GSM based Communication-Sensor (CommSense) System

Abhishek Bhatta

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical Engineering

UNIVERSITY OF CAPE TOWN

Cape Town

South Africa

“I have no special talents. I am only passionately curious.”

-Albert Einstein

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in Engineering in the University of Cape Town. It has not been submitted before for any degree or examination in any other university.

Signature of Author Signed by candidate .....

Cape Town
May 28, 2018

Abstract

Using communication signals for radar applications has been a major area of research in radar engineering. In the recent years, due to the widely available wireless signals, a new area of research called commensal radars has emerged. Commensal radars use available wireless Radio Frequency (RF) signals to detect and track targets of interest. This is achieved by placing two antennas, one towards the transmitting base station and the other towards the surveillance area. The signal received by these two antennas are correlated to determine the location and velocity of the target.

When a signal passes through a channel, it reflects off the obstacles within its path. These reflections usually degrade quality of the signal and cause interference to the telecommunication systems. To mitigate the effects of the channel on a signal these systems transmit a known bit sequence within each frame. Our goal, with this thesis, is to design and implement a working prototype of a novel architecture for the commensal radar system, which uses these known bit sequences to extract the channel information and determine events of interest.

The major novelties of the system are as follows. Firstly, this system will be built upon existing communication systems using Software Defined Radio (SDR) technology. Secondly, this design eliminates the need for a reference antenna, which reduces the cost of the system and creates an opportunity to make the system portable. We name this system Communication-Sensing (CommSense). Since, our plan is to use Global System for Mobile Communication (GSM) as the parent system for the prototype development, we decide to update the name to GSM based Communication-Sensing (GSM-CommSense) system.

This thesis begins with theoretical analysis of the feasibility of the GSM-CommSense system. First of all, we perform a link budget analysis to determine the power requirements for the system. Then we calculate the ambiguity function and Cramér-Rao Lower Bound (CRLB) for a two-path received signal model. With encouraging theoretical results, we design a prototype of the system that can capture real GSM base station broadcast signals. After the design of the GSM-CommSense system, we capture channel data from multiple locations with varying environmental conditions. The aim for this set of experiment is to be able to distinguish between different environmental conditions. Then, we performed statistical analysis on the data by means of Probability Density Function (PDF) fitting, a goodness-of-fit test called chi-square test and a clustering algorithm called Principal Components Analysis (PCA). We have presented the results from each analysis and discussed them in detail. Upon, receiving positive results in each step we have decided to move towards using learning algorithms to categorise the data captured by the system. We have compared two widely accepted supervised learning algorithms, called Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP). The results showed that with the current hardware capabilities of the system and the amount of data available per GSM frame, the performance of SVM is better than MLP. Thus, we have used SVM to classify two events of detection and classification across a wall. We have presented our findings and discussed the results in detail.

We conclude our current work and provide scope for future work in development and analysis of the GSM-CommSense system.

Acknowledgements

I wish to express my gratitude to the following individuals for assistance and support during my research.

My parents: For providing unconditional support and continuous guidance for my endeavours in life. Without them I would never have been able to attain the success in life as it stands today.

My supervisor Associate Professor Amit Mishra: Who provided me technical as well as financial support throughout my postgraduate study. With his expert guidance and valuable feedback, I am inspired to put in more effort in order to attain my research goals.

University of Cape Town Personal: Notably, Dr. Simon Winberg for all the valuable discussions and ideas that we shared over the years regarding Software Defined Radio, Communication systems and C++ programming. Also, Associate Professor Daniel O'Hagan for all the insightful conversations in both the fields, technical and personal.

My fellow postgraduate students: All the members of the Radar Remote Sensing Group at the University of Cape Town, past and present who provided support, motivation, inspiration and company during my studies. Notably, Ashiv Dhondea, Po-Kai (Randy) Cheng and Stephen Paine for helping me with all the data capture campaigns and providing moral support in times of need. Also, Dane Du Plessis, Simon Lewis, Daniel Czech and John-Philip Taylor for all the interesting conversations that we shared over the years.

My childhood teacher Reba Chakravarty: For helping me improve the overall quality of this thesis, in terms of language.

I would also like to thank the authors of all the great free and open source libraries and software tools that I used during the course of my research.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

2G Second Generation. 18

ADC Analogue-to-Digital Converter. 21, 71, 76, 78

ARD Amplitude-Range-Doppler. 26

ARFCN Absolute Radio Frequency Number. 18, 65, 76

ASIN Application Specific INstrumentation. 7, 8, 62, 127

AWGN Additive White Gaussian Noise. xiii, 41, 42, 66, 68, 70, 132, 133

BCCH Broadcast Channel. 73

BER Bit Error Rate. 139

BPSK Binary Phase Shift Key. 50, 61

BT Time Bandwidth product. xvii, 132

BTS Base Transceiver Station. 17, 39, 72, 74

C0 Carrier Index Zero. 73

CIR Channel Impulse Response. xvii, 66–68, 70–72, 74, 132

CommSense Communication-Sensing. iii, xi, 2, 4

CPU Central Processing Unit. 79

CR Corner Reflector. xiii, xiv, 84, 91, 92, 98, 102, 104

CRLB Cramér-Rao Lower Bound. iv, xi, xii, xvi, 7, 9, 54–61, 124, 125

CW Continuous Wave. 1

DAB Digital Audio Broadcast. 24

DAC Digital-to-Analogue Converter. 21

DC Direct Current. 79, 80, 82

DF Degree of Freedom. 101

DVB-T Digital Video Broadcasting – Terrestrial. 22, 25, 128

EDGE Enhanced Data Rates for GSM Evolution. 18

ETSI European Telecommunications Standards Institute. 16

FCCH Frequency Correction Channel. 73

FDMA Frequency Division Multiple Access. 18, 65, 76

FIR Finite Impulse Response. 132

FM Frequency Modulation. 15, 23–27

GMSK Gaussian Minimum Shift Key. xiii–xv, xvii, 20, 52, 61, 69, 70, 130–133

GPRS General Packet Radio Service. 18

GPS Global Positioning System. 26

GRC GNU Radio Companion. 22

GSM Global System for Mobile Communication. iii, iv, xi–xiii, xx, xxi, 2–5, 8, 10, 13, 15–20, 27–30, 52, 58, 59, 61, 63, 65, 66, 68–74, 76, 78, 81, 82, 109, 120, 124–126, 128, 130, 131, 139

GSM-CommSense GSM based Communication-Sensing. iii, iv, xii, 2–4, 7, 8, 10–15, 22, 28–32, 39, 40, 46, 59–61, 63–66, 72, 75–79, 81–83, 93, 95, 107, 109, 113–115, 117, 120, 123–128

GSM900 GSM working in the frequency band of 900 MHz. 18

GUI Graphical User Interface. 22

HDMI High-Definition Multimedia Interface. 81

I In-Phase. xi, xiii, 10, 64, 69, 70, 73, 75, 81, 131–133

ISI Intersymbol Interference. 65

LTE Long-term Evolution. 128, 137, 138

MIMO Multiple Input Multiple Output. 137

MLE Maximum Likelihood Estimation. 87–90, 107

MLP Multi-Layer Perceptron. iv, xvii, 12, 109, 111, 115, 116, 123, 125, 127

mph mile per hour. 28

MSc Master of Science. 23

MSK Minimum Shift Key. 130

NRZ non-return-to-zero. 131, 133

OSR Oversampling Ratio. 71, 76, 132

PBR Passive Bistatic Radar. 1, 23, 25, 28

PC Principal Component. 11, 86, 108, 127

PCA Principal Components Analysis. iv, xi, xiv, 6, 11, 86, 101, 103–105, 107, 108, 120, 127

PCL Passive Coherent Location. 23–25

PDF Probability Density Function. iv, xiii, xiv, 83, 87–89, 92–98, 100, 101

PhD Doctor of Philosophy. 23

PLL Phase Locked Loop. 73

Q Quadrature. xi, xiii, 10, 64, 69, 70, 73, 75, 81, 131–133

QPSK Quadrature Phase Shift Key. 39

RAM Random Access Memory. 79

RBF Radial Basis Function . 111, 115

RCS Radar Cross Section. 26, 28, 37, 38, 58, 117, 127

RF Radio Frequency. iii, 2, 8, 18, 21–23, 76, 78, 79

SCH Synchronization Channel. 73

SDR Software Defined Radio. iii, xii, xvi, 2, 5, 7, 10, 13, 15, 20–22, 25, 29, 64, 72, 73, 78, 81, 124, 126

SI International System of Units. xxv, xxvi

SNR Signal-to-Noise Ratio. 4, 17, 24, 42, 65, 139

SVD Singular Value Decomposition. 103

SVM Support Vector Machines. iv, xvii, 12, 109–111, 115, 116, 120, 123, 125, 127

TDMA Time Division Multiple Access. 18, 65, 76

UCT University of Cape Town. 25, 85, 95, 105

UE User Equipment. 39

USB Universal Serial Bus. 79, 81

USRP Universal Software Radio Peripheral. 21, 25, 78, 79

VHF Very High Frequency. 25

WCDMA Wideband Code Division Multiple Access. 139

Wi-Fi Wireless Fidelity. 79, 80, 126

XML Extensible Markup Language. 22

List of Symbols

μs Micro Seconds. xi, xii, 9, 48–54, 60, 65, 126

A Ampere. 80, 82

bit basic unit of information used in computing. xxiv, 50, 78

bits plural of bit. 71, 74, 78

bits/s bits per second. 137

dB Decibel. 31–35, 39, 58, 135, 136

dBμV/m Decibel-micro-Volt per meter squared. 35, 37

dBi Decibel-isotropic. 31, 39

dBm Decibel-milliwatts. 31, 34–39, 58, 139

dBm/m^2 Decibel-milliwatts per meter squared. 35, 36

dBW Decibel-Watts. 35, 37

GB Giga bytes. 77, 79

GHz Giga Hertz. 24, 72, 79

h hour. 20

Hz Hertz. 137, 138

J/K Joules per Kelvin. 138

K Kelvin. 39, 138

kHz Kilo Hertz. 18, 65, 76, 78, 137

km Kilo Meter. 25, 32, 33, 35, 59, 60, 136

ksps kilo samples per second. 137

kW Kilo Watt. 25

m Meter. xiii, xiv, 28, 32, 33, 35, 39, 59, 60, 84–86, 92, 95, 97, 98, 104, 105, 114, 118, 119

m/s meters per second. 137

mA milli Ampere. 79

MB Mega Bytes. 77

Mbps Mega bits per second. 39, 137

MHz Mega Hertz. xxi, 18, 23, 24, 32, 33, 36, 39, 59, 60, 65, 72, 76, 78, 79, 81, 136–138

min minutes. 20

mm milli Meter. 117

ms milli seconds. 18, 20

Msps Mega samples per second. 39, 137

mW milli-Watt. 58

s second. 20, 110, 114, 119

V International System of Units (SI) unit of amplitude of a signal. xi, xii, 9, 48–53, 79, 80, 82

Vs Volt second. 43

Watts SI unit of power. 39, 58, 134, 137, 138

Chapter 1

Introduction

Radar systems rely on the principle that a signal propagating through a channel is affected by the properties of the medium and the obstacles in its path. In typical radar systems, a transmit signal is passed through a wireless channel. This signal, when received at the receiver is used to identify and classify the targets of interest in the channel. This, is achieved by correlating the received signal, also known as surveillance signal, with the reference signal and determining the differences. The field of radar system is ever growing and new types of radars are being introduced of late, such as passive radars [2, 3, 4] also known as commensal radars [5, 6, 7], cognitive radars [8, 9], Continuous Wave (CW) radars [10, 11], etc..

Since, radar systems were initially designed for military specific applications which were built as a stand-alone system, consisting of a dedicated transmitter and a receiver. In the recent times, with the increase in the demand of mobile communication systems, wireless signals are available all around us. This has led to a specific field of radar research, known as commensal radar [12, 13]. Commensal is a name borrowed from biology that means co-existence of two species out of which one is benefited and the other remains unaffected. This type of radar system is also known as Passive Bistatic Radar (PBR) [14, 15, 16]. In this document we will use the term commensal radar to describe such a system.

Commensal radars use wireless Radio Frequency (RF) signals transmitted by surrounding base stations to detect and track targets. These transmitters in the scope of commensal radar systems are known as illuminators-of-opportunity. In the recent years, with the abundance of wireless signal available in the environment, the popularity of research in the field of commensal radar has increased.

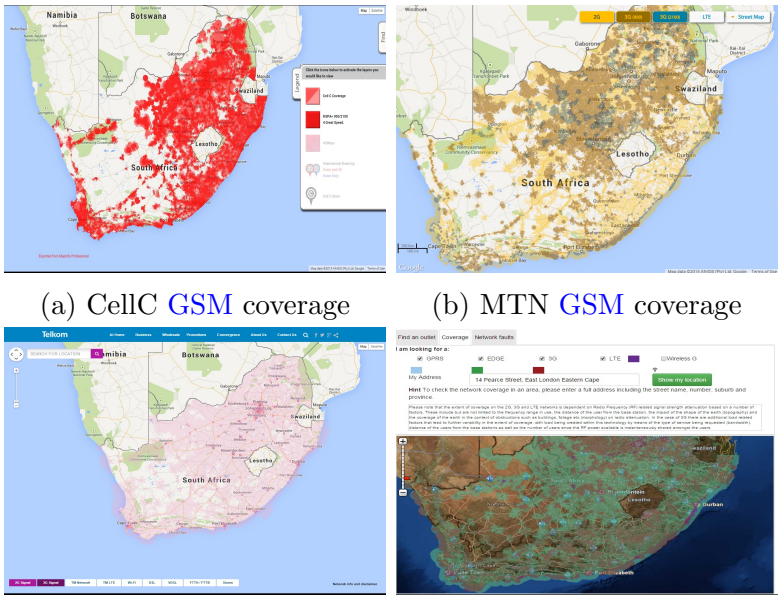
In this thesis, we aim to introduce a novel architecture of a commensal radar, which we call Communication-Sensing (CommSense). It uses channel estimation processing of telecommunication systems to estimate the changes in the environment. This system is built upon Software Defined Radio (SDR) platform [17, 18, 19, 20]. Here, the components that are traditionally built on hardware, such as mixers, filters, amplifiers, modulators, etc. are implemented using software on any processing device.

The major novelties of the system are as follows. First of all, this system is built upon existing communication systems using SDR hardware and software. The cost of implementation of such a system is considerably low. Secondly, unlike commensal radars, it does not process the information using correlation. Rather, it uses the known bit-sequences transmitted by communication systems to estimate the changes in the environment. This, eliminates the need for a reference antenna pointing towards the transmitter. We have used Global System for Mobile Communication (GSM) as the parent system, to implement CommSense. Thereby it has been updated to GSM based Communication-Sensing (GSM-CommSense). From here on, we will refer to this system as GSM-CommSense.

In this chapter, we present the motivation behind the design of GSM-CommSense system and reflect upon the reason for implementing the system on GSM standards. Next, we elaborate on the scope of our research and present in details our contributions in this field.

1.1 Motivation

The GSM coverage data from some of the most prominent network service providers in South Africa are given in Figure 1.1. The data is obtained from the official website of the service providers [21, 22, 23, 24].



(a) CellC GSM coverage (b) MTN GSM coverage

(c) Telkom GSM coverage (d) Vodacom GSM coverage

Figure 1.1: GSM coverage region for some of the most prominent communication service providers in South Africa, obtained from the official websites of individual service provider in December 2014.

From Figure 1.1, we can infer that the GSM standard provides wireless coverage to almost the entire Republic of South Africa. Since, we have been developing GSM-CommSense in South Africa, GSM has been the most suitable parent communication system.

