and 600. We found that for a sliding window of 600, the signal achieved 95% of confidence interval for 1st moment.

Fig. 47 shows the mean (1st moment) for a sliding window size of 600. The mean values were obtained within a 95% confidence interval. Fig. 48 shows the variance (2nd moment) for a sliding window size of 600. The variance is highly fluctuating over the given range, considering a 95% confidence interval. Hence, the RSS is not weak sense stationary (WSS) in a whole window. Thus multi-fractal analysis is not possible for the whole window and we chose VFD to obtain the complexity of the signal.

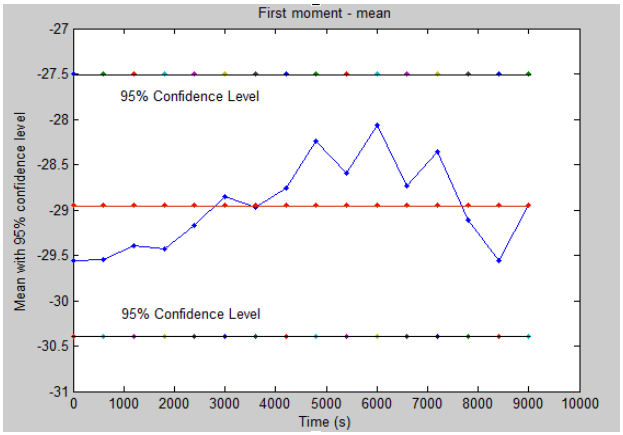


Fig. 47. First moment (mean) with 95% confidence interval.

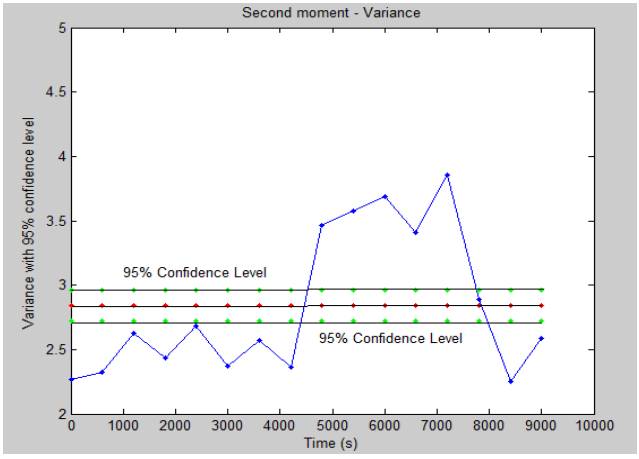


Fig. 48. Second moment (variance) with 95% confidence interval.

Skewness is a measure of regularity of the flow of the time series. Skewness determines the symmetry or asymmetry of a signal. Kurtosis is a measurement that compares the signal with its normal distribution. Kurtosis classifies if the signal is saturated or peaked. Fig. 49 and Fig. 50 shows the variability of third moment and forth moment with time. Both skewness and kurtosis changed with time intervals.

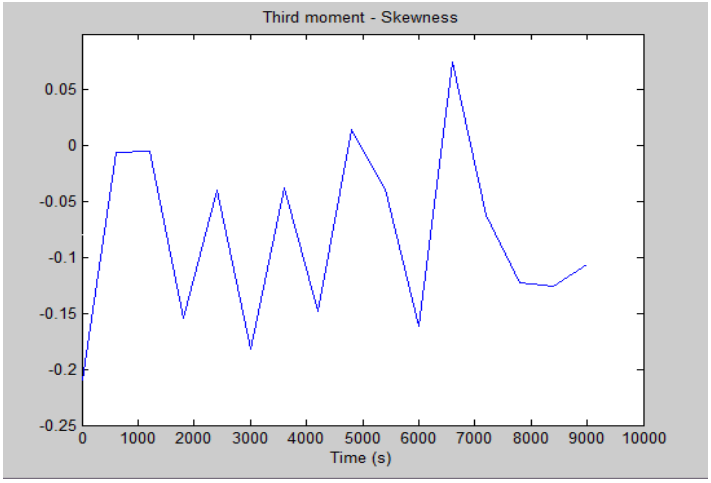


Fig. 49. Third moment (skewness).

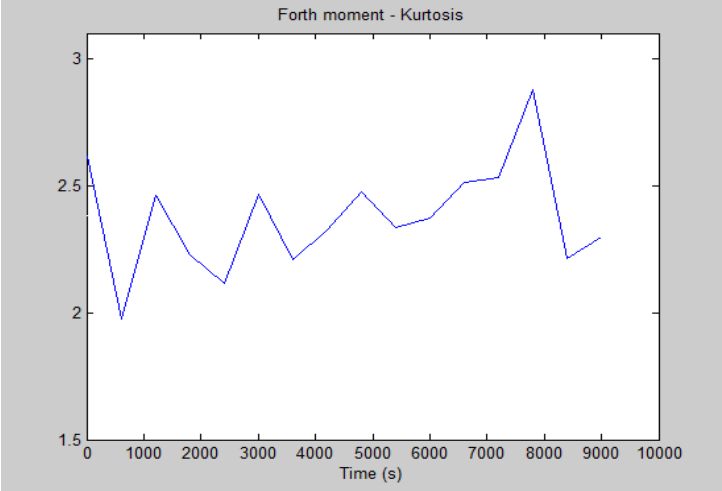


Fig. 50. Forth moment (kurtosis).

5.4 Multi-fractal Analysis of Wi-Fi Signal

Multi-fractal analysis of the stochastic, non-stationary Wi-Fi signal requires obtaining the power-law relationship in the log-log plot for RSS, as shown in Fig. 51.

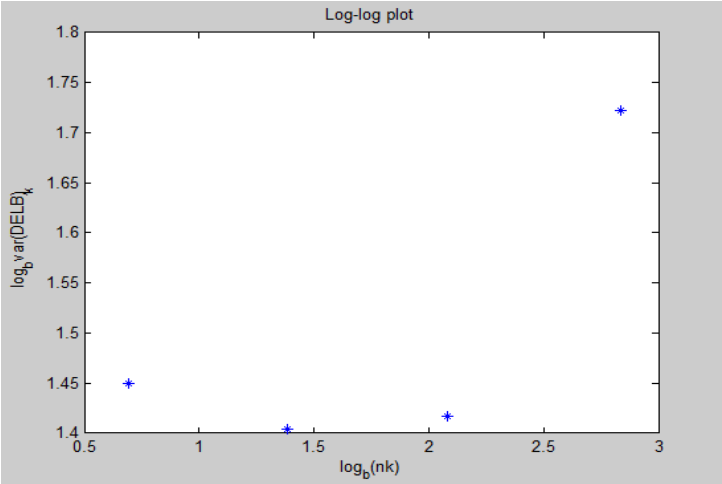


Fig. 51. Log-log plot for RSS.

All the required parameters for VFDT algorithm are listed in Table 16.

Table 16 Parameters for VFDT Algorithm.