A signal travelling through a wireless channel, reflects off multiple obstacles within its path as shown in Figure 1.2. These reflections, considered as interference, are eliminated by a method known as channel equalization, in case of communication systems. In order to equalize, the channel effects, wireless communication system transmits a known bit-sequence within its frames. The

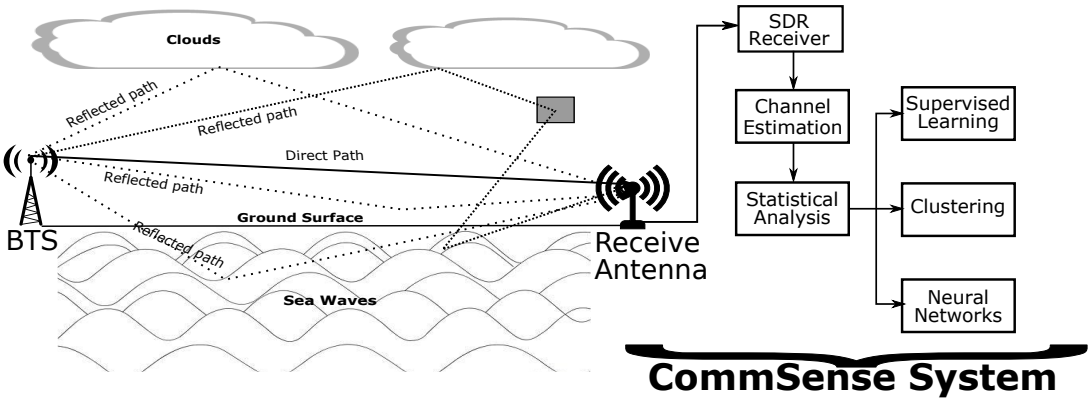


Figure 1.2: Concept diagram showing the system architecture for CommSense system.

purpose of these sequences is to estimate the effects of the channel from the received signal and extract the part with maximum Signal-to-Noise Ratio (SNR) and minimum interference. In case of GSM these are known as training sequence [25].

The GSM-CommSense system is designed to extract the channel information using the training sequence transmitted within each GSM frame. This extracted information then can be analysed statistically in order to characterise different environmental conditions.

1.2 Problem Description

Telecommunication systems estimate and equalize the effects of channel on a signal at the receiver. Designing a system that can extract the channel information from these blocks and characterise the environment accordingly can revolutionise the field of commensal radar. With this idea as a starting point, we started our investigations with the following hypothesis and research questions.

1.2.1 Research Hypothesis

1.2.1.1 Formulation

The research hypothesis for this thesis is as follows:

GSM based channel equalization modules and algorithms can be used to design and implement a commensal radar system to monitor the environment in real-time.

It is suggested that the following questions shall be answered in order to prove this hypothesis:

Q1: Is it possible to use the known bit-sequences transmitted by communication systems to visualize changes in the environment? Is there any other way to observe the channel effects (e.g. Viterbi decoder's trellis)?

- Survey different open source tools available.
- Find out the limitations of the system.
- Calculate and simulate various analytical tools to check feasibility of the system

Q2: Can the system be implemented in a hand-held portable system?

- Check accessibility of physical layer information in the possible implementation platforms such as SDR.
- Perform a user requirement analysis for all the equipment needed.
- Implement the system to receive data in real-time.
- Minimize the size of the system and make it portable.

Q3: What does the data capture from the system represent? Can it be used successfully to find difference amongst different environmental conditions?

- Use the designed system to capture data in multiple different environmental conditions.

- Statistically analyse the data and attempt to characterise it depending on the capture conditions.
- Check if there is separability of the data and comment on possible applications.

Q4: What is the possible application for the proposed system? How effective will the system be in real world applications?

- Determine possible applications of the designed system.
- Attempt to characterise the captured data using advanced machine learning algorithms.
- Find out solutions to improve the detection and classification accuracy of the system.

1.2.1.2 Justification

One of the early works in using radar systems for limited data communication has been done in 1969 by Ritterbach [26]. Since then, a lot of effort has been made to integrate the two systems [7, 27]. The hypothesis presented above proposes a novel commensal radar system which will be based on current communication system. Communication systems transmits a known bit-sequence to help the receiver in mitigating the channel effects and extracting the signal. The process of mitigating the channel impairments, known as channel equalization, involves estimation of the channel information and to remove it from the received data.

The purpose of this proposed work is to extract the channel information and analyse it in order to monitor the environmental parameters such as wind speed, humidity, land terrain etc.. In this proposed work linear estimation techniques are to be used for estimating the channel values, as they are computationally simple, efficient and easily implementable in real-time [28, 29]. After channel estimation, the captured data needs to be analysed in order to determine the differences between them. This can be done by implementing a statistical pattern recognition algorithms such as Principal Components Analysis (PCA) [30, 31,

32]. When, the clustering of the data is distinguishable between different types of channel that would validate the concept.

Once, the concept is validated the next challenge is to implement it on hand-held systems. For the proposed hypothesis to be tested on hand-held systems it is advisable to implement it on microprocessor based boards such as a Raspberry Pi or BeagleBone Black that can perform all the real-time processing. The receiver unit for this kind of implementation can be implemented using any open source SDR hardware such as Airspy, BladeRF or RTL-SDR and the software implementation can be done on open source SDR software platform known as GNU Radio.

It can also be noted here that there are two major challenges in implementing the system. First, the concept of resolution for conventional radar systems is not valid here because the available bandwidth is very narrow. This makes the measurement system an ill-posed inverse problem. Secondly, it is a forward-looking non-coherent radar system, extracting the information from part of the received data. This, theoretically limits the amount of information that it can capture. To overcome these challenges we will base the system on Application Specific INstrumentation (ASIN) principle [33, 34, 35] where, the system is trained to detect only a particular kind of event.

The solution to all the questions above and implementation of the system to operate in real-time producing the desired result, will prove the hypothesis.

1.2.2 Statement of Originality

The candidate believes that the following part of this work constitute original contributions to the field of commensal radars.

- Theoretical analysis of the GSM-CommSense system along with calculations and simulations for analytical tools such as link budget analysis, ambiguity function and Cramér-Rao Lower Bound (CRLB).

- Design of a hand-held prototype of the system that can receive and extract the channel information from GSM frames in real-time.
- Analysis of the captured data to classify different environmental conditions using GSM-CommSense.
- Possible real-time application of the system with test results.
- Provide insight into ASIN.

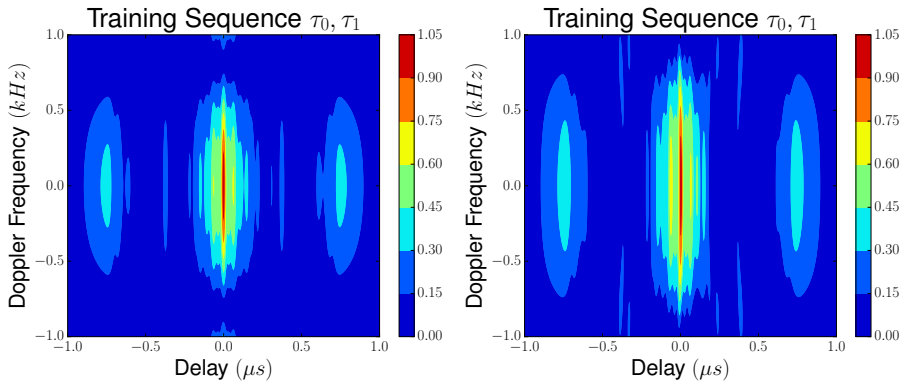
1.3 Thesis Contribution

In this research, we hypothesised a novel architecture to sense the environment using available RF bands [36]. A system that we named GSM-CommSense, is a special type of commensal radar. This thesis aims to prove the hypothesis by building a system that works in real-time and analyses the data captured with it. In the following, we summarise our contribution in this thesis.

1.3.1 Analytical tools to understand the feasibility of the system

Here we started with the link budget calculations to check the power requirements for the GSM-CommSense system and if the available GSM base stations were sufficient to implement the system. With the conclusion that there was enough power we moved on to calculate the ambiguity function.

We plotted the ambiguity function for a two path ground reflection model by passing different types of transmit waveforms. Figure 1.3 contains an example of the ambiguity function obtained by passing a GSM training sequence. We observed separable differences in the ambiguity function and concluded that a system extracting the channel information from the training sequence of a GSM system can be feasible.



(a) $a_1 = 0.11$ V and $\tau_1 = 0.012$ μs (b) $a_1 = 0.38$ V and $\tau_1 = 0.017$ μs

Figure 1.3: Normalized ambiguity function for training sequence based transmit signal in linear scale. The direct signal parameters are $a_0 = 0.5$ V, $\tau_0 = 0.01$ μs.

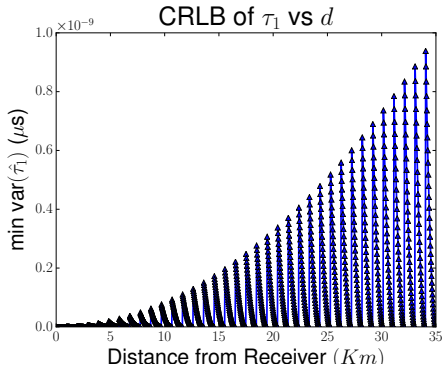
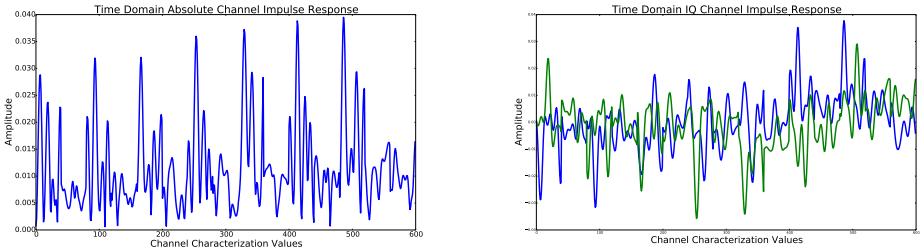


Figure 1.4: CRLB simulation to show the minimum variance of time delay for the scattered path in the presence of direct path with respect to the distance of the reflection point from the receiver.

We then, calculated the minimum variance in time of the received signal due to the scattered path in the presence of direct path signal. The simulation results proved that the delay in the received signal due to the scattered path was proportional to the distance of the point of reflection with respect to the receiver as shown in Figure 1.4. Although in the figure the variation was of the order of μs, the calculations were for very idealistic conditions. We believed, the signal would vary much more in real applications. With this belief we moved forward to designing the system.

1.3.2 Design and real-time implementation of the proposed system

We designed the GSM-CommSense system to receive data in real-time. We started with a top-level system architecture showing the multipath interference and used the GSM channel estimation blocks to extract the channel values from GSM broadcast signals transmitted by nearby base stations. Then, we moved to build a prototype where we used a SDR hardware called BladeRF to receive the GSM signals and performed the channel estimation on a laptop. At this point, we could plot and visualise the estimated channel values in real-time. The time-domain captured result is shown in Figure 1.5.



(a) Absolute channel value. (b) The I and the Q components.

Figure 1.5: Estimated normal channel values in terms of the In-Phase (I) and Quadrature (Q) components.

With the success of this design, we built a hand-held system that we could carry to different locations and capture data. The final system was built on a Raspberry Pi 3 running raspbian operating system and the BladeRF receiver. The software used for this implementation was GNU radio, which is an open-source SDR software. As a power source, we used individual off-the-shelf power banks to power each of the devices. At this point, we could capture data using the GSM-CommSense system. Next step was to analyse the data in order to understand the differences in the channel due to change in environmental parameters.

1.3.3 Statistical analysis of the captured data

In order to perform statistical analysis of GSM-CommSense data, we captured multiple sets of data in various locations as well as changing a few parameters within a location, such as location of a vehicle, humidity, etc.. We performed the analysis in three steps. First, we generated the histograms of the captured data and fitted it with known distributions, such as log-normal, Gaussian, Rayleigh, etc.. We observed visually that log-normal distribution provided the most consistent fit on all the datasets. To verify the visually observed information we performed a goodness-of-fit test, called chi-square test, on the empirical and fitted distributions. This reconfirmed the visually observed distribution fitting.

At this point, we decided to visualise if the data was separable within its components. Thus, we passed this through a clustering algorithm, called PCA, and observed the clusters formed due to the different environmental conditions. Figure 1.6 contains the plot showing one example of the separation between the datasets. Here, we plotted the first two Principal Component (PC) against each other and the data shows separation due to different humidity conditions.

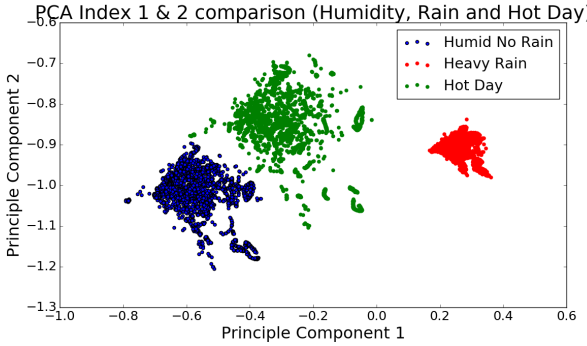


Figure 1.6: PCA of different climatic conditions at a single location. The conditions include high humidity without rain, heavy rain and Hot day with very low humidity.

1.3.4 Application of the designed system

With positive results in each step, we moved forward towards applying the system in real-time environment. In the process of applying the designed system, we decided to implement supervised learning algorithms and classify the environment. We compared two well-known algorithms, called Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP) and presented the results. During this comparison we found that SVM was by far a better classifier and predictor as compared to MLP with the current system capabilities. Therefore, we decided to implement SVM for real-time applications.

We investigated the operation of the GSM-CommSense system with SVM classifier to identify events across a wall. We conducted two different sets of experiments. First one is detecting a person across a wall and the second one is detecting a person carrying a weapon across a wall. We obtained encouraging results with a minimum average classification for detecting of a person across a wall as 77.458% and detecting of a person carrying a weapon across a wall as 95.208%.

In the future, we can investigate different applications of the GSM-CommSense system and also find ways to increase the prediction accuracy. A video providing brief introduction of the system along with a short demonstration in real-time is available at [37].

1.4 List of Publications

The research detailed in this thesis has contributed to the following publications:

1.4.1 Journal

1. A. Bhatta and A. K. Mishra, *GSM-based CommSense system to measure and estimate environmental changes*, in IEEE AESS Magazine, Special Edition, Feb 2017. [29]

2. A. Bhatta and A. K. Mishra, *GSM−CommSense-based through-the-wall sensing*, in Taylor and Francis Remote Sensing Letters, 2018. [38]

3. A. Bhatta and A. K. Mishra, *Ambiguity Function and Cramér-Rao Lower Bound calculation for CommSense system*, in IEEE Transactions on Aerospace and Electronic Systems, (Accepted)[39]

1.4.2 Conference

1. A. Bhatta and A. K. Mishra, *Implementation of GSM channel estimation using open-source SDR environment*, in ICMOCE, 2015 International Conference (pp. 322-325), IEEE, 2015. [28]

2. A. Bhatta and A. K. Mishra, *GSM based Hand-held CommSense for Environment Monitoring*, in ICIIS, 2016 International Conference (pp. 360-364), IEEE, 2016. [40]

3. A. Bhatta, A. K. Mishra and J. Pidanic, *Classification of CommSense data using learning algorithms*, in International Conference on Radar Systems, IET, 2017. [41]

1.5 Outline of the Thesis

The rest of the thesis is organised as follows. We first review some basic concepts of commensal radars, GSM and SDR in Chapter 2. In Chapter 3, we present some analytic tools that will help in the design of the GSM-CommSense system. Then, in Chapter 4, we present the design and implementation of the system to receive information in real-time. The data captured using the GSM-CommSense system is then analysed and presented in Chapter 5. At this point the system is designed and the data is analysed. In Chapter 6, we present some potential applications proving the capabilities of the GSM-CommSense system to work in real life. Finally, in Chapter 7, we summarise our contributions and discuss possibilities of future work.

1.6 Summary

In this chapter, we provided a brief introduction to the GSM-CommSense system along with the motivation that led to the conception of the idea. We formulated a research hypothesis and presented with four questions to prove the hypothesis. These questions act as the foundation of this research. We then presented parts of our findings, which we believed were original contribution to the field of commensal radars. We presented a list of publications in peer reviewed journals and reputed conferences that arose from this research. Lastly, we presented the outline of this thesis. Overall, in this chapter we provided a comprehensive overview of the research that led to the design of the GSM-CommSense system.