VFDT Algorithm Parameters	
Symbol	**Values**
N	512
d	32
N_{min}	$N_1 > 30$
K_{low}	$K_{low} \geq 1$
E	1
n_k	2,4,6,….

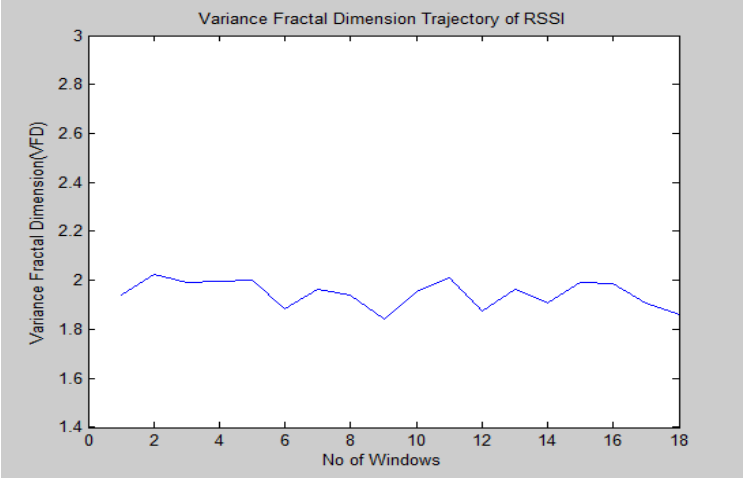


Fig. 52. Variance fractal dimension trajectory for RSS.

Fig. 52 shows how the VFD of the received signal strength changes over time. For a monofractal signal, the same complexity would be obtained everywhere, and the plot of the VFD would be a flat line. Thus, Fig. 52 shows that the RSS is a multi-fractal signal, i.e., a mixture of monofractals. Moreover, values of VFDT are normalized dimensions between 1 and 2, which is required to identify the unique signal. The global fractal dimension obtained was 1.94 for the given RSS signal. Most importantly, we observed that VFDT reflected the same trend as the time series plot of the RSS. In the initial period there was saturation in the time series signal. The variance fractal dimension trajectory remained stable at the beginning. Onwards both plots show fluctuation and irregularity. Thus, this provides motivation and rational for us to use the VFD of the RSS signal as an input in our proposed ANN model.

We obtained the stationarity analysis and VFDT for captured RSS for 25 positions in 5mX5m grid. The detail results of global fractal dimensions are shown in Appendix B.

5.5 RSS Based Triangulation Technique for Localization

A MATLAB simulation was created to determine the location of the wireless device using the trilateration method for indoor Location-1 (E2-229). We captured the RSS readings from three of the wireless receivers and analyzed the data. We captured 250 samples for each grid position. We repeated the following process a total of 10 times. We performed statistical analysis of the readings. We calculated the mean, standard deviation, maximum and minimum of RSS values for each position. We discarded the outliers.

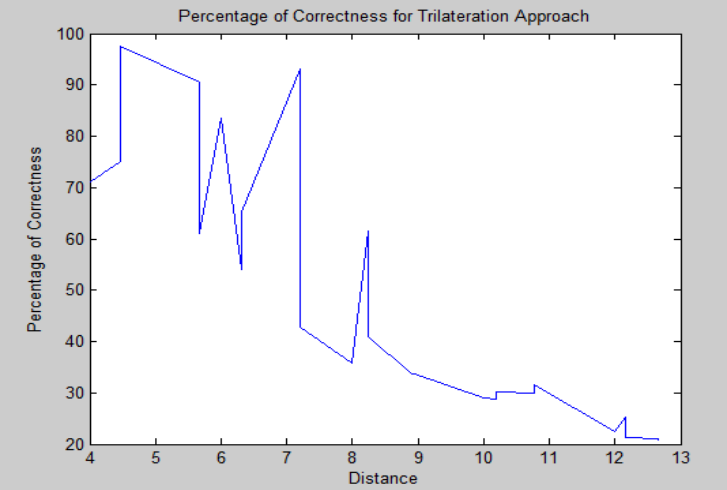


Fig. 53. Percentage of correctness for trilateration approach.

The high standard deviation was observed at each point. Finally, we obtained the amount of correct classification by the trilateration technique and found that indoor localization was strongly environment-dependent.

Observation on Trilateration

- The RSS readings are not stable due to environmental impairments. There is fluctuation in RSS values for each of the grid points. The distance estimation based on higher variance of RSS readings caused significant errors.

- RSS diversification is mainly caused by the isotropic properties of the antenna, antenna orientation and media through which the radio propagates.
- The variations in distance were about 10 meters for a specific position. Percentage of correctness decreases with the increase of distances for the trilateration approach. However, it did not smoothly decrease.
- Average percentage of correctness using Trilateration approach was 47.97% (Fig. 53). So, RSS cannot be used to measure accurate distance.

5.6 Artificial Neural Network for WSN Localization

A MATLAB simulation was created using neural network toolbox to determine the location of the wireless device using ANN for indoor Location-1 (E2-229). We found that the ANN achieved 60% of correctness during testing for the data obtained for indoor Location 1 shown in Table 17.

Table 17 Performance of Trilateration and ANN.

Trilateration	Trilateration	ANN
Classification Rate	47.97%	60.00%

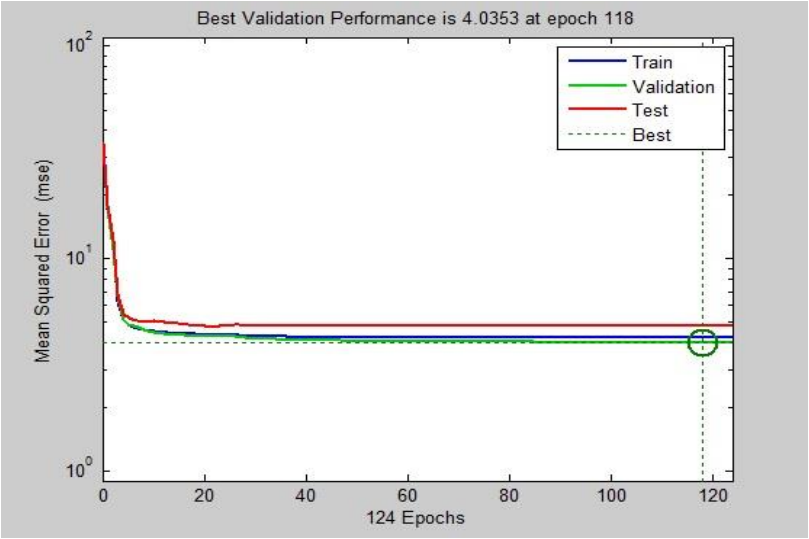


Fig. 54. MSE Vs Epoch in indoor location-1.

The network performed best validation performance of MSE = 4.035 at epoch 118 shown in Fig. 54.

5.7 Performance Evaluation for Synthetic Data

A MATLAB simulation was created to determine the location of the wireless device using the artificial neural network (ANN). We generate synthetic data for a 5mX5m grid and fed the triplet set of data to the ANN. We tested the system with 60,000 possible input-output arbitrary combinations. We needed to verify that the ANN did not prefer any combination of test data, so we randomly chose 5 test data points from 25 points, 60,000 times. We plotted the MSE vs. iterations and observed the variance of the MSE for training and testing phase in Fig. 55.

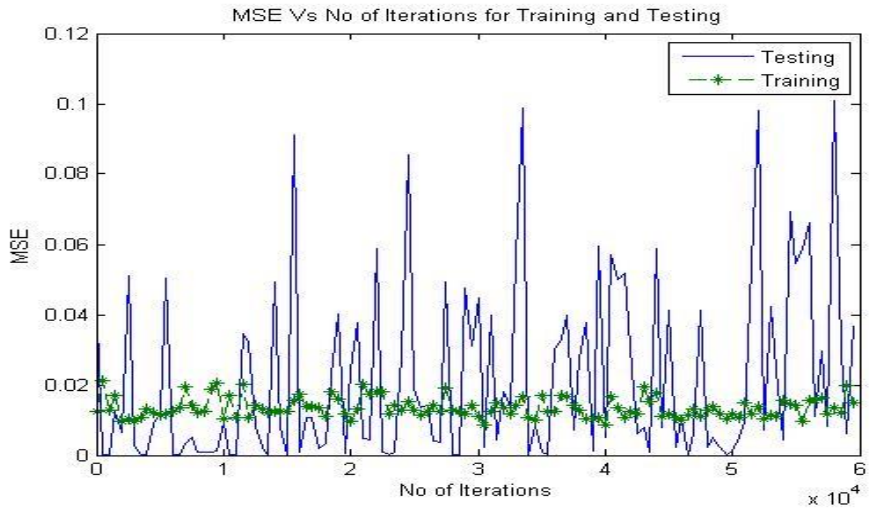


Fig. 55. Variation of MSE for synthetic data set in different iterations.

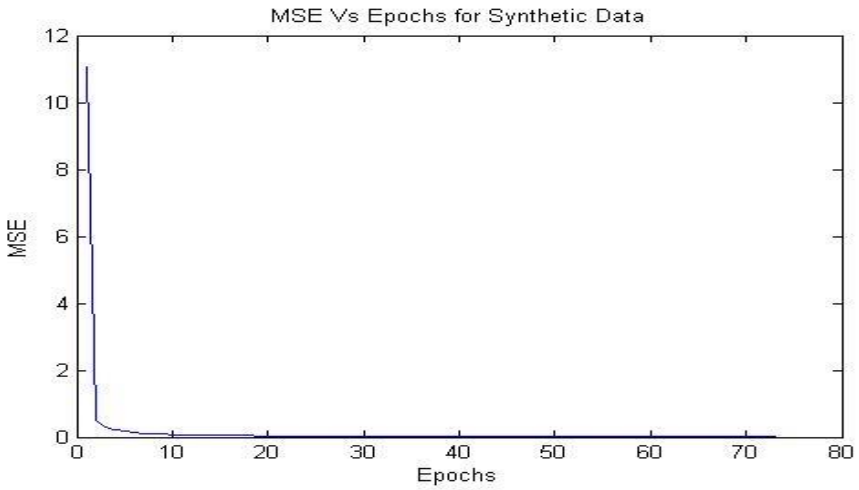


Fig. 56. MSE vs Epoch for synthetic data.

For synthetic data we see that network learns very quickly and the MSE obtained was 0.0138. We ran the experiment for 1000 epochs and observed after 73 epochs network stopped learning (Fig. 56).

We observed the performances for two different architectures to train the network. We consider 3-3-2, 3-6-2, 6-3-2 and 6-6-2 architectures to evaluate the performance of our

model shown in Fig. 57. We found that the change in hidden layer neurons did not affect our model significantly shown in

Table 18. But with the increase of hidden layers the mean square error changed. Thus for simulation, we considered a 3 neuron hidden layer model.

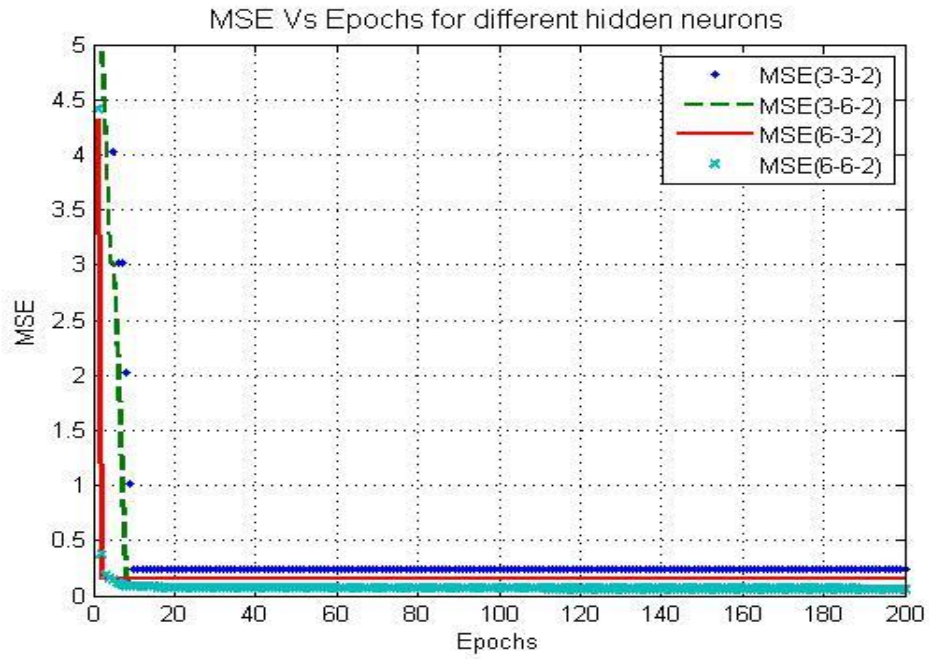


Fig. 57. MSE vs Epoch for different hidden layer neurons.

Table 18 MSE for different hidden neurons.

Architecture	3-3-2	3-6-2	6-3-2	6-6-2
MSE	0.23504	0.1517	0.154226	0.062088

5.8 Integrated ANN Approach for WSN Localization

We observed the MSE vs. epochs in Fig. 58 for different learning rate. We find that incorporating fractal dimension in the input elements gave better results than the RSS input based ANN model. Higher learning rate results in faster convergence.