Chapter 2

Theoretical Background

2.1 Introduction

The goal of this chapter is to familiarise the reader with the basic theoretical background which is necessary to understand the underlying concepts of GSM-CommSense system. Here, we present a brief overview of commensal radars, mobile communication system focussing on GSM technology and discuss SDR. We introduce the basic theory and provide a review of the current literature on commensal radars and its implementation on GSM signals.

2.2 Commensal Radar

Commensal radar is a class of radar systems, that can detect and track objects utilising the reflection of the signal transmitted by illuminators-of-opportunity. Some examples of such transmitters are, Frequency Modulation (FM) radio, telecommunication base stations, digital audio broadcasting, digital video broadcasting, etc..

Implementation of a commensal radar is dependent on the wireless transmitter available in a region. This technology requires two antennas as shown in

Figure 2.1. Here, the transmit antenna is an illuminators-of-opportunity. A reference antenna is pointed towards the transmitter and receives the line of sight signal. The information received by this antenna is used as the reference signal. A surveillance antenna is pointed towards the region under study. The target echo signal is the signal that is reflected off a target. The reference signal is correlated with the surveillance signal to extract the range and the Doppler information about the target. [7, 12, 5]

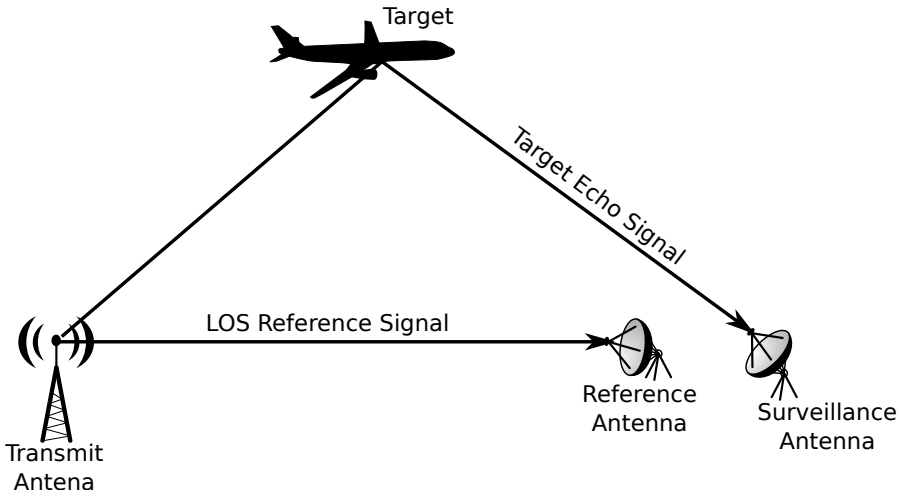


Figure 2.1: Commensal Radar description.

2.3 Mobile Communication System

The invention of telephones gave rise to a new field of research which, combined with the study of wireless data transfer, has created mobile communication systems. Over the years, there have been many versions of this system, also referred to as generations, which are developed and maintained by European Telecommunications Standards Institute (ETSI) [42]. ETSI is a standardisation institution that provides globally applicable standards for information and communications technologies. Out of all the communication protocols, the one of interest for this research is GSM [43, 44].

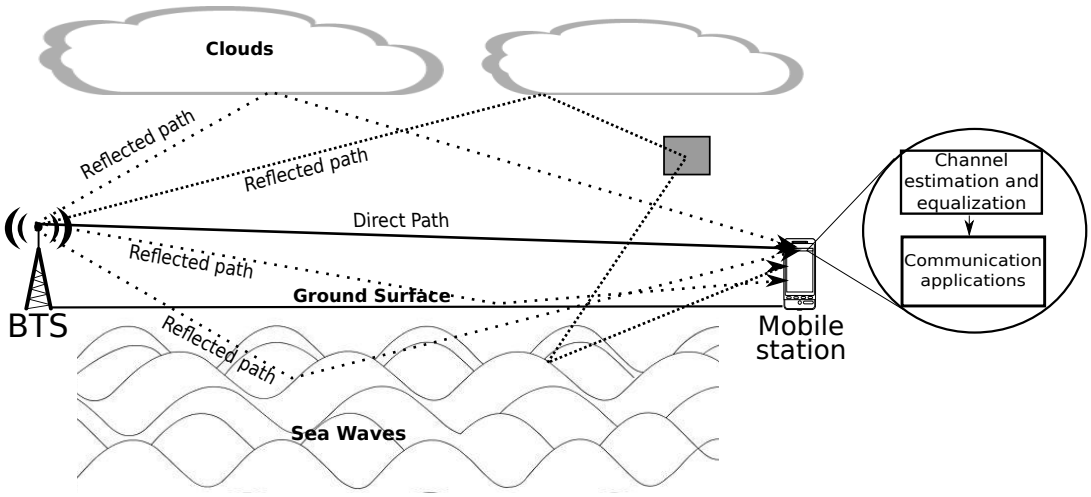


Figure 2.2: Overview of mobile communication system.

A generic mobile communication system is shown in Figure 2.2. Here, when the Base Transceiver Station (BTS) transmits a signal, it passes through a wireless channel and reflects off any obstacles in its path. This effect is commonly known as multipath effect [45, 46, 47]. Multipath effect causes constructive or destructive interference and phase shifting in a signal. Constructive interference causes the signal to gain power whereas destructive interference causes fading. Phase shifting changes the phase of the received signal which in case of phase coded signals is considered data loss. The effect of multipath on a signal is dependent on multiple factors like the modulation technique, distance between transmitter and receiver, the bandwidth of the signal, the medium of communication, etc..

Multipath effect is unwanted in telecommunication systems as it diminishes the overall quality of communication. In order to reduce the multipath effect, telecommunication systems use a technique called channel equalization [25, 48]. Channel equalization is a two-step process. Firstly the channel parameters are estimated [49, 50, 51] and secondly the signal is equalized to reduce the multipath effect and increases the SNR.

Since, this work is based on GSM signals, we present a brief overview of the protocol.

2.3.1 Global System for Mobile Communications

GSM is a standard for wireless telephonic communication. It is also referred to as Second Generation (2G) digital cellular networks. GSM was developed as a replacement of analogue communication systems and originally comprised of digital, circuit switched network with the capability of full duplex voice transfer. In course of time, it expands to incorporate data communication over circuit switched network, then by packet data via General Packet Radio Service (GPRS) [44, 52] and Enhanced Data Rates for GSM Evolution (EDGE) [53, 54].

GSM operates in multiple different carrier frequency ranges, which are separated by location. In general most of the GSM bands lie in the range of 900 MHz or 1800 MHz. At places where these frequencies are already allocated for some other purpose, the bands 850 MHz and 1900 MHz are allocated.

Regardless of the frequency of operation GSM uses Time Division Multiple Access (TDMA) [55, 56] and Frequency Division Multiple Access (FDMA) [57] in order to incorporate multiple users and multiple network providers. In order to operate in full duplex mode GSM uses different frequency bands for uplink and downlink. For example, GSM900 (GSM working in the frequency band of 900 MHz) uses $880 - 915$ MHz for uplink and $925 - 960$ MHz for downlink. These bands are further divided into 200 kHz bands each containing a single RF carrier and is referenced by a number called Absolute Radio Frequency Number (ARFCN). This division of frequency pools is called FDMA.

Each of the wireless carriers denoted by an ARFCN is further divided into TDMA frames. The frame structure establishes a schedule and pre-determines the use of each timeslot. Figure 2.3 contains a general orientation of frames in GSM. The GSM frame structure is designated as hyperframe, superframe, multiframe and frame.

The basic building block for a GSM frame is called a burst period, which lasts for approximately 0.577 ms. A group of 8 of these bursts is known as a TDMA frame. The duration of a frame is 4.615 ms and forms a basic unit which is used to define the logical channels. One burst period allocated in each TDMA frame

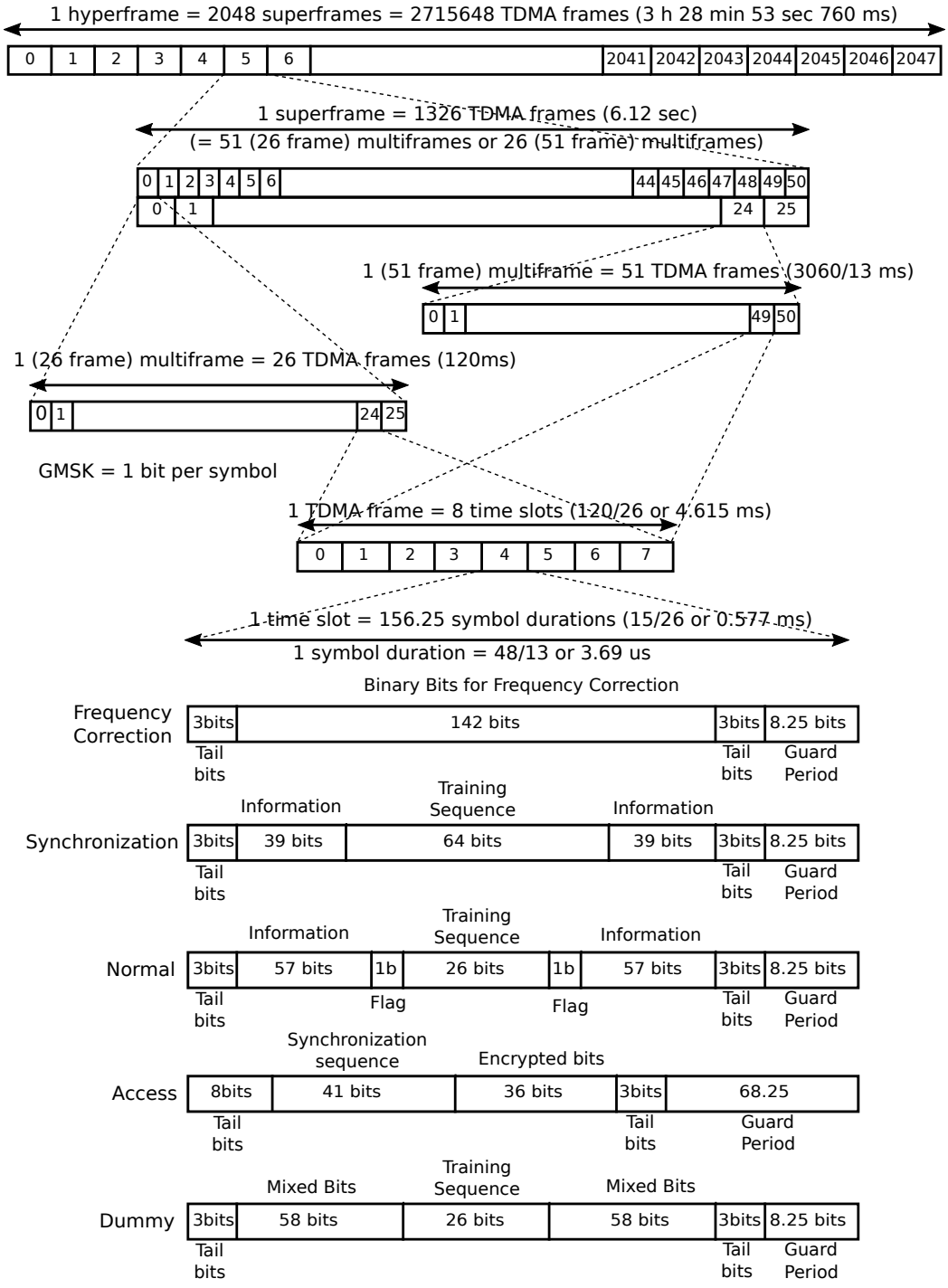


Figure 2.3: GSM frame structure.

is the definition of one physical channel. A GSM base station transmits two types of channels, named traffic and control. A set of GSM frames are grouped together to form a multiframe depending on the type of channel.

- Traffic multiframe: The traffic channel is used to transmit the communication information. It consists of multiframes with 26 frames and takes a total time of 120 ms.
- Control multiframe: The control channel is used to transmit control signal or broadcast signal. It consists of multiframes with 51 frames and takes a total time of 235.4 ms. The control channel is further divided into multiple logical channels. More details on each of these logical channels is presented in [58]

The multiframes are then packed together to form superframes taking 6.12 s. These consist of 51 traffic multiframes or 26 control multiframes. Since, the traffic and the control multiframes are of different durations, that brings them back in line taking the same time interval for all superframes.

A group of these superframes is called as a hyperframe. There are 2048 superframes grouped together creating a hyperframe. Each hyperframe lasts for 3 h 28 min 53 s 760 ms. This is the largest interval within a GSM frame structure.

More details on the GSM standard can be found in [59, 60, 58, 52]. Some details on the Gaussian Minimum Shift Key (GMSK) modulation scheme used in GSM along with channel characteristics is given in Appendix A.

2.4 Software Defined Radio

The basic concept of SDR is to build a completely configurable radio, so that a common hardware platform can be used in different wireless domains [18, 19, 20, 17]. It is a radio communication system, where the components that are traditionally designed using hardware, such as mixers, filters, amplifiers,

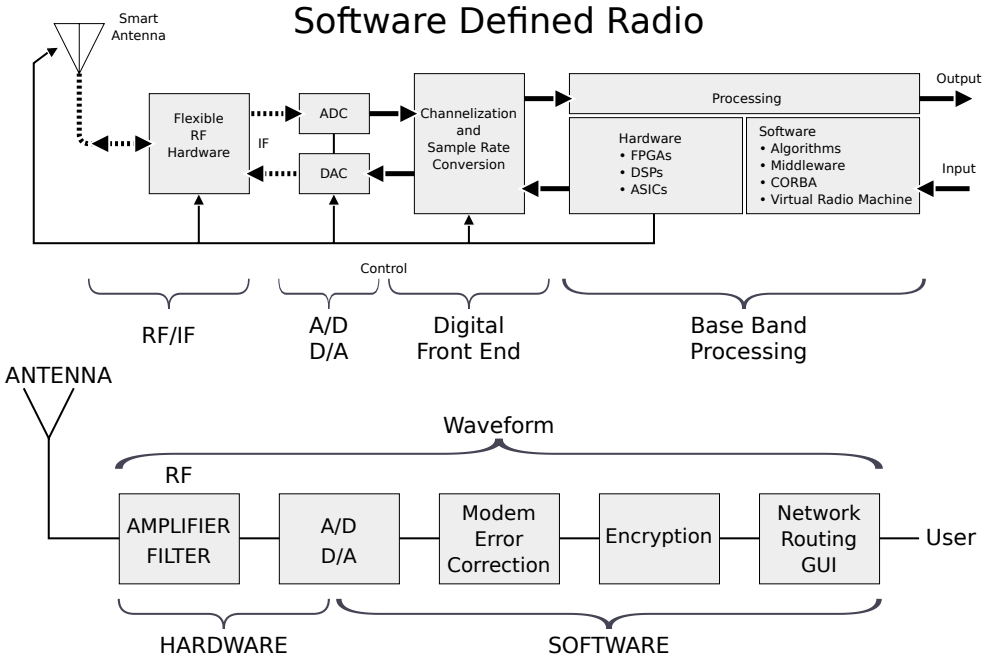


Figure 2.4: SDR block diagram (Image source [1]).

modulators/demodulators, etc. are built using software on a processing device. A traditional SDR contains a processing unit connected to an Analogue-to-Digital Converter (ADC), which is connected to an RF front end. The major part of signal processing is performed on a general purpose hardware. This, produces a design of a radio that can receive a wide variety of waveforms depending on the implemented software.

Figure 2.4 contains a block diagram representation of the concept of SDR. It shows an ideal receiver scheme, where an ADC is connected to an RF front end. The information from the ADC is read by a digital signal processor and then the software implementation processes the data according to the application. In case of a transmitter the signal is generated on a processing unit and this information is transmitted through the Digital-to-Analogue Converter (DAC), which is connected to the antenna.

Lately, with the growing market of SDR applications multiple SDR hardware are available in the market. Some such hardwares are, a Universal Software Radio

Peripheral (USRP) board developed by Ettus Research, a BladeRF developed by Nuand, a Digital Video Broadcasting – Terrestrial (DVB-T) dongle commonly known as the RTL-SDR, an Airspy, etc.. All of these devices work with a free SDR toolkit called GNU Radio [61, 62].

2.4.1 GNU Radio

One of the most widely used toolkit for implementation of SDR is called GNU Radio. It is a free toolkit that provides signal processing blocks to implement SDR and signal processing applications. Almost all SDR hardware drivers are available to be connected with this toolkit. It also provides support to create a simulation based SDR environment that does not require any RF devices to be connected. It has a wide community that provides support to anyone who is interested in the field of SDR. It contains multiple libraries to perform signal processing applications and also provides the ability to write out-of-tree modules. An out-of-tree module is a custom module that can perform any task programmed by the user.

It contains an easy to use GUI (Graphical User Interface) with a flowgraph type structure to connect different blocks, called GNU Radio Companion (GRC). A flowgraph on GNU Radio can be implemented either using c++ or python as the programming language and the graphical element for each of the defined blocks is created using Extensible Markup Language (XML).

The final implementation of the GSM-CommSense is based on the GNU Radio platform.

2.5 Literature Review

Having presented an overview of the commensal radar, mobile communication systems and SDR, we now review the existing work in the field of commensal radars. To determine the type of effort, we attempt to refer to the characteristic

of the publication whether they are Master of Science (MSc), Doctor of Philosophy (PhD) theses, series of papers or isolated articles. To facilitate the reading we sort them according to topics and chronologically within each topic.

2.5.1 Commensal Radars, Passive Bistatic Radars or Passive Coherent Location Radar

In this section, we present details about the existing work in the field of commensal radars. The term commensal was introduced by Professor Michael Inggs to the field of radar systems, in the year 2012. Commensal was a name borrowed from biology that meant co-existence of two species out of which one was benefited and the other remained unaffected. Some other commonly used names for this type of radar system are PBR and Passive Coherent Location (PCL). All these names are used interchangeably in this section as per mentioned in the original published document.

1974 A patent was granted to the United States Navy, detailing a system, they named "Bistatic Passive Radar" [63]. It is pointed out that in the active radar systems an RF transmitter is an integral part. Since, transmitting a signal could give away the location of the radar system, They proposed a system which could receive the wireless signals from uncontrollable, uncooperative transmitters to detect and track targets.

1997 A PBR for observation of the upper atmosphere was designed at University of Washington. This system was named "The Manastash Ridge radar" [64]. This relied on commercial FM broadcast signals in 100 MHz frequency range.

1998 The Manastash Ridge radar, mentioned above, was built and used to observe the auroral E-region irregularities [65].

1999 An analysis of the waveform to check the viability of a passive radar system was presented in [66]. It was mentioned here that with additional process-

ing, such as adaptive beamforming and application of temporal tracking to consecutive scans, it can act as an alternate surveillance system.

2002 In [67], a demonstrator developed by Dynetics Inc., was presented. It was a PCL radar system that used commercially available FM broadcast signals as illuminator-of-opportunity.

2003 An analysis of ambiguity function due to the waveform characteristics of off-air signals for a PCL was presented in [68].

2005 The performance of a PCL system due to different waveforms, such as FM signals, cellular phone base stations and digital audio broadcast, was analysed in [69]. This was a two part paper and in the second part different waveform properties were analysed [70].

2006 A multiband passive radar system demonstrator was introduced in [71].

2007 An improvement to the above mentioned multiband passive radar system was presented along with the results from trials in detection of a civil aircraft were discussed in [72].

In another publication, a detailed analysis of the different available transmitters that could be used as an illuminator-of-opportunity was presented [73]. In this, the authors had considered the type of signals and the interference environment to check the feasibility of using these signals for passive radar applications. The signals that have been studied are with different modulation schemes and ranges from 10 MHz to 2 GHz.

2008 A passive radar was used to detect and track targets with low SNR with a lot of false detections [74]. It was shown that tracking a target in this type of environment was also possible using passive radars. The system was not perfect then, there were multiple false tracks detected which needed to be minimised.

Another passive radar was built using Digital Audio Broadcast (DAB) signals as mentioned in [75]. It was proved that the system had a potential

to detect targets to a distance of 30 km with an experimental campaign using the radar.

Meanwhile, a passive radar demonstrator, called PaRaDe, was being developed at Warsaw University of Technology [76]. This system was based on the FM radio signals as illuminator-of-opportunity. The concept of digital beamforming in a PCL system was studied as part of a passive system development using an array of antennas in [77].

2009 A multistage processing algorithm was proposed to address the issue of canceling the direct signal, multipath and clutter echoes in [78]. The effectiveness of the algorithm was presented in simulation as well as on data acquired from an experimental Very High Frequency (VHF) PBR.

2010 The SDR technology, with an USRP, was used to design a multiband passive radar system [79].

The performance of an airborne PBR was analysed, depending on the geometrical configuration, to detect targets, in [80].

2012 High range resolution was obtained using a multichannel passive radar [81, 82]. Theoretical analysis was performed by means of ambiguity function and practical results were obtained using SDR technology to exploit multichannel DVB-T signals and were published.

The term commensal radar was introduced to describe an FM broadcast based air traffic control radar designed at University of Cape Town (UCT). This defined a prototype system which used certain FM transmitters with a power rating of 1.3 kW to yield a monostatic range in the order of 100 km. The system was being expanded into a multistatic configuration with the aim to position and track targets using multilateration, including multiple transmitters at diverse frequencies.

A commensal radar was proposed [7], where the reference and the surveillance channels were widely separated. This separation reduced the direct path interference by means of terrain shielding, allowing the reference and surveillance antennas to be placed far apart. The time synchronization for

this system was maintained by means of Global Positioning System (GPS) clocks.

Modelling and simulation of a multistatic, multi-frequency commensal radar was presented in [13]. This system consisted of multiple spatially distributed receivers. The direct signal was not processed at the receivers, instead it was shared through the network. Simulation results were presented for a multi-site radar with multistatic Radar Cross Section (RCS).

2013 Details of an FM based airborne passive radar system demonstrator was presented in [83].

Simulation and measurement results in target classification using commensal radar were presented in [84]. The authors provided proof of an application of commensal radars to classify the rotation rate of a Cessna 172 aircraft from different aspect angles.

An alternative method of using fourier transform for Doppler processing to compute Amplitude-Range-Doppler (ARD) data was presented in [12]. The work proposed the use of a recursive Fourier technique which allowed select channels to be computed. This introduced significant memory savings, and offered very fine time frequency decomposition.

2014 A plan for a commensal multistatic radar was developed and published in [6]. This radar was dependent on commercial FM broadcast bands. Here, the reason for choosing the term commensal to describe passive bistatic radars is explained in detail.

Details investigating theoretical placement of receivers for a commensal radar, that performs Doppler only tracking with a single transmitter and multiple receivers was presented in [85]. Here, theoretical concepts such as Shannon entropy and Cramér-Rao analysis were explained and used in the selection process of receiver positions. The developed theory was analysed by means of simulation to select the receivers that will minimise the error performance of a Doppler only tracking system.

In another research [86], the capabilities of commensal radar to adopt the standards of white space communication was investigated.

2015 Some trials were conducted to detect the performance of an FM radio band based commensal radar against smaller aircraft [87]. The detection range was presented along with the accuracy to detect the rotation rate of the aircraft's propellers.

Commensal radars were initially proposed to address the spectrum congestion issue. The current communication waveforms were not optimum for radar detection. The idea of designing communication waveforms that fulfils its primary purpose along with providing properties were in some sense optimised for radar signals were explored in [5].

2016 Another application of using commensal radars as ground based sensors to detect targets of interest in low earth orbit was investigated in [88]. A planning tool was developed to assess and simulate passes of objects such as the International Space Station. The results were analysed in detail.

2.5.1.1 GSM based commensal radars

In this section, we present some details about the existing work in using GSM signals for commensal radar applications.

2003 Analysis of feasibility and suitability of using the GSM downlink signals from the commercial base station as the 'illuminator-of-opportunity' for passive radar system was presented in [89]. It was proved by means of simulation and preliminary results that GSM signals could be used for radar application in detection of ground moving targets.

2005 Complete theoretical analysis with implementation of a prototype system using GSM signals for passive radar operation was presented in [90]. The results gathered by conducting numerous measurements to investigate the detection capability of such a system for various ground moving targets were presented. It was concluded from the preliminary processing results that GSM signals had the capability to be used as a radar waveform in order to detect and track ground moving targets.

2008 After proving the feasibility of using GSM broadcast signals for the purpose of passive radar, more data was analysed to detect aircrafts and presented in [91]. Here, based on the signals received by passive radar, the target RCS was estimated and the prediction performance of the system was analysed for different operating conditions.

2011 Another feasibility study was conducted at Warsaw University of Technology, about using GSM base station broadcast signals as passive radars [4]. Real life signals were recorded in measurement campaigns to extract traffic parameters such as average speed of vehicles and road capacity.

2012 A concept of determining the velocity of ground vehicles, such as passenger cars, trucks, buses, etc. was presented in [3, 92]. Detailed algorithm description with a full signal processing scheme dedicated for an experimental GSM based PBR was presented here along with real velocity measurements and tracking of ground moving vehicles. The system was able to simultaneously distinguish between different-sized objects, for example a truck and a cyclist.

2016 A train monitoring technology using passive radars was investigated in [93]. This system was based on the GSM-Railway radio communication technology and had the capability to determine the position and velocity of trains over any section of the railways with GSM coverage. A theoretical ambiguity function analysis on real measured waveforms suggested a range resolutions of approximately 850 m, and velocities down to less than 1 mile per hour (mph). The proof of concept was demonstrated using software defined passive radar system.

2.6 Summary

In this chapter, we present basic theoretical concepts necessary in understanding the GSM-CommSense system. The concept of commensal radars along with some details about mobile communication systems is explained here. Since the

GSM-CommSense system is based on GSM signals, a brief introduction of the GSM protocol is also presented here along with an explanation of the GSM frame structure. The hardware and software implementation of the GSM-CommSense system is done on SDR platform. Therefore, some theoretical background of the technology is presented here along with an introduction to GNU Radio.

In the literature review section, we present a chronological list of the relevant literature in the field of commensal radar. Here, we list the GSM based commensal radar work in a different subsection, as it is key to understand the GSM-CommSense system.

Chapter 3

Analytical GSM-CommSense[1]

3.1 Introduction

Using communication systems for radar applications have been an active topic of research for a long time [14, 7, 27, 2, 3]. This type of radar system, also known as commensal or passive radar, usually contains two antennas. One of the antennas is pointed towards the transmitter and the other towards the surveillance region. The signals from each of these antennas are correlated in order to isolate the targets under observation as shown in [3, 4].

In the quest to design a more optimized radar system, we decided to check the feasibility of extracting the environmental parameters from the known bit stream transmitted by communication system. Communication systems transmit these sequences for the purpose of equalizing the channel effects [36]. The proposed system should be capable of using the data captured by a single antenna to extract the information about the environment. Since we have been planning to implement the system on GSM standards, we have been deciding to name this system GSM-CommSense.

[1] Based on Abhishek Bhatta and Amit Kumar Mishra, "Ambiguity function and Cramér-Rao Lower Bound calculation for CommSense system," in *IEEE Transactions on Aerospace and Electronic Systems* (Accepted).

In this chapter, we present a set of analytical tools that will act as a basic guideline for the design of a GSM-CommSense system. Our contribution here, is a study of feasibility to implement such a system.

3.2 Link Budget Analysis

Link budget analysis of a system includes calculation of all the gains and losses between the transmitter and the receiver. This provides a general introduction to the various parameters of the system such as the power requirement or in this case whether the transmitted power is enough to implement a GSM-CommSense system, range of operation of the system, etc.. The various parameters of link budget are explained along with a sample calculation in [94], detailing the importance of the analysis. Since the proposed GSM-CommSense system is a receive-only system, there is no control on the transmit power of the parent communication system. Thus, an analysis of the available power and overall effectiveness of such an implementation is necessary.

The simple link budget equation, in the logarithm scale, is given as:

$$P_{rx} = P_{tx} + G_{tx} - L_{tx} - L_{fs} - L_m + G_{rx} - L_{rx} \quad . \tag{3.1}$$

Here, P_{rx}, represents the received power and is measured in dBm. P_{tx}, is the power transmitted by the communication system in dBm. The gain of the transmitting antenna is represented by G_{tx} and the gain of the receiving antenna is represented as G_{rx} with a unit of dBi. Various losses are mentioned such as losses in the transmitter, L_{tx}, due to connectors, cables, etc.. Similarly, the losses in the receiver is represented by, L_{rx}. Free space path loss is represented by, L_{fs}, and some miscellaneous losses, L_m are caused due to fading margin, body loss, etc.. All the losses are measured in dB.

Definitions of some terms used for link budget analysis are explained in Appendix B.

3.2.1 Path loss calculations

In order to calculate the link budget for the GSM-CommSense system, each of the individual parameters needs to be calculated. One such parameter is the free space path loss, also known as path loss or propagation loss, between the transmitter and the receiver. The path loss is dependent on the path taken by the signal to reach the receiver and the type of environment through which the signal is transmitted. For example the path loss for a signal being transmitted over a rural region will be different from the path loss over an urban region.

From literature, it can clearly be stated that the path loss shows logarithmic behaviour to the distance between the transmitter and the receiver. Multiple statistical analysis on empirical data for different propagation conditions are available as reference in [95, 96, 97, 98]. Out of all the available literature, Okumura's report [95] is still considered the most practical, as it arranges field strength and service area. Hata took Okumura's prediction methods and created an empirical formula for propagation loss as shown in [96]. The empirical formula created by Hata, also known as Hata model, is used in this work to calculate the path loss.

Hata model provides a formula for calculating the path loss between a transmitter and a receiver placed at certain heights. The standard formula for path loss is given as:

$$\begin{aligned} L_p(\text{dB}) =& 69.55 + 26.16\log_{10}(f_c) - 13.82\log_{10}(h_t) - \\ & a(h_m) + (44.9 - 6.55\log_{10}(h_t))\log_{10}(D) \quad . \end{aligned} \tag{3.2}$$

Here, L_p, represents the path loss, calculated in dB. f_c, is the frequency of operation, calculated in MHz. The height of base station is given by h_t, in m. The distance between the transmitter and the receiver antenna is given as D, in km. The correction factor for the mobile station antenna of height, h_m, in m, is given by $a(h_m)$, in dB. Range of variables for the model, as presented in [96] are:

$$f_c = 150 - 1500 \text{ MHz}$$

$h_t = 30 - 200$ m

$D = 1 - 20$ km

$a(h_m) = 0$ dB for $h_m = 1.5$ m

The correction factor value changes according to the size and structure of the cities.

- Correction factors in a medium-small city and is represented as:

$$a(h_m) = (1.1\log_{10}(f_c) - 0.7)h_m - (1.56\log_{10}(f_c) - 0.8) \quad . \tag{3.3}$$

  Here:

  $h_m = 1 - 10$ m

  $f_c = 150 - 1500$ MHz.

- Correction factor in large city is given as:

$$\begin{aligned} a(h_m) &= 8.29(\log_{10}(1.54h_m))^2 - 1.10(\text{dB}) \text{ for } f_c \leq 200 \text{ MHz} \\ a(h_m) &= 3.2(\log_{10}(11.75h_m))^2 - 4.97(dB) \text{ for } f_c \geq 400 \text{ MHz} \end{aligned} \tag{3.4}$$

- Suburban correction factor, generally denoted by $K_r(dB)$, is given as:

$$K_r = 2\log_{10}\left(\frac{f_c}{28}\right)^2 + 5.4 \tag{3.5}$$

- Open area correction factor, represented by $Q_r(dB)$, is given as:

$$Q_r(dB) = 4.78(\log_{10}(f_c))^2 - 18.33\log_{10}(f_c) + 40.94 \tag{3.6}$$

Path loss for suburban area is calculated as:

$$L_{ps} = L_{pu}(dB) - K_r \tag{3.7}$$

Here, $L_{pu}(dB)$, is obtained from Equation 3.2 with the correction factor of urban area.