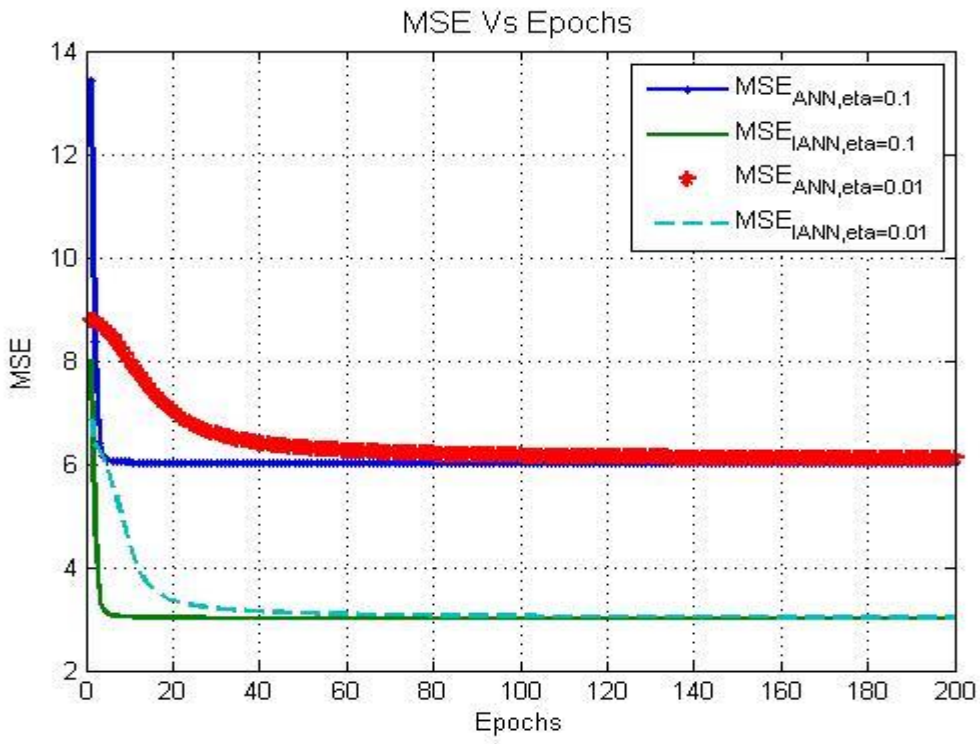


Fig. 58. MSE vs Epoch for ANN and proposed IANN models for indoor location.

ANN and IANN successfully classify 84.78% and 88.64% respectively during training and 79.92% and 86.88% respectively during testing phase. IANN has 8.7% higher classification rate than the ANN shown in Table 19. MSE obtained for training phases for ANN and IANN model are 6.003 and 3.002 respectively. MSE obtained for testing phases for ANN and IANN model are 6.1 and 3.0 respectively.

Table 19 Correct classification rate for training and testing phases.

ANN Model	**Training Phase**	**Testing Phase**
MSE	6.003	6.1
Classification Rate	84.78%	79.92%
IANN Model	**Training Phase**	**Testing Phase**
MSE	3.002	3.0
Classification Rate	88.64%	86.88%

We observed the model for different learning rates. We varied the learning rate from 0.01 to 0.1. We found that the accuracy of the network remained almost unchanged in Fig. 59.

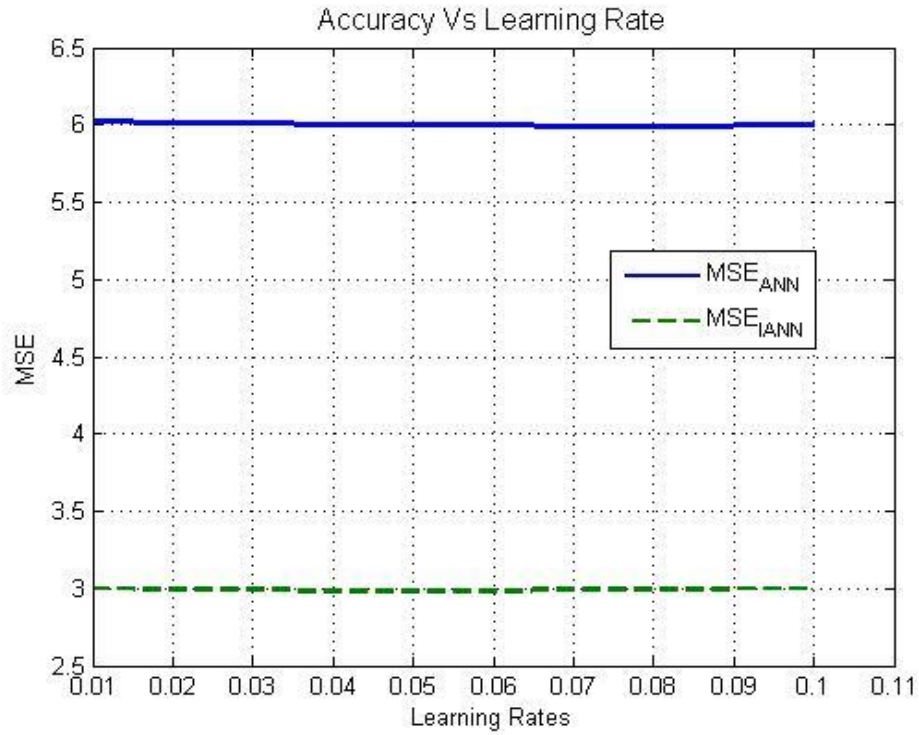


Fig. 59. Accuracy for different learning rates.

The MSE results obtained for different learning rates are shown in Table 20.

Table 20 MSE for different learning rates.

ANN					
LR	0.1	0.08	0.05	0.02	0.01
MSE	6.003051	5.994961	6.007947	6.016223	6.031215
Integrated ANN					
LR	0.1	0.08	0.05	0.02	0.01
MSE	3.002969	3.000967	2.98994	3.000494	3.003923

5.9 Summary

We observed that our proposed IANN model outperforms the traditional Trilateration technique and simple RSS based ANN model. However, the accuracy of the system remained almost unchanged with different learning rates. We determined that performance of the model does not change significantly with varying number of hidden neurons. Our proposed IANN model achieved 8.7% better correct classification than the simple RSS based ANN model.

Chapter 6

6. Conclusions and Recommendations for Future Work

This chapter concludes the thesis; it reviews experimental results and describes the contributions of the thesis. We outline the possible extensions of our work.

6.1 Conclusions

This thesis identified and discussed the pros and cons of various node localization techniques for wireless sensor networks. Many issues were identified that significantly impacted the performance of indoor node localization in these networks. We chose two types of indoor test bed locations: one with average or higher interference and the other with minimum interference. We observed and studied the impact of interference on the received signal strength readings. For an indoor location with a large amount of interference, the readings varied erratically.

We performed a stationarity analysis of the Wi-Fi signal to determine how the statistical properties of the Wi-Fi signal change over time. We found that, despite the signal showing non-stationary character as a whole, the statistical properties of the signal remained unchanged for a small time-window size. Within the small time-window, the fluctuations of the signal's mean and variance remained within a ±95% confidence interval. Thus, we conclude that the Wi-Fi signal is a weak sense stationary process considering small window.

We captured the receive signal strength of a Wi-Fi signal using a spectrum analyzer to provide the greatest possible amount of signal strength precision, given the available equipment. We required high precision signal strength readings for the multi-fractal analysis of the indoor Wi-Fi signal, as well as for the one dimensional inputs of the ANN.