3.2.1.1 Implementation

The $L_p(dB)$ is dependent on the distance between the transmitter and the receiver. The total path is split into two separate transmit channels, to simplify calculations, and the values are added in the end to reveal the total path loss. Figure 3.1 contains a diagram showing the paths for which the losses are being calculated here. It can be identified that the two different paths are named d_1 and d_2. Thus, the total path loss will be:

$$L_p = L_{pd_1}(dB) + L_{pd_2}(dB) \quad . \tag{3.8}$$

Here, $L_{pd_1}(dB)$, is the calculated loss for the path d_1, ranging from the base station (assumed to be at the point of origin for this calculation) to the point x. $L_{pd_2}(dB)$ is the loss for path d_2, ranging from x to D.

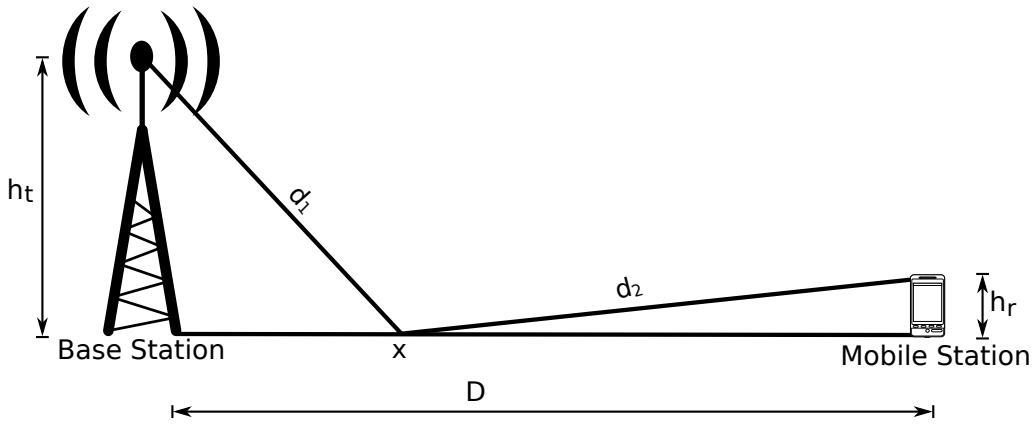


Figure 3.1: Diagrammatic representation of the path loss calculations.

3.2.1.1.1 Path loss calculations for "d_1"

Received power, $P_{r1}(dBm)$, at the point x is a function of received power

density, $P_u(\text{dBm/m}^2)$, and is given as:

$$P_{r1}(dBm) = P_u(\text{dBm/m}^2) \quad . \tag{3.9}$$

Here, the effective aperture of the antenna is not included which will be included later, in the calculation of path loss for d_2. The effective aperture is given as:

$$A_{eff} = \frac{\lambda^2}{4\pi} \quad . \tag{3.10}$$

Here, λ is the wavelength of the signal calculated in m.

The received power density is given as:

$$P_u(\text{dBm/m}^2) = E(\text{dB}\mu\text{V/m}) - 10\log_{10}(120\pi) - 90 \quad . \tag{3.11}$$

Since the propagation loss is the difference between the radiated power and the received power, using Equation 3.9 the loss L_{p1} for path d_1 is given by:

$$\begin{aligned} L_{p1} &= P_t - P_{r1} \\ &= P_t(dBW) - E(\text{dB}\mu\text{V/m}) + 145.8 \quad . \end{aligned} \tag{3.12}$$

The propagation loss, $L_{p1}(\text{dB})$, between two isotropic antennas is given by the Okumura's prediction curves in [95] and is represented as:

$$L_{p1} = 178 - 10\log_{10}\left(\frac{\lambda^2}{4\pi}\right) - E(\text{dB}\mu\text{V/m}) \quad . \tag{3.13}$$

The field strength as obtained from [96] as a function of distance, D, measured in km, is:

$$E(\text{dB}\mu\text{V/m}) = \gamma + \beta\log_{10} D \quad . \tag{3.14}$$

Here, γ and β are constants determined by the height of base station (h_t in m)

and the frequency (f in MHz). The general equation of path loss is given by:

$$L_p(dB) = A + B \log_{10} D \quad . \tag{3.15}$$

Substituting Equation 3.13 and 3.14 in Equation 3.15 we get:

$$A = 178 - \gamma + a(h_{m1}) \tag{3.16}$$

$$B = -\beta \tag{3.17}$$

Here, $a(h_{m1})$ is the correction factor for the mobile antenna height.

Introducing the empirical formula from the tables in [96] we get:

$$A = \alpha - 13.82 \log_{10}(h_t) - a(h_{m1}) \tag{3.18}$$

$$B = 44.9 - 6.55 \log_{10}(h_t) \tag{3.19}$$

Here,

$$\alpha = 31 + 46.16 \log_{10}(f_c) \quad .$$

Loss in signal strength for the path d_1 is given as:

$$\begin{aligned} L_{p1} =& 31 + 46.16 \log_{10}(f_c) - 13.82 \log_{10}(h_t) - \\ & a(h_{m1}) + (44.9 - 6.55 \log_{10}(h_t)) \log_{10}(d_1) \end{aligned} \tag{3.20}$$

3.2.1.1.2 Path loss calculations for "d_2"

Considering d_2 as a complete new path the whole formula for L_{p2} is derived. Received power, P_{r2}, is similar to the P_{r1} along with the effective aperture of the receiver antenna given as:

$$P_{r2}(dBm) = P_{u2}(\mathrm{dBm/m^2}) + 10 \log_{10}(A_{eff}) \quad . \tag{3.21}$$

The formula for effective aperture of the antenna is shown in Equation 3.10. The received power density is given in Equation 3.11.

The transmit power in this case will be a factor of the RCS of the reflecting surface as:

$$P_{t2}(dBm) = P_{t2}(dBW) + \sigma(dBm) + 30 \quad . \tag{3.22}$$

Here, σ is the RCS measured in dBm. Path loss L_{p2} is given by:

$$\begin{aligned} L_{p2} =& P_{t2} - P_{r2} \\ =& P_{t2}(dBW) - E_2(\text{dB}\mu\text{V/m}) - 10\log_{10}(120\pi) \\ &+ \sigma(dBm) + 145.8 \quad . \end{aligned} \tag{3.23}$$

Using the Hata model from [96] and calculating the path loss for d_2 we get:

$$\begin{aligned} L_{p2} =& 69.55 + 26.16\log_{10}(f_c) - 13.82\log_{10}(h_{b2}) - \\ & a(h_{m2}) - \sigma(dBm) + (44.9 - 6.55\log_{10}(h_{b2}))\log_{10}(d_2) \quad . \end{aligned} \tag{3.24}$$

Since in this case the height of the base station is non-existent we can assume it to be a minimal value (as $\log_{10}(0)$ is not defined let it be 10^{-3}).

3.2.1.1.3 Final Path Loss "(L_p)"

The final path loss can be given as the sum of the loss in each path. This can be obtained by adding Equation 3.20 and 3.24 as shown in Equation 3.8. Substituting all the values and deriving the equation for path loss we get the final equation as:

$$\begin{aligned} L_p =& 141.95 + 72.32\log_{10}(f_c) - 13.82\log_{10}(h_t) + \\ & (44.9 - 6.55\log_{10}(h_t))\log_{10}(d_1) + 64.5\log_{10}(d_2) \\ & - a(h_m) - \sigma(dBm) \quad . \end{aligned} \tag{3.25}$$

3.2.1.1.4 Calculation considering the correction factor for suburban region

In case of a suburban region, the correction factor is calculated as shown in

Equation 3.7. The value of correction factor is calculated as:

$$a(h_m) = a(h_{m1}) + a(h_{m2}) \quad , \tag{3.26}$$

$$a(h_m) = (1.1\log_{10}(f_c) - 0.7)(h_m + 10^{-3}) - (3.12\log_{10}(f_c) - 1.6) \quad . \tag{3.27}$$

In this particular case the value of K_r is given by:

$$K_r = 4\log_{10}\left[\frac{f_c}{28}\right]^2 + 10.8 \quad . \tag{3.28}$$

Substituting the values of Equation 3.25, 3.27 and 3.28 in Equation 3.7 provides the final value of the signal strength lost in the channel and is given as:

$$\begin{aligned} L_{ps} =& 141.95 + 72.32\log_{10}(f_c) - 13.82\log_{10}(h_t) + (44.9 - 6.55\log_{10}(h_t))\log_{10}(d_1) \\ &+ 64.5\log_{10}(d_2) - ((1.1\log_{10}(f_c) - 0.7)(h_m + 10^{-3}) - (3.12\log_{10}(f_c) - 1.6)) \\ &- \sigma(dBm) - \left(4\log_{10}\left[\frac{f_c}{28}\right]^2 + 10.8\right) \quad . \end{aligned} \tag{3.29}$$

Equation 3.29 is dependent on the frequency of operation, f_c, height of the transmitter, h_t, the height of the receiver, h_m, the RCS of the point of reflection σ and the path taken by the signal to reach the receiver. The path is divided into two parts d_1 and d_2. The derived equation shows the relationship of the path loss to various parameters. Now we present a sample calculation of the link budget analysis.

3.2.2 Sample calculation

A sample link budget analysis calculation is presented here. This calculation is based on a particular set of parameters obtained from [99].

Table 3.1: Parameters used for the link budget analysis calculations.

The transmit power	50 Watts (47 dBm) Downlink and 1 Watts (30 dBm) Uplink
Bandwidth	20 MHz at a frequency of 2300 MHz
Bit Rate	100 Mbps Downlink and 50 Mbps Uplink for a 20 MHz channel
Symbol Rate	50 Msps Downlink and 25 Msps Uplink for Quadrature Phase Shift Key (QPSK) modulation
Noise Temperature	300 K
Io/No	−2.2 dB
Receiver Sensitivity	−104 dBm
BTS Antenna Gain	0 dBi
User Equipment (UE) Antenna Gain	10 dBi
BTS Antenna Height	15 m
UE Antenna Height	2 m
Transmission Loss	2 dB
Fading Margin	9.8 dBm
Reception Losses	14 dBm

3.3 Signal Model

The link budget analysis provides a basic understanding of the gains and losses of the wireless communication system. Although it should be noted that the concept of GSM-CommSense system is to design a non-radiating, receive-only system. An understanding of the link budget provides us with an idea of the range of operation.

Since the idea of GSM-CommSense is to extract the channel information from the received signal it is conceived to work on the principle of finding the time difference between the direct path signal and the scattered path signal. In this

Summary

	Transmit power	47.0 dBm
+	Gains	10.0 dB
-	Losses	191.4 dB
=	Received power	-134.4 dBm
-	Noise + interference power	-94.8 dBm
=	Median received SNR	-39.6 dB
+	Processing gain	-3.0 dB
=	Median received EbNo	-42.6 dB
-	Required EbNo	-10.2 dB
=	Excess	-32.4 dB
-	Margin	9.8 dB
=	SURPLUS	-42.2 dB
	Desired link reliability	90 %
	Effective link reliability	0 %
	Specified link distance	20.000 km
	Distance for desired reliability	1.469 km
	Reliability mode	fading only

Downlink Budget

	Transmit power	30.0 dBm
+	Gains	10.0 dB
-	Losses	171.9 dB
=	Received power	-131.9 dBm
-	Noise + interference power	-94.8 dBm
=	Median received SNR	-37.1 dB
+	Processing gain	-3.0 dB
=	Median received EbNo	-40.1 dB
-	Required EbNo	-10.2 dB
=	Excess	-30.0 dB
-	Margin	9.8 dB
=	SURPLUS	-39.7 dB
	Desired link reliability	90 %
	Effective link reliability	0 %
	Specified link distance	20.000 km
	Distance for desired reliability	2.372 km
	Reliability mode	fading only

Uplink Budget

Figure 3.2: Link Budget calculation summary.

section, we outline a basic signal model that can be used to better understand the GSM-CommSense system.

3.3.1 Transmit signal model

The closed form representation of a narrow bandpass signal is given as:

$$s(t) = g(t)\cos(\omega_c t + \phi) \quad . \tag{3.30}$$

Here, $s(t)$ represents the transmit signal, $g(t)$ is the natural envelope of $s(t)$. $\omega_c = 2\pi f_c$ is the angular frequency and ϕ is the instantaneous phase.

The geometry of a two path ground reflection model used for performing the analysis of the system is shown in Figure 3.3.

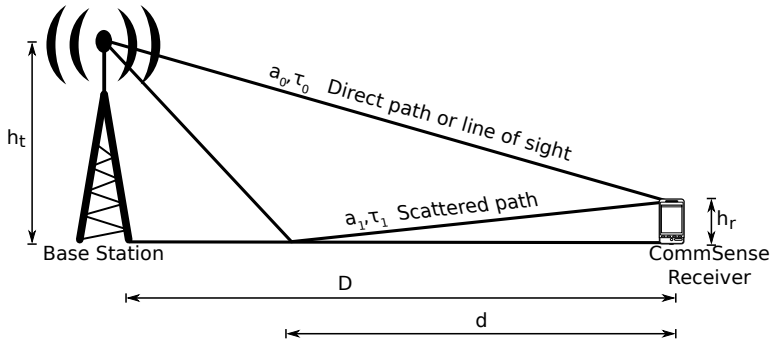


Figure 3.3: Representation of a two path ground reflection model.

The complex envelope of the signal $s(t)$ is given as:

$$u(t) = g(t)e^{j\phi} \quad . \tag{3.31}$$

Using the complex envelope another form of the transmit signal model is devised as given as:

$$s(t) = \Re\{u(t)e^{j\omega_c t}\} \quad . \tag{3.32}$$

Expanding Equation 3.32 provides the third representation of the signal given as:

$$s(t) = \frac{1}{2}\left(u(t)e^{j\omega_c t} + u^*(t)e^{-j\omega_c t}\right) \quad . \tag{3.33}$$

Here, $u^*(.)$ is the complex conjugate of the complex envelope $u(t)$. This representation will be used to calculate the matched filter response of the system.

3.3.2 Receive signal model

A signal, transmitted over a channel, passes through multiple obstacles in its path. These obstacles introduce scattering also known as multipath effect as well as some noise which in this case is assumed to be Additive White Gaussian Noise (AWGN) and is denoted by $w(t)$. The scattering introduces two main factors in the signal namely attenuation factor, a_n, and propagation delay, τ_n. The subscript n represents the corresponding path number considering the signal

scatters through more than one path. The received bandpass signal, passing through a multipath channel, is given in Equation 3.34. [47, 100]

$$x(t) = \sum_n a_n s(t - \tau_n) \tag{3.34}$$

Substituting Equation 3.30 in Equation 3.34 and expanding it gives the received signal model with AWGN as:

$$x(t) = \sum_n a_n g(t - \tau_n) \cos(\omega_c(t - \tau_n) + \phi) + w(t) \quad . \tag{3.35}$$

The complex envelope representation of the received signal is presented as:

$$x(t) = \frac{1}{2} \sum_n a_n \left[u(t - \tau_n) e^{j\omega_c(t-\tau_n)} + u^*(t - \tau_n) e^{-j\omega_c(t-\tau_n)} \right] + w(t) \quad . \tag{3.36}$$

3.4 Matched Filter

A matched filter is obtained by correlation of a known signal with an unknown signal. The objective of this filter is to increase the SNR and thereby detect the presence of the known signal within the unknown signal. In this section, we calculate the matched filter for the signal representations in Section 3.3, for the two path ground reflection model shown in Figure 3.3 and then generalized for n-paths.

3.4.1 Calculation

The matched filter calculations shown in [101] for a complex transmit signal, $s(t)$, derives a time domain filter response, $h(t)$, as:

$$h(t) = Ks^*(t_o - t) \quad . \tag{3.37}$$

The arbitrary constant K has a dimension of $(\mathrm{Vs})^{-1}$ and t_o is a predetermined delay which is greater than or equal to the period of the signal.

Output of a matched filter $s_o(t)$ in time domain is calculated as the convolution of the received signal with the matched filter response as:

$$s_o(t) = x(t) * h(t) \quad , \tag{3.38}$$

$$s_o(t) = K \int_{-\infty}^{\infty} x(\alpha) s^*(t_o - (t - \alpha)) d\alpha \quad . \tag{3.39}$$

Expanding the convolution in Equation 3.39, substituting the received signal model Equation 3.35 and the matched filter response Equation 3.37 provides the output of the matched filter as:

$$s_o(t) = \frac{K}{4} \sum_n a_n \int_{-\infty}^{\infty} \left[u(\alpha - \tau_n(\alpha)) e^{j\omega_c(\alpha - \tau_n)} + u^*(\alpha - \tau_n(\alpha)) e^{-j\omega_c(\alpha - \tau_n)} \right] \left[u^*(t_o - (t - \alpha)) e^{-j\omega_c(t_o - (t-\alpha))} + u(t_o - (t - \alpha)) e^{j\omega_c(t_o - (t-\alpha))} \right] d\alpha \ . \tag{3.40}$$

Steps involved in expanding Equation 3.40 and solving it is given as:

$$\begin{aligned} s_o(t) = \frac{K}{4}\sum_n a_n \Bigg[& \int_{-\infty}^{\infty} u(\alpha-\tau_n(\alpha))u^*(t_o-(t-\alpha))e^{j\omega_c(\alpha-\tau_n-t_o+(t-\alpha))}d\alpha \\ & + \int_{-\infty}^{\infty} u^*(\alpha-\tau_n(\alpha))u^*(t_o-(t-\alpha))e^{-j\omega_c(\alpha-\tau_n+t_o-(t-\alpha))}d\alpha \\ & + \int_{-\infty}^{\infty} u(\alpha-\tau_n(\alpha))u(t_o-(t-\alpha))e^{j\omega_c(\alpha-\tau_n+t_o-(t-\alpha))}d\alpha \\ & \int_{-\infty}^{\infty} u^*(\alpha-\tau_n(\alpha))u(t_o-(t-\alpha))e^{-j\omega_c(\alpha-\tau_n+t_o-(t-\alpha))}d\alpha \Bigg] \quad , \end{aligned} \tag{3.41}$$

$$\begin{aligned} s_o(t) = \frac{K}{4}\sum_n a_n \Bigg[& e^{j\omega_c(t-\tau_n-t_o)} \int_{-\infty}^{\infty} u(\alpha-\tau_n(\alpha))u^*(t_o-(t-\alpha))d\alpha \\ & + e^{-j\omega_c(t-\tau_n-t_o)} \int_{-\infty}^{\infty} u^*(\alpha-\tau_n(\alpha))u(t_o-(t-\alpha))d\alpha \Bigg] \\ & + \Bigg[e^{j\omega_c(t+\tau_n-t_o)} \int_{-\infty}^{\infty} e^{-2j\omega_c\alpha} u^*(\alpha-\tau_n(\alpha))u^*(t_o-(t-\alpha))d\alpha \\ & + e^{-j\omega_c(t+\tau_n-t_o)} \int_{-\infty}^{\infty} e^{2j\omega_c\alpha} u(\alpha-\tau_n(\alpha))u(t_o-(t-\alpha))d\alpha \Bigg] \quad . \end{aligned} \tag{3.42}$$

Using the formula $(a+jb)+(a-jb)=2a=2\Re\{a+jb\}$ in Equation 3.42 gives the matched filter output for multipath received signal including the direct path signal as:

$$\begin{aligned} s_o(t) = \frac{K}{2}\sum_n a_n \Re\Bigg\{ & e^{j\omega_c(t-\tau_n-t_o)} \int_{-\infty}^{\infty} u(\alpha-\tau_n)u^*(t_o-(t-\alpha))d\alpha \\ & + e^{j\omega_c(t+\tau_n-t_o)} \int_{-\infty}^{\infty} u^*(\alpha-\tau_n)u^*(t_o-(t-\alpha))e^{-j2\omega_c\alpha}d\alpha \Bigg\} \quad . \end{aligned} \tag{3.43}$$

The second integral in Equation 3.43 is the Fourier transform of $u^*(\alpha-\tau_n)u^*(t_o-(t-\alpha))$ at $\omega=\omega_c$ and if this signal is narrow bandpass around ω_c then the

spectrum of its complex envelope is cut off well below ω_c. The remaining of the expression is:

$$s_o(t) \approx \Re\left\{ \left[\frac{K}{2} \sum_n a_n e^{-j\omega_c(t_o-\tau_n)} \int_{-\infty}^{\infty} u(\alpha-\tau_n) u^*(t_o-(t-\alpha)) d\alpha \right] e^{j\omega_c t} \right\} \quad . \tag{3.44}$$

The complex envelope $u_o(t)$ of the matched filter output signal $s_o(t)$ can be represented as Equation 3.45. Here K_u is a constant for a particular path as represented in Equation 3.46.

$$u_o(t) = \sum_n a_n K_u \int_{-\infty}^{\infty} u(\alpha-\tau_n) u^*(t_o-(t-\alpha)) d\alpha \tag{3.45}$$

$$K_u = \frac{K}{2} e^{-j\omega_c(t_o+\tau_n)} \tag{3.46}$$

Using Equation 3.44,Equation 3.45 and Equation 3.46 the matched filter output can be represented as:

$$s_o(t) = \Re\{u_o(t) e^{j\omega_c t}\} \quad . \tag{3.47}$$

To calculate the value of the complex envelope of the signal a simple rectangular natural envelope $g(t)$ was chosen with the length of the rectangular function given by T as:

$$g(t) = \text{rect}\left(\frac{t}{T}\right) \quad . \tag{3.48}$$

In order to calculate the matched filter output with zero Doppler shift, a constant phase ϕ has to be maintained. Thus, the complex envelope can be written as:

$$u_o(t) = \sum_n K_u e^{j\phi} \begin{cases} t - t_o & -\frac{T}{2} < t \leq 0 \\ T - t + t_o & 0 < t < \frac{T}{2} \end{cases} \quad . \tag{3.49}$$

3.5 Ambiguity Function

Ambiguity function is a two-dimensional function of time delay and the Doppler shift showing the changes in the received signal due to the matched filter.

3.5.1 Calculation

The ambiguity function $\chi(t, \nu)$ of a basic radar system in terms of its complex envelope is given as:

$$\chi(t, \nu) = \Re\{u_{t,\nu}(t)e^{j\omega_c t}\} \quad . \tag{3.50}$$

Here, $u_{t,\nu}(t)$ gives the complex envelope of the signal as shown in Equation 3.51 with a Doppler shift, ν.

$$u_{t,\nu}(t) = u_o(t)e^{j2\pi\nu t} \tag{3.51}$$

Calculating the value of $u_{t,\nu}(t)$ for the GSM-CommSense system using Equation 3.39, Equation 3.50 and Equation 3.51 gives:

$$u_{t,\nu}(t) = \sum_n K_u e^{j\phi} \begin{cases} e^{j2\pi\nu(t-t_o)} & -\frac{T}{2} < t \leq 0 \\ e^{j2\pi\nu(T-t+t_o)} & 0 < t < \frac{T}{2} \end{cases} \quad . \tag{3.52}$$

3.5.2 Simulation

The Ambiguity function for different types of transmit waveform is presented in this section. Each of these waveform is simulated when the received signal contains a direct path and one scattered path. Here a_0, τ_0 represent the amplitude and time delay of the direct path signal and a_1, τ_1 represent the amplitude and time delay of one scattered path signal with respect to the transmit signal respectively. No noise component is considered in this simulation.

Each set of plots is simulated using a specific modulation scheme. This serves

the purpose of proving the concept that the ambiguity function gets affected if multipath components are available in the received signal irrespective of the modulation or the type of signal. It must be noted that the ambiguity function is not symmetrical along the zero delay line which occurs due to the presence of an added scattered path in the received signal waveform.

Time domain representation of one pulse of the transmit signal is shown in Figure 3.4. The cosine wave is passed directly without any pulse shaping filter, the barker code is passed through a sinc filter and the training sequence is passed through a Gaussian filter. The pulse shaping filters constrain the bandwidth of the signals and provide higher fidelity results.

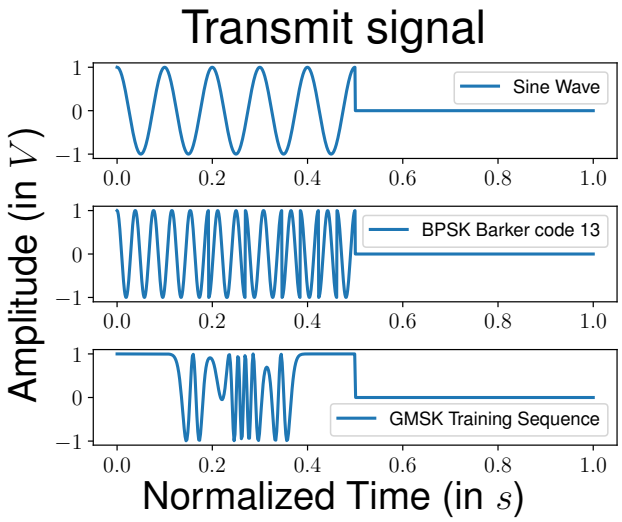


Figure 3.4: Normalized time domain model of the transmit signals used for the simulation results presented here.

3.5.2.1 Sinusoidal signal

This set of simulation results use a cosine wave without any modulation. The transmit waveform is represented as:

$$\text{Tx} = \Re\{e^{j\omega_c t}\} \quad . \tag{3.53}$$

An unmodulated cosine wave is used to show that there is difference in the received signal due to the presence of a scattered path component but not as a result of the modulation scheme used. For the purpose of this simulation the signal strength and time delay of the direct path signal are kept constant at $a_0 = 0.5$ V and $\tau_0 = 0.1$ μs respectively.

Figure 3.5 contains the ambiguity function plots showing different signal strength and delay conditions for the scattered signal.

Figure 3.5a shows the behaviour of the ambiguity function when the scattered path delay, τ_1, is 0.12 μs with a considerably low signal strength, a_1 when compared to the direct signal strength, a_0. The signal is mostly focused at the center of the plot and the edges are not very distinguishable.

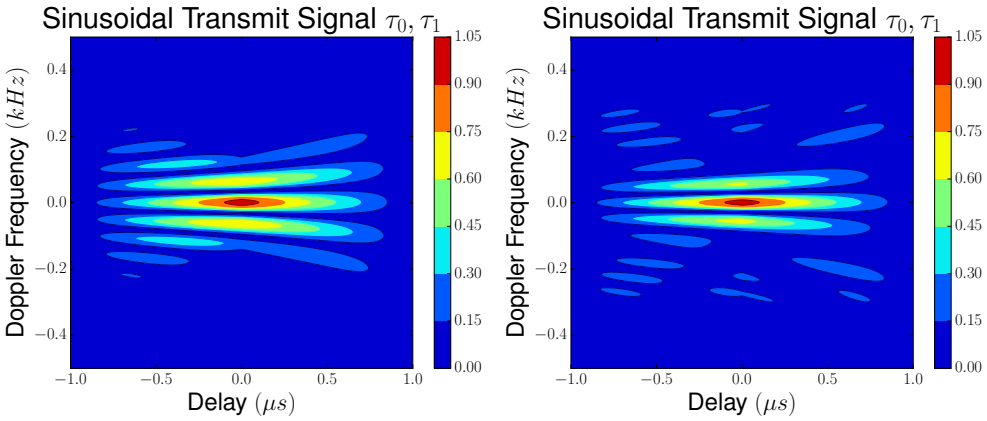
Figure 3.5b shows the plot where the delay τ_1 is 0.14 μs and a_1 is close to half of a_0. The ambiguity function in this case has started spreading more, which is due to the presence of scattered signal.

Figure 3.5c shows the plot where the delay τ_1 is 0.15 μs and a_1 is a little more than half of a_0. Since the time variation between Figure 3.5b and 3.5c is 0.01 μs the plots are similar despite the change in the scattered signal strength. This means that the ambiguity function gets affected significantly by a minor change in the delay parameter as opposed to the signal strength.

Figure 3.5d shows the plot where the delay τ_1 is 0.17 μs and a_1 is very almost similar to a_0. Here the spread of the ambiguity function is similar to the Figure 3.5b but the differences in the signal can be better resolved.

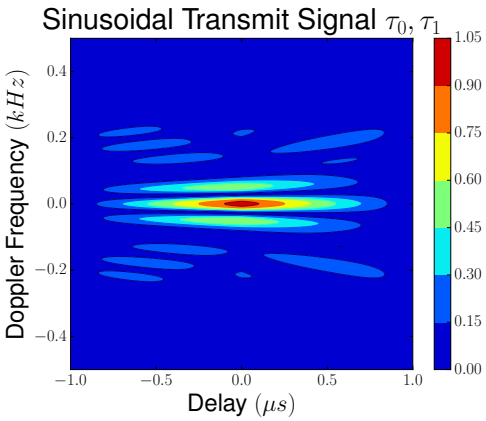
Figure 3.5e shows the plot where the delay τ_1 is 0.18 μs and a_1 is further closer to a_0. Here the Doppler spread of the ambiguity function is clearly visible and the differences in the signal can be resolved much finer.

Figure 3.5f shows the plot where the τ_1 is 0.19 μs and a_1 is exactly same as a_0. Here the fine Doppler spread is visible but the separation due to the signal strength of each path is not possible as the values are same.

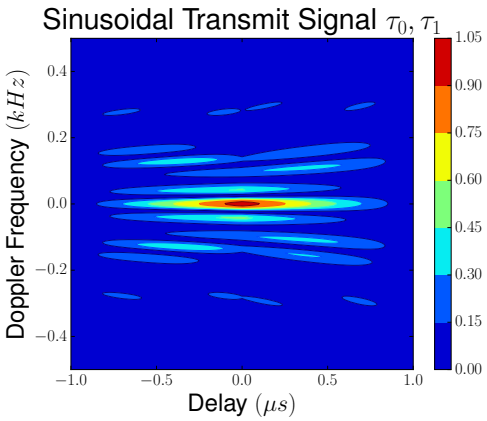


(a) $a_1 = 0.11$ V and $\tau_1 = 0.12$ μs

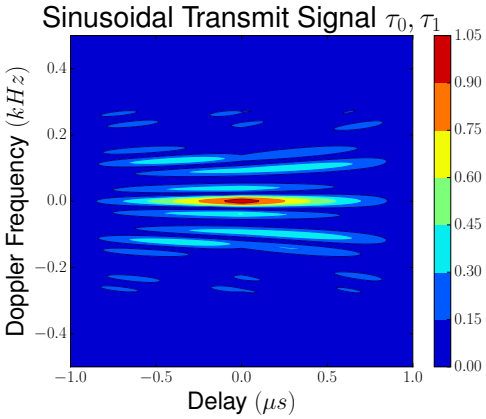
(b) $a_1 = 0.22$ V and $\tau_1 = 0.14$ μs



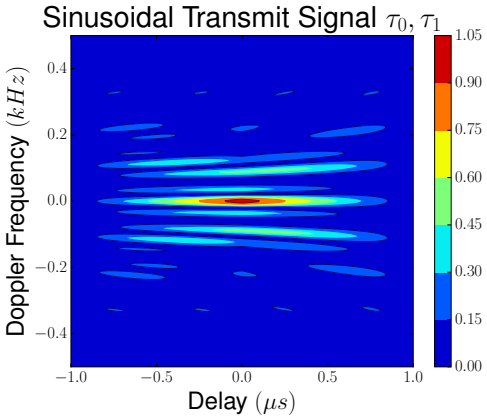
(c) $a_1 = 0.27$ V and $\tau_1 = 0.15$ μs



(d) $a_1 = 0.38$ V and $\tau_1 = 0.17$ μs



(e) $a_1 = 0.44$ V and $\tau_1 = 0.18$ μs



(f) $a_1 = 0.5$ V and $\tau_1 = 0.19$ μs

Figure 3.5: Normalized ambiguity function for cosine wave transmit signal in linear scale. The direct signal parameters are $a_0 = 0.5$ V, $\tau_0 = 0.1$ μs.