6.2 Contributions

1. This thesis proposes a *feature integration* approach to the machine learning of node localization. The localization features of a Wi-Fi signal were learned (extracted) from a combination, or integration, of two dimensions of the signal's character. First, the raw received signal strength was used to represent the one dimensional features of the Wi-Fi signal. Second, the variance fractal dimension trajectory was used to represent the multi-fractal dimensional features of the Wi-Fi signal. Finally, this integrated combination of dimensions was input to an artificial neural network for learning and classification of localization features.
2. A stationarity analysis of the Wi-Fi signal determined the signal to be a weakly sense stationary process for small time windows.
3. A procedure and methodology using Labview and Agilent N9320B was developed to capture received signal strength, and this methodology will be published in the Canadian Microelectronics Corporation student library.
4. Training and testing data has been captured for three indoor locations and one outdoor location. This data can be used for further studies.
5. A conference paper [3] on the topic of this thesis has been published at WorldComp, 2012 which is the largest conference in this field in the world. This

conference was chosen to showcase the thesis to industry experts and more than 20,000 researchers from more than 80 countries of the world.

6.3 Numerical Results

The results of the experiments show that integrated ANN approach (IANN) performed 8.7% better classification than the “one dimensional input” ANN approach. The IANN approach achieved an 86.88% correct classification rate, while the one dimension input ANN approach achieved a 79.92% correct classification rate. Moreover, Trilateration method achieved only a 47.97% correct classification rate. ANN achieved 25% better classification rate than Trilateration technique.

6.4 Recommendations for Future Work

In future the model can be improved to achieve faster training, so that it may be used in real time node localization applications. We can implement other neural network techniques to model the system. A probabilistic neural network technique can be implemented in the future. The Wi-Fi signal was captured in the frequency domain using a spectrum analyzer. The spectrum analyzer could only be configured to output the magnitude of the spectrum, and so, it was not possible to perform an inverse Fourier transform without loss of signal character. In the future another type of network analyzer can be used, to either provide the real and imaginary parts of the Fourier transform, or provide time domain samples directly.

Our experiment was conducted by using high volume UDP traffic with random payload. In the future we could consider capturing TCP traffic or other traffic that represents actual client data in a commercial deployed area. We expect to get different complexity values for actual client TCP or other traffic.

Moreover, we can feed more information to the ANN, like antenna gain at the transmitter and receiver ends, propagation channel, noise, and period. We could compute TDOA, TOA algorithms and feed the proposed IANN more information about the environment.

Another area of improvement could be incorporating mobility for the localization approach. We can track any moving target in this regards. Presently we didn't consider mobility in our algorithm. A support for wider indoor experimental setup can give better accuracy and more realistic results of our experiment. We did our experiments on 2 dimensional models. We can also consider three dimensional localization environments in future.

References

[1] B. t. Commission, "Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems," RM-8143,CC Docket 94-102, FCC., Washington, DC, 1996.

[2] S. G. Glisic, Advanced Wireless Communications and Internet: Future Evolving Technologies, West Sussex, United Kingdom: Wiley; 3 edition , 2011.

[3] T. Kaiser, P. Card and K. Ferens, "Environment Feature Map for Wireless Device Localization," in *The 2012 World Congress in Computer Science, Computer Engineering, and Applied Computing (WORLDCOMP'12)*, Las Vegas, USA., July 2012.

[4] "Global Positioning System," [Online]. Available: http://en.wikipedia.org/wiki/Global_Positioning_System. [Accessed 10 October 2013].

[5] J. Wang , Q. Gao, Y. Yu , P. Cheng, L. Wu and H. Wang, "Robust Device-Free Wireless Localization Based on Differential RSS Measurements," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 12, pp. 5943 - 5952, Dec. 2013.

[6] C. Ma, "Integration of GPS and Cellular Networks to Improve Wireless Location Performance," in *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS 2003)*, Portland, September 9 - 12, 2003.

[7] S. Venkatraman, J. Caffery and H.-R. You, "A novel ToA location algorithm using LoS range estimation for NLoS environments," Sept. 2004 .

[8] C.-S. Chen, "Artificial Neural Network for Location Estimation in Wireless Communication System," vol. 12, pp. 2798--2817, 2012.

[9] L. Doherty, K. Pister and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings* , Anchorage, AK , 2001.

[10] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization: part I: a general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956-4980, October 2010.

[11] Y. Shen, H. Wymeersch and M. Z. Win, "Fundamental limits of wideband localization: part II: cooperative networks," *IEEE Transactions on Information Theor*, vol. 56, no. 10, pp. 4981--5000, October 2010.

[12] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses and N. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," vol. 22, no. 4, pp. 54 - 69 , July 2005.

[13] . J. A. Costa, N. Patwari and A. O. Hero,III, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 2, no. 1, pp. 39-64, February 2006.

[14] A. Ihler, J. Fisher, R. Moses and A. Willsky, "Nonparametric belief propagation for

self-localization of sensor networks," *IEEE Journal on Selected Areas in Communications,* vol. 23, no. 4, pp. 809 - 819, April 2005.

[15] U. Khan, S. Kar and J. Moura, "Distributed Sensor Localization in Random Environments Using Minimal Number of Anchor Nodes," *IEEE Transactions on Signal Processing,* vol. 57, no. 5, pp. 2000 - 2016, May 2009.

[16] H. Karl and A. Willig, Protocol and Architecture for Wireless Sensor Networks, J. Wiley & Sons, Ltd, 2005.

[17] N. Patwari, A. Hero, M. Perkins, N. Correal and R. O'Dea, , "Relative location estimation in wireless sensor networks," *Signal Processing, IEEE Transactions,* vol. 51, no. 8, pp. 2137 - 2148, 2003.

[18] B. Hofmann-Wellenhof, H. Lichtenegger and . J. Collins, Global Positions System: Theory and Practice, Springer-Verlag, 1997.

[19] A. Savvides, C.-C. Han and M. B. Strivastava, "Dynamic fine-grained localization in Ad-Hoc networks of sensors," in *Proceedings of the 7th annual international conference on Mobile computing and networking*, Rome, Italy, 2001.

[20] A. Savvides, H. Park and M. B. Srivastava, "The bits and flops of the n-hop multilateration primitive for node localization problems," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, Atlanta, Georgia, USA, 2002.

[21] D. Humphrey and M. Hedley, "Super-Resolution Time of Arrival for Indoor Localization," in *IEEE International Conference on Communications, 2008. ICC '08.*, Beijing, 19-23 May 2008.

[22] A. Nasipuri and K. Li, "A directionality based location discovery scheme for wireless sensor networks," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications,* Atlanta, Georgia, USA, 2002.

[23] S. Chan and G. Sohn, "Indoor Localization Using Wi-Fi Based Fingerprinting and Trilateration Techniques For LBS Application," in *7th International Conference on 3D Geoinformation*, Quebec, Canada., May 16-17, 2012.

[24] D. Niculescu and . B. Nath, "Ad Hoc Positioning System (APS) using AoA," 30 March-3 April 2003 .

[25] C. Savarese, J. M. Rabaey and K. Langendoen, "Robust Positioning Algorithms for Distributed Ad-Hoc Wireless Sensor Networks," in *In Proceedings of the General Track of the annual conference on USENIX Annual Technical Conference (ATEC '02),* Berkeley, CA, USA, 2002.

[26] A. Savvides, H. Park and M. B. Srivastava, "The n-hop multilateration primitive for node localization problems," *Mobile Networks and Applications,* vol. 8, no. 4, pp. 443--451, August 2003.

[27] S. Y. Wong , J. G. Lim, S. Rao and W. Seah, "Density-aware hop-count localization (DHL) in wireless sensor networks with variable density," in *Wireless Communications and Networking Conference, 2005 IEEE*, 2005.

[28] S. Yang , J. Yi and H. Cha, "HCRL: A Hop-Count-Ratio based Localization in Wireless Sensor Networks," in *4th Annual IEEE Communications Society Conference*

on Sensor, Mesh and Ad Hoc Communications and Network, San Diego, CA, 2007.

[29] O. Jegede and K. Ferens, "A Genetic Algorithm for Node Localization in Wireless Sensor Networks," in *The 2013 World Congress in Computer Science, Computer Engineering, and Applied Computing (WORLDCOMP'13)*, Las Vegas, USA., July 22-25, 2013.

[30] I. Vilovic and B. Zovko-Cihlar, "WLAN Location Determination Model Based on the Artificial Neural Network," in *47th International Symposium ELMAR-2005*, Zadar, Croatia, 08-10 June 2005.

[31] S. Bhardwaj, "ANN For Node Localization in Wireless Sensor Network," *International Journal of Advanced Research in Electrical, Electronic and Instrumentation Engineering*, vol. 2, no. 5, May 2013.

[32] S. Yun, J. Lee, W. Chung, E. Kim and S. Kim, "A soft computing approach to localization in wireless sensor networks," *Expert Systems with Applications*, vol. 36, no. 4, p. 7552–7561, 2009.

[33] Q. Zhang , J. Wang , C. Jin and Q. Zeng, "Localization Algorithm for Wireless Sensor Network Based on Genetic Simulated Annealing Algorithm," in *4th International Conference on Wireless Communications, Networking and Mobile Computing.*, Dalian, 12-14 Oct. 2008.

[34] J. Saxe, "Embeddability of weighted graphs in k-space is strongly np-hard," in *17th Allerton Conference in Communications, Control and Computing*, 1979.

[35] M. Noel, P. Joshi and T. Jannett, "Improved Maximum Likelihood Estimation of Target Position in Wireless Sensor Networks using Particle Swarm Optimization," in *Third International Conference on Information Technology: New Generations*, Las Vegas, NV, 10-12 April 2006.

[36] Y. Chen and V. Dubey, "Ultrawideband source localization using a particle-swarm-optimized capon estimator," in *IEEE International Conference on Communications, 2005*, 16-20 May 2005.

[37] A. Kumar, A. Khosla, J. Saini and S. Singh, "Meta-heuristic range based node localization algorithm for Wireless Sensor Networks," in *International Conference on Localization and GNSS (ICL-GNSS), 2012*, Starnberg, 25-27 June 2012.

[38] G. Giorgetti, S. . K. S. Gupta and G. Manes, "Wireless Localization Using Self-Organizing Maps," in *Proceedings of the 6th International Conference on Information Processing in Sensor Networks*, 2007.

[39] N. Patwari and A. Hero, "Manifold learning algorithms for localization in wireless sensor networks," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings.*, 17-21 May 2004.

[40] A. Shareef, Y. Zhu and M. Musavi, "Localization Using Neural Networks in Wireless Sensor Networks," in *In Proceedings of the International Conference on MOBILe Wireless MiddleWARE, Operating Systems, and Applications*, London, UK,, 22–24 June 2007.

[41] "Multilayer perceptron," [Online]. Available: http://en.wikipedia.org/wiki/Multilayer_perceptron . [Accessed 10 October 2013].

[42] M. F. Møller , "A scaled conjugate gradient algorithm for fast supervised learning," *NEURAL NETWORKS*, vol. 6, no. 4, pp. 525--533, 1993.

[43] M. Borenovic and A. Neskovic, "Positioning in WLAN environment by use of artificial Neural Networks and Space partitioning," *Institut TELECOM and Springer-Verlag France*, p. 665–676, 2009.

[44] M. Hosoun, Fundamentals of Artificial Neural Networks, MA: MIT Press, 1995.

[45] M. Gholami and R. W. Brennan,, "Localization of Industrial Wireless Sensor Networks: An Artificial Neural Network Approach," *Holonic and Multi-Agent Systems for Manufacturin*, vol. 6867, pp. 62-71, 2011.

[46] W. Kinsner, "A unified approach to fractal dimensions," *Intern. J. Cognitive Informatics and Natural Intelligence*, vol. 1, no. 4, pp. 26-46, October-December 2007.

[47] W. Kinsner and W. Grieder, "Speech segmentation using multifractal measures and amplification of signal features," in *7th IEEE International Conference on Cognitive Informatics, 2008. ICCI 2008*, Stanford, CA, 14-16 Aug. 2008.

[48] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt and P. Raymond, "Comparative study of heart sound localization algorithms," *Proc. SPIE*, vol. 8068, pp. 80680P-80680P-9, 2011.

[49] W. Kinsner, "Batch and real-time computation of a fractal dimension based on variance of a time series," UofM, Winnipeg, June 15, 1994.

[50] C. Fu and F.-f. Zhang, "On the Nonlinear Properties Analysis of Multipath Fading Channel," in *IEEE Region 10 Conference TENCON 2007*, Taipei, 2007.

[51] W. Kinsner, *Class notes, Fractal and Chaos Engineering*, Winnipeg: Dept. of Electrical and Computer Engineering, University of Manitoba, 2010, p. 900.

[52] Cisco, Wi-Fi Location-Based Services 4.1 Design Guide, San Jose, CA: Cisco Systems, Inc, May 20, 2008.

[53] W. Kinsner , V. Cheung, K. Cannons, J. Pear and T. Martin, "Signal classification through multifractal analysis and complex domain neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 36, no. 2, pp. 196 - 203, March 2006.

[54] A. . D. Buschbom, *Restricting wireless network access within the classroom*, Ames, Iowa: Information Assurance and Computer Engineering, Iowa State University, 2007.

[55] "umanitoba.ca/campus/physical_plant/fortgarry/pdfs/," University of Manitoba, [Online]. Available: umanitoba.ca/campus/physical_plant/fortgarry/pdfs/.

[56] "WLAN Fundamentals," [Online]. Available: http://www.ciscopress.com/articles/article.asp?p=1876001. [Accessed 10 October 2013].

[57] L. Cheng, C. Wu, Y. Zhang, H. Wu, M. Li and C. Maple, "A Survey of Localization in Wireless Sensor Network," *International Journal of Distributed Sensor Network*, vol. 2012, no. Article ID 962523, 2012.

[58] K. Benkic, M. Malajner, P. Planissic and Z. Cucej , "Using RSSI Value for Distance

Estimation in Wireless Sensor Networks Based on Zigbee," SPaRS Laboratory, UM-FERI Maribor.