3.5.2.2 Barker code

This simulation is to show the effect on the signal when the transmit waveform is a 13−bit Barker code with Binary Phase Shift Key (BPSK) modulation. The 13−bit barker code used here is $[1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1]$ and the constant direct path signal parameters are $a_0 = 0.5$ V and $\tau_0 = 0.01$ μs.

Figure 3.6 contains the ambiguity function plots showing different signal strength and delay conditions for the scattered signal.

Figure 3.6a shows the behaviour of the ambiguity function when the scattered path delay, τ_1 is 0.012 μs and the scattered signal strength, a_1 is very low compared to the direct signal strength, a_0. The signal is mostly focused at the center of the plot and the edges are not very distinguishable.

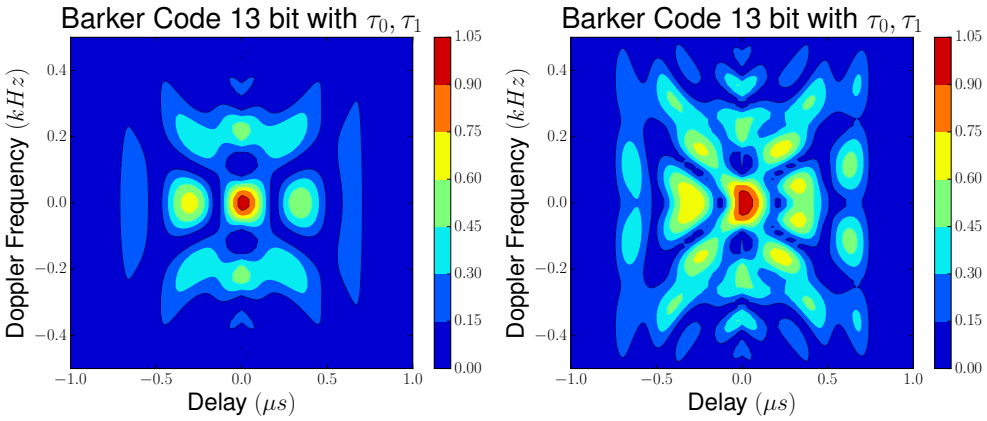
Figure 3.6b shows the plot where the delay τ_1 is 0.014 μs and a_1 is close to half of a_0. The ambiguity function in this case has started spreading more which is due to the presence of scattered signal.

Figure 3.6c shows the plot where the delay τ_1 is 0.015 μs and a_1 is a little over half of a_0. Even though the difference between τ_0 and τ_1 is incredibly low the ambiguity function changes significantly which means that BPSK modulated Barker code can resolve the signals at a very low difference in time. This mostly happens due to the correlation properties of the Barker code.

Figure 3.6d shows the plot where the delay τ_1 is 0.017 μs and a_1 is very close to a_0. Here the difference in the direct path and the scattered path is visible.

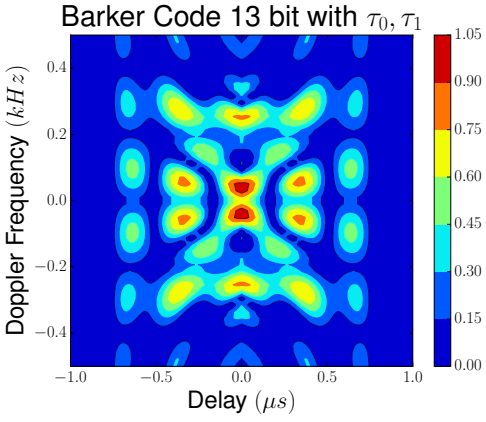
Figure 3.6e shows the plot where the delay τ_1 is 0.018 μs and a_1 is further closer to a_0. Here the difference in the direct path and the scattered path is still visible.

Figure 3.6f shows the plot where the delay τ_1 is 0.019 μs and a_1 is exactly same as a_0. Here the difference due to the amplitude of the direct path and the scattered path is not visible but the Doppler spread is clearly observable.

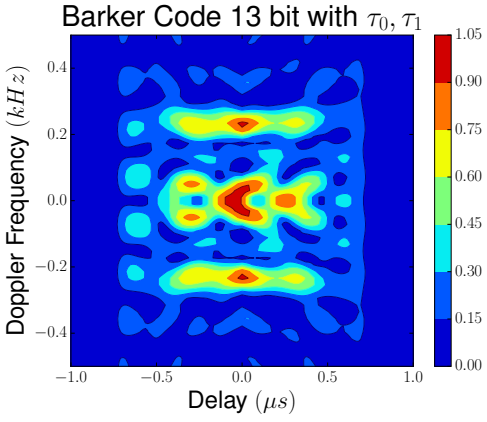


(a) $a_1 = 0.11$ V and $\tau_1 = 0.012$ μs

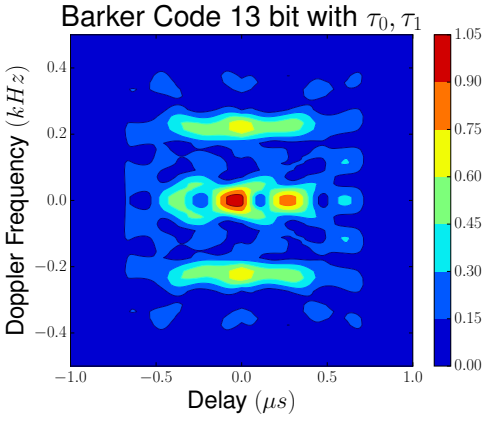
(b) $a_1 = 0.22$ V and $\tau_1 = 0.014$ μs



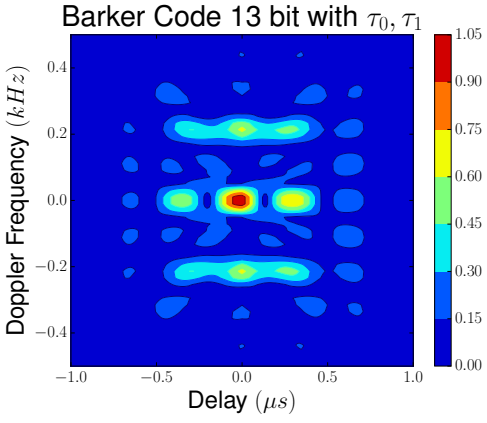
(c) $a_1 = 0.27$ V and $\tau_1 = 0.015$ μs



(d) $a_1 = 0.38$ V and $\tau_1 = 0.017$ μs



(e) $a_1 = 0.44$ V and $\tau_1 = 0.018$ μs



(f) $a_1 = 0.5$ V and $\tau_1 = 0.019$ μs

Figure 3.6: Normalized ambiguity function for 13 bit Barker code based transmit signal in linear scale. The direct signal parameters are $a_0 = 0.5$ V, $\tau_0 = 0.01$ μs

3.5.2.3 GSM training sequence

This simulation is performed on one of the training sequences transmitted by GSM system. The transmit signal is modulated based on GMSK as it is implemented in GSM systems. The training sequence used in this simulation is $[0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, \ 1, 0, 1, 1, 1]$ and the constant direct path parameters are $a_0 = 0.5$ V and $\tau_0 = 0.01$ μs.

Figure 3.7 contains the ambiguity function plots varying the signal strength and the delay of the scattered signal.

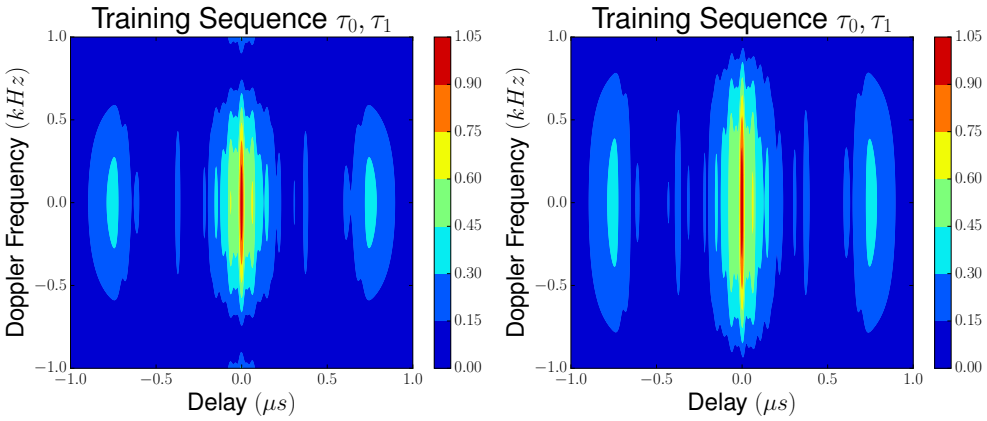
Figure 3.7a shows behaviour of the ambiguity function when the scattered signal delay, τ_1 is 0.012 μs and the scattered signal strength, a_1 is very low as compared to the direct signal strength, a_0. The signal is mostly focused on the zero delay line of the plot and the edges are not very distinguishable.

Figure 3.7b shows the plot where the delay τ_1 is 0.014 μs and a_1 is close to half of a_0. The ambiguity function in this case has started spreading more which is due to the changes of signal strength and delay of the scattered signal.

Figure 3.7c shows the plot where the delay τ_1 is 0.015 μs and a_1 is a little over half of a_0. Even though the difference between τ_0 and τ_1 is low, the ambiguity function spreads in Doppler. It should be noted here, that in this specific case the ambiguity function is symmetrical across the zero delay line as opposed to all the other cases. Also, the ambiguity function is spread to show two peaks in the same region which means that the contribution of the direct signal and the scattered signal is separated in this particular case, creating the symmetry and the peaks.

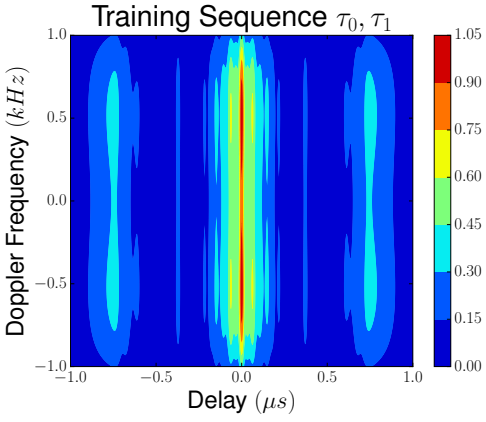
Figure 3.7d shows the plot where the delay τ_1 is 0.017 μs and a_1 is very close to a_0. The symmetry from Figure 3.7c is no more visible and only one peak is visible.

Figure 3.7e shows the plot where the delay τ_1 is 0.018 μs and a_1 is even closer to a_0. The signal strength is not clearly separable, although the difference due to the delay is visible.

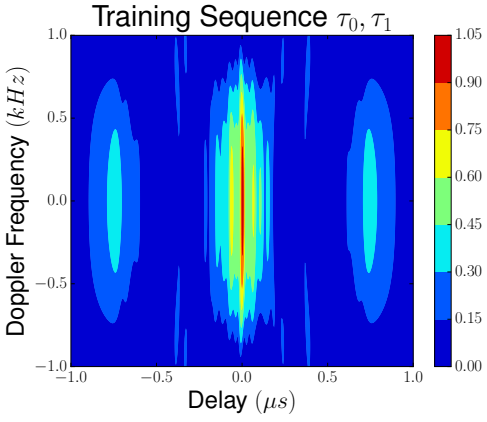


(a) $a_1 = 0.11$ V and $\tau_1 = 0.012$ μs

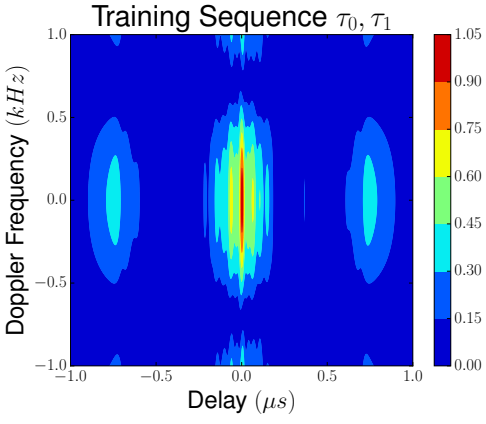
(b) $a_1 = 0.22$ V and $\tau_1 = 0.014$ μs



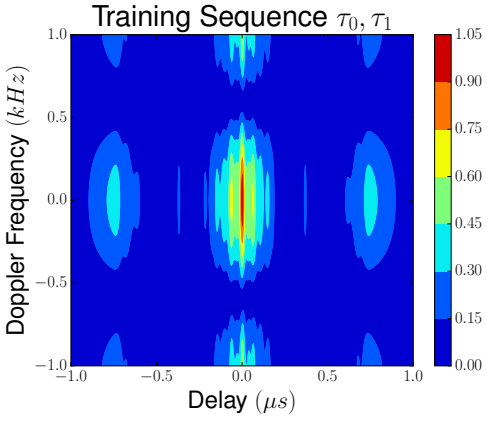
(c) $a_1 = 0.27$ V and $\tau_1 = 0.015$ μs



(d) $a_1 = 0.38$ V and $\tau_1 = 0.017$ μs



(e) $a_1 = 0.44$ V and $\tau_1 = 0.018$ μs



(f) $a_1 = 0.5$ V and $\tau_1 = 0.019$ μs

Figure 3.7: Normalized ambiguity function for training sequence based transmit signal in linear scale. The direct signal parameters are $a_0 = 0.5$ V, $\tau_0 = 0.01$ μs.

Figure 3.7f shows the plot where the delay τ_1 is 0.019 μs and a_1 is exactly same as a_0. The separation due to the signal strength is not visible, but the difference due to delay is visible if observed closely.

Here, we presented the ambiguity function calculations for a two path received signal model along with the simulation results. Three different input signal types with different modulation schemes are used to show that the ambiguity function of a signal changes when a scattered path information is introduced in the received signal model. An unmodulated sinusoidal signal proves that the change in the ambiguity function is not due to the modulation scheme. The modulated signals show the effect of the scattered path in the ambiguity function for the respective modulation schemes.

Figures 3.5, 3.6 and 3.7 contain the necessary plots to conclude that the delay of the scattered signal plays a more significant role in the ambiguity function as compared to the signal strength.

3.6 Cramér-Rao Lower Bound

CRLB gives the minimum estimated variance for an unbiased estimator by calculating the inverse of the Fisher information. In this case, we calculate the minimum variance of the time delay in a scattered path in the presence of direct path within the signal. The final calculated CRLB is simulated and the results are discussed.

3.6.1 Calculation

Using the received signal model from Equation 3.35, we shall now calculate the CRLB [102] for $n = 0, 1$. This includes the direct path and one scattered path

signal, respectively. The modified receive signal model is given as:

$$\begin{aligned} x(t) =& a_0 g(t-\tau_0)\cos(\omega_c(t-\tau_0)+\phi)+ \\ & a_1 g(t-\tau_1)\cos(\omega_c(t-\tau_1)+\phi)+w(t) \quad . \end{aligned} \tag{3.54}$$

To perform the CRLB calculations a two path model is used where a_0 is the direct path attenuation factor, τ_0 the direct path delay, a_1 is the first scattered path attenuation factor and τ_1 the first scattered path delay.

In calculating the CRLB for τ_n we assume that τ is the only unknown parameter and $w(t)$ is zero mean ($\mu = 0$) and variance of σ^2. Length of the sampled data is given by N and the natural envelope $g(t)$ of the signal is considered constant for the entire duration of sampling (thus $g(t) = 1$).