[59] D. Niculescu and B. Nath, "Ad hoc positioning system (APS)," in *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, San Antonio, TX, 2001.

[60] M. Rahman, Y. Park and K.-D. Kim, "Localization of Wireless Sensor Network using artificial neural network," in *9th International Symposium on Communications and Information Technology*, Icheon, 28-30 Sept. 2009.

[61] C. Nerguizian, C. Despins and S. Affes, "Indoor geo-location with received signal strength fingerprinting technique and neural networks," in *Telecommunications and Networking - ICT 2004*, 2004.

[62] T. He, C. Huang, B. M. Blum, . J. A. Stankovic and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *9th Annual International Conference on Mobile Computing and Networking, (MobiCom '03),*, San Diego, Calif, USA, September, 2003.

[63] J. Wang, R. K. Ghosh and S. K. Das, "A survey on sensor localization," *Journal of Control Theory and Applications,* vol. 8, no. 1, p. 2–11, 2010.

[64] "GPS Resolution, Sensitivity and Repeatability," [Online]. Available: GPS Resolution, Sensitivity and Repeatability . [Accessed 10 October 2013].

APPENDIX A

A.1 Floor Maps of Indoor Test Locations

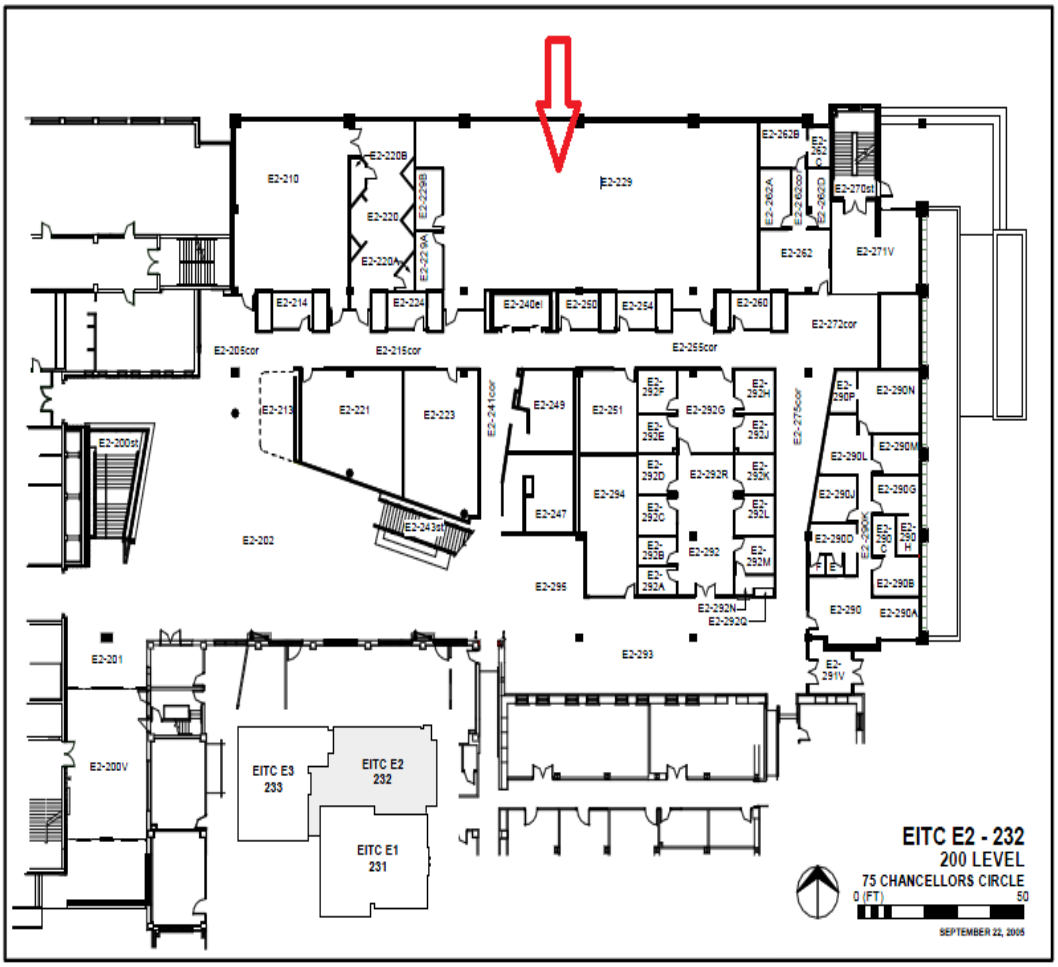


Fig. 60. Floor map of indoor test location 1(E2-229) [55]

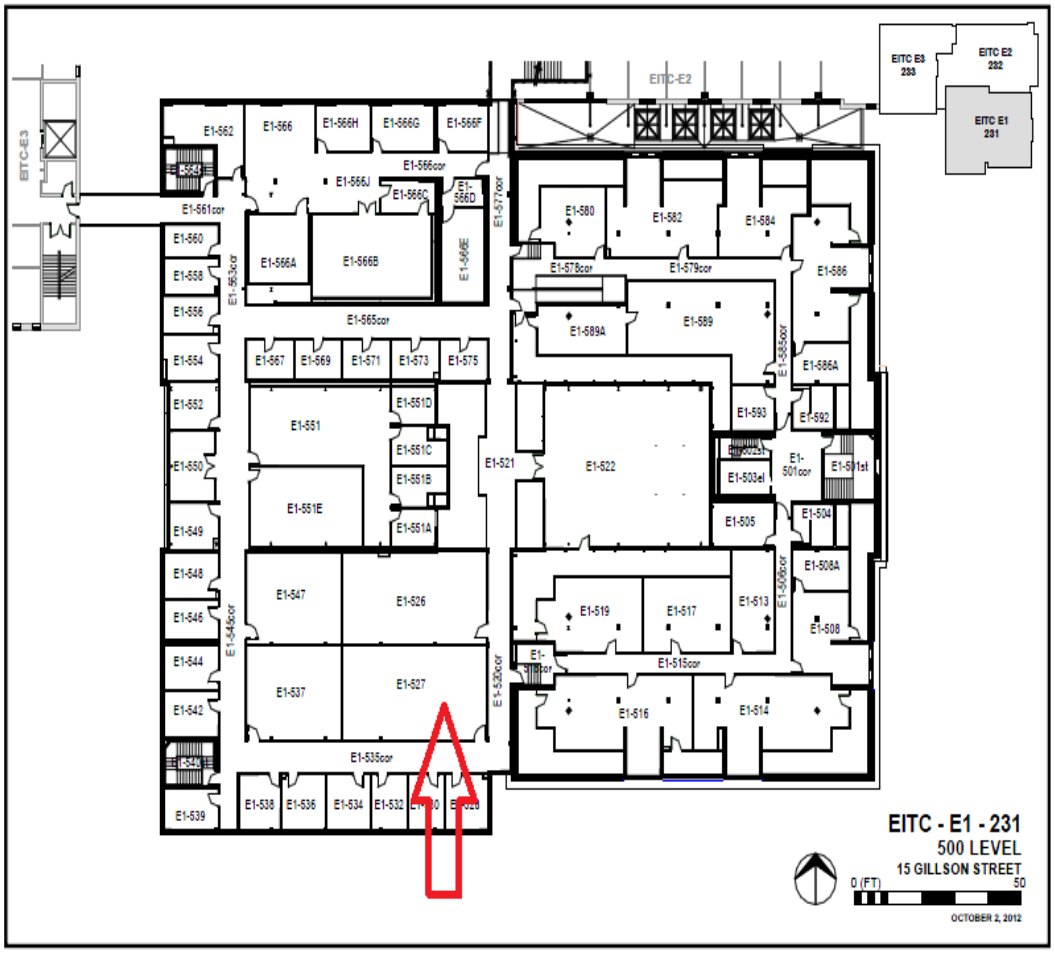


Fig. 61. Floor map of indoor test location 2(E1-527) [55]

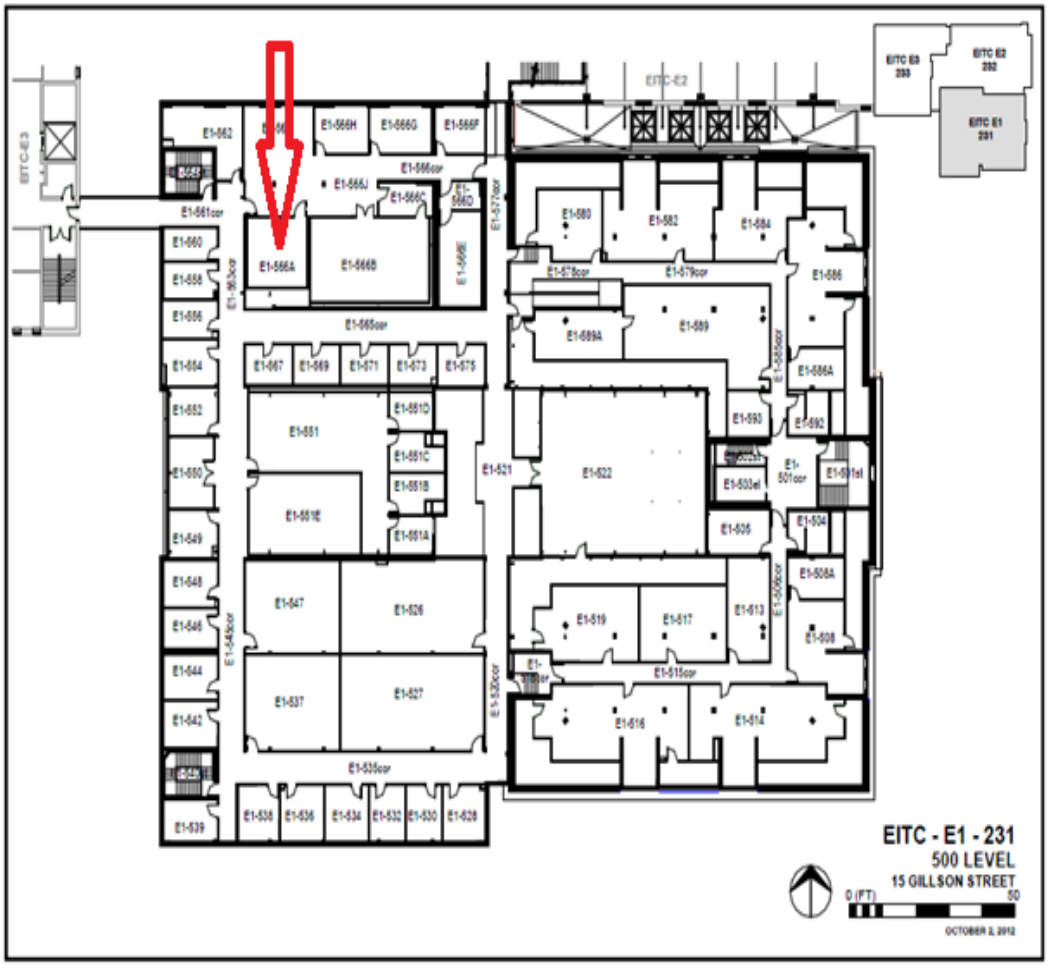


Fig. 62. Floor map of indoor test location 3(E1-566A) [55]

APPENDIX B

B.1 Variance Fractal Dimension for Different Locations Signal

Grid Points for Transmitter Position	Distance Between RX1 and TX	Distance Between RX2 and TX	Distance Between RX3 and TX	Points	Global VFD of RSS using Receiver1 (5,0)	Global VFD of RSS using Receiver2 (5,5)	Global VFD of RSS using Receiver3 (0,5)
(0,0)	5.00	7.07	5.00	Point1	1.9868	1.9872	1.99
(0,1)	5.10	6.40	4.00	Point2	1.9827	1.9837	1.9864
(0,2)	5.39	5.83	3.00	Point3	1.9802	1.9884	1.98545
(0,3)	5.83	5.39	2.00	Point4	1.9688	1.9866	1.9911
(0,4)	6.40	5.10	1.00	Point5	1.99	1.9881	1.97
(1,0)	4.00	6.40	5.10	Point6	1.981	1.9841	1.98
(1,1)	4.12	5.66	4.12	Point7	1.9872	1.99	1.9861
(1,2)	4.47	5.00	3.16	Point8	1.9875	1.9844	1.99
(1,3)	5.00	4.47	2.24	Point9	1.986	1.983	1.98
(1,4)	5.66	4.12	1.41	Point10	1.9866	1.9916	1.9888
(2,0)	3.00	5.83	5.39	Point11	1.99	1.987	1.98
(2,1)	3.16	5.00	4.47	Point12	1.99	1.9814	1.98
(2,2)	3.61	4.24	3.61	Point13	1.9853	1.99	1.9892
(2,3)	4.24	3.61	2.83	Point14	1.9827	1.9882	1.98
(2,4)	5.00	3.16	2.24	Point15	1.9873	1.9857	1.99
(3,0)	2.00	5.39	5.83	Point16	1.98	1.9889	1.99
(3,1)	2.24	4.47	5.00	Point17	1.9817	1.9885	1.984
(3,2)	2.83	3.61	4.24	Point18	1.99	1.9822	1.9894
(3,3)	3.61	2.83	3.61	Point19	1.99	1.9865	1.9854

(3,4)	4.47	2.24	3.16	Point20	1.9874	1.9886	1.99
(4,0)	1.00	5.10	6.40	Point21	1.99	1.9904	1.99
(4,1)	1.41	4.12	5.66	Point22	1.9881	1.99	1.985
(4,2)	2.24	3.16	5.00	Point23	1.9869	1.99	1.9832
(4,3)	3.16	2.24	4.47	Point24	1.9841	1.99	1.9876
(4,4)	4.12	1.41	4.12	Point25	1.9847	1.9871	1.9874