The likelihood function of $x(t)$ with respect to τ_1 is given in Equation 3.55. The term $\exp\{.\}$ represents $e^{\{.\}}$, it is used to avoid fraction in superscript. The log-likelihood function is shown in Equation 3.56.

$$\begin{aligned} p(x;\tau_1) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\Bigg\{ -\frac{1}{2\sigma^2} \sum_{m=0}^{N-1} \Bigg[x[m] - \Big(& a_0\cos(\omega_c(m-\tau_0)+\phi) \\ & + a_1\cos(\omega_c(m-\tau_1)+\phi)\Big)\Bigg]^2 \Bigg\} \end{aligned} \tag{3.55}$$

$$\begin{aligned} \ln p(x,\tau_1) = -\frac{1}{2\sigma^2} \sum_{m=0}^{N-1} \Bigg[x[m] - \Big(& a_0\cos(\omega_c(m-\tau_0)+\phi) \\ & + a_1\cos(\omega_c(m-\tau_1)+\phi)\Big)\Bigg]^2 - \frac{N}{2}\ln(2\pi\sigma^2) \end{aligned} \tag{3.56}$$

Differentiating the log-likelihood function with respect to the scattered path τ_1

gives:

$$\frac{d\ln p(x,\tau_1)}{d\tau_1} = \frac{a_1\omega_c}{\sigma^2}\sum_{m=0}^{N-1}\Bigg[x[m]\sin(\omega_c(m-\tau_1)+\phi) - \frac{a_1}{2}\sin(2\omega_c(m-\tau_1)+2\phi) - \frac{a_0}{2}\Big(\sin(\omega_c(\tau_0-\tau_1)+\sin(\omega_c(m-\tau_1)) + \omega_c(m-\tau_0)+2\phi)\Big)\Bigg]. \tag{3.57}$$

To calculate the CRLB the log-likelihood function needs to be differentiated twice with respect to the time parameter, in this case τ_1. The second order differentiation provides:

$$\frac{d^2\ln p(x,\tau_1)}{d\tau_1^2} = -\frac{a_1\omega_c^2}{\sigma^2}\sum_{m=0}^{N-1}\Bigg[x[m]\cos(\omega_c(m-\tau_1)+\phi) - a_1\cos(2\omega_c(m-\tau_1)+2\phi) - \frac{a_0}{2}\Big(\cos(\omega_c(\tau_0-\tau_1)) + \cos(\omega_c(m-\tau_1)+\omega_c(m-\tau_0)+2\phi)\Big)\Bigg]. \tag{3.58}$$

To obtain the expected value $E(.)$ of Equation 3.58, $x[m]$ should be substituted by $a_0\cos(\omega_c(m-\tau_0)+\phi) + a_1\cos(\omega_c(m-\tau_1)+\phi)$ as shown below.

$$\begin{aligned}-E\left[\frac{d^2\ln p(x,\tau_1)}{d\tau_1^2}\right] = \frac{a_1\omega_c^2}{\sigma^2}\sum_{m=0}^{N-1}\Bigg[&(a_0\cos(\omega_c(m-\tau_0)+\phi) \\ &+ a_1\cos(\omega_c(m-\tau_1)+\phi))\cos(\omega_c(m-\tau_1) \\ &+\phi) - a_1\cos(2\omega_c(m-\tau_1)+2\phi) - \frac{a_0}{2} \\ &\Big(\cos(\omega_c(\tau_0-\tau_1)) + \cos(\omega_c(m-\tau_1)+ \\ &\omega_c(m-\tau_0)+2\phi)\Big)\Bigg]\end{aligned} \tag{3.59}$$

$$
\begin{aligned}
-E\left[\frac{d^2 \ln p(x,\tau_1)}{d\tau_1^2}\right] = \frac{a_1\omega_c^2}{\sigma^2}\sum_{m=0}^{N-1}\Big[a_1\cos^2(\omega_c(m-\tau_1)+\phi) \\
+ a_0\cos(\omega_c(m-\tau_0)+\phi)\cos(\omega_c(m-\tau_1) \\
+\phi) - a_1\cos(2\omega_c(m-\tau_1)+2\phi) - \frac{a_0}{2} \\
\Big(\cos(\omega_c(\tau_0-\tau_1)) + \cos(\omega_c(m-\tau_1)+ \\
\omega_c(m-\tau_0)+2\phi)\Big)\Big]
\end{aligned}
\tag{3.60}
$$

$$
\begin{aligned}
-E\left[\frac{d^2 \ln p(x,\tau_1)}{d\tau_1^2}\right] = \frac{a_1\omega_c^2}{\sigma^2}\sum_{m=0}^{N-1}\Big[\frac{a_1}{2}\big[1+\cos(2\omega_c(m-\tau_1) \\
+2\phi)\big] - a_1\cos(2\omega_c(m-\tau_1)+2\phi)\Big]
\end{aligned}
\tag{3.61}
$$

Upon simplification of Equation 3.61 we get:

$$
-E\left[\frac{d^2 \ln p(x,\tau_1)}{d\tau_1^2}\right] = \frac{a_1^2\omega_c^2}{2\sigma^2}\sum_{m=0}^{N-1}\Big[1-\cos(2\omega_c(m-\tau_1)+2\phi)\Big] \tag{3.62}
$$

The negative of the expectation value, also known as the fisher information, is given in Equation 3.62. In this case, since the signal is periodic and the sum is zero, the expectation value after performing the summation is given as:

$$
-E\left[\frac{d^2 \ln p(x,\tau_1)}{d\tau_1^2}\right] = \frac{a_1^2\omega_c^2 N}{2\sigma^2} = I(\hat{\tau_1}) \quad . \tag{3.63}
$$

In Equation 3.63, $I(.)$ is the Fisher information and $\hat{\tau_1}$ is the estimated value of the parameter τ_1. The CRLB, also is represented by the variance $\text{var}(\hat{\tau_1})$ of the

scattered path and is given as the inverse of the Fisher information as:

$$\text{var}(\hat{\tau_1}) \geq I(\hat{\tau_1})^{-1} = \frac{2\sigma^2}{a_1^2 \omega_c^2 N} \quad . \tag{3.64}$$

3.6.2 Simulation

The CRLB calculated in Equation 3.64 shows that the $\text{var}(\hat{\tau_1})$ depends on the amplitude a_1 of the scattered signal. The amplitude can be calculated based on the two path model from Figure 3.3 in terms of height of the transmitter h_t, height of the receiver h_r. Separation between the transmitter and receiver is denoted by D and the distance of the point from the receiver is where the signal reflects d.

$$a_1^2 = \frac{Z}{(d^2 + h_r^2)(h_t^2 + (D-d)^2)} \tag{3.65}$$

The value of CRLB depends on a_1^2 which is represented in terms of the distance from the receiver d as shown in Equation 3.65. Here, Z is a constant dependent on transmit power per unit area of the transmitter, P_t, gain of the transmitter, G_t, the receiver gain, G_r, and the reflectivity of the surface from which the signal reflects, ρ, using the formula in linear scale $Z = P_t G_t G_r \rho$.

We can consider a case where the transmit power of a GSM base station is 10 Watts (40 dBm) and the receiver antenna gain is 3 dB. The ground surface RCS for a farmland and a dessert as given in [103] is in the range of -15 dB to -45 dB. Calculating the range of Z from the above mentioned values yields 0.63 mW $< Z <$ 630 mW approximately.

For the purpose of this calculation the value of Z is assumed to be unity for the ease of calculations as this will act as a scaling factor for the CRLB and will not affect the overall performance.

The simulated CRLB is presented in Figure 3.8. This contains the variance of time τ_1 with respect to the distance of the target from the receiver. It shows the

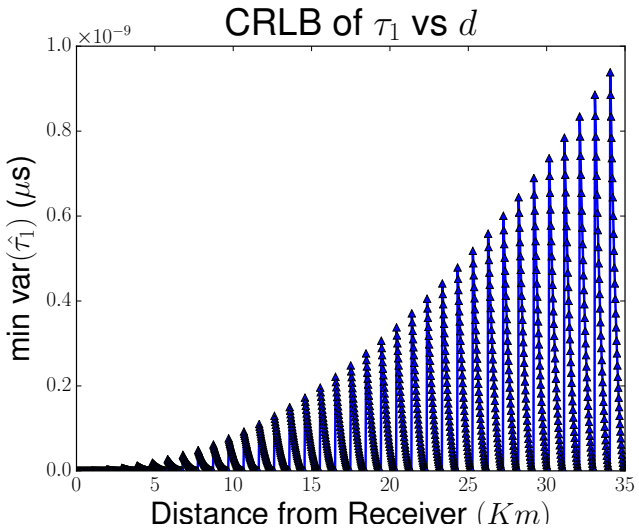


Figure 3.8: CRLB simulation is to show the minimum variance of time delay for the scattered path in the presence of direct path with respect to the distance of the reflection point from the receiver.

variance of a scattered path time in the presence of the direct path signal. The plot contains peaks which occur when the signals are out of phase and there is a phase difference in the direct path signal and the scattered path signal and the troughs occur when the signals are in phase.

The parameters used for this particular simulation are given in Table 3.2. The height of transmitter h_t is chosen to match a generic base station height, the height of receiver h_r is set as all the captures using the designed system based on this technology is taken keeping the receiver at a height of 1.12 m. Maximum cell size of a base station is about 35 km thus the separation D is kept at that value, d is a variable as the plot is based on the distance and time. The value of N changes the scale of the CRLB thus it is kept at a value of 100. The variance of noise is set to 1 as this also acts as a scaling factor for the CRLB and changing the noise variance does not affect the overall structure of the graph. GSM in South Africa works on 900 MHz. Therefore, for the purpose of this calculation the operating frequency is set at 900 MHz.

The purpose of these simulations is to show the trend of the performance of GSM-CommSense system with respect to distance of the point of reflection from

Table 3.2: Parameters used to simulate CRLB.

Parameter	Value
Transmitter Height (ht)	30 m
Receiver height (hr)	1.12 m
Separation between transmitter and receiver (D)	35 km
Distance of reflection point from receiver (d)	0 to D km
Length of sampled data (N)	100
Variance of noise (σ^2)	1
Frequency (f_c)	900 MHz

the receiver. GSM-CommSense is an under-determined system. This means that the amount of data received by the system will not be enough to provide an overview of the surroundings. Thus, the conventional resolution, as defined for the radar systems, is not applicable. The trend shown in Figure 3.8 can be used to calculate the trend of degradation of the sensitivity of GSM-CommSense system with range. It is observed that the minimum variance, $\hat{\tau_1}$, is of order 10^{-9} μs which is due to the presence of one scattered path and a direct path information, with the constrains at the beginning of this section. The variance of time shown here is purely analytical and the events that can be distinguished will be dependent on factors such as the type of signal under test, the surroundings of the receiver, the sensitivity of the actual hardware system, to name a few.

3.7 Summary

The analysis presented in this chapter includes the study of the link budget analysis. This helps in understanding the power received at any point within the entire system, thereby getting us a step closer for identifying the range of operation. Next, we present an analysis of the ambiguity function and CRLB for a two path reflection model with direct path and one scattered path for forward scattering geometry.

The GSM-CommSense system is inherently under-determined, which means only a part of a communication frame is used to extract the channel information. The

details presented here will provide us with an analytical tool that will aid in understanding the system performance before it is used for a particular application. The calculations are provided and the relevant simulation results are shown.

The ambiguity function is simulated for multiple input signal type, each with a different modulation scheme, to show that different signals get affected in a similar way when a particular geometry is used. The simulation results provide insight into the ambiguity function for GSM-CommSense system where three different input signal types are used in the presence of a time delayed and attenuated scattered path information. The different types of input signals that used, are unmodulated sinusoidal signal, BPSK modulated Barker code and GMSK modulated GSM training sequence. The variations in the ambiguity functions are observed and discussed. The information presented here shows, that the received signal changes, if there is one scattered path information along with the direct path information in the received data. The change in the received signal is due to the time delay as well as the signal strength but the simulation results prove that the effect of the time delay of the scattered path is dominant.

The CRLB simulation result is also shown, which provides an analysis of the variance of time parameter for the scattered path with respect to the distance from the receiver. The parameters used for the CRLB simulation are mentioned and the variance of time for one scattered path in the presence of the direct path is shown. From the CRLB we came to a conclusion that the variance of the time of the scattered path is dependent on the distance of the point of reflection from the receiver.

There are few limitations of the current work. The work presented here, is limited to a single scattered path analysis in the presence of a direct path signal at the receiver. The calculations can be extended to multiple scattered paths using the work here as reference as, the matched filter is already calculated for n-paths. Since at this point the system is completely analytical and designed to portray the idea of the system. Thus the values, used for the simulations, are idealistic. GSM-CommSense system can not be compared to traditional radar systems as the amount of information is non coherent and limited, although it can be used

as an ASIN [33] system and steered for specific applications. The traditional definition of resolution of a radar system also does not apply to this system. Thus, an empirical resolution has to be defined for it, if necessary.

Chapter 4

Implementation of GSM-CommSense[2]

4.1 Introduction

To verify the validity of the hypothesis of using the GSM based channel equalization modules to monitor the environment, a working prototype of the proposed system is necessary. At this point, looking at the analytical results the implementation of the system appears to be feasible. With the help of the results presented in Chapter 3, we will implement a commensal radar system that we have named GSM-CommSense.

In this chapter, we introduce the design parameters of the GSM-CommSense system and the implementation of the system in real-time. The system will be implemented to receive the GSM base station broadcast signals with an attempt

[2] Based on Abhishek Bhatta and Amit Kumar Mishra, "GSM-based CommSense system to measure and estimate environmental changes," in *IEEE Aerospace and Electronic Systems Magazine* 32, no. 2 (2017): 54-67; Abhishek Bhatta and Amit Kumar Mishra, "GSM based hand-held CommSense-Sensor for environment monitoring," in *IEEE 11th International Conference on Industrial and Information Systems (ICIIS)*, 2016, pp. 360-364, IEEE, 2016 and Abhishek Bhatta and Amit Kumar Mishra "Implementation of GSM channel estimation using open-source SDR environment," in *International Conference on Microwave, Optical and Communication Engineering (ICMOCE)*, 2015 , pp. 322-325, IEEE, 2015.

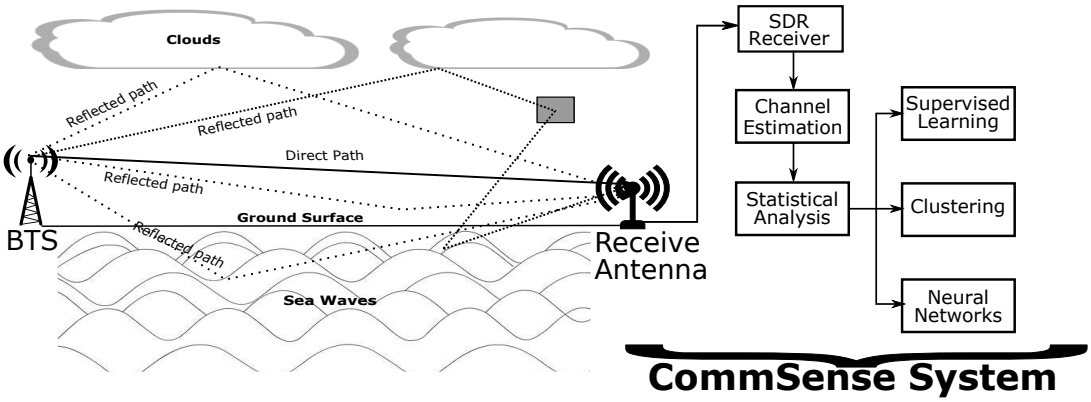


Figure 4.1: Concept diagram showing the system architecture for GSM-CommSense system.

to sense the immediate environment near the receiver.

4.2 CommSense System Architecture

An idea that originated by studying a wireless communication system now needs to be realized in the real world. The proposed GSM-CommSense system is mainly based on estimating the multipath channel parameters from the received signal with the help of the digital frame structure transmitted by the parent system. In this section, we introduce the system blocks necessary to design the GSM-CommSense system.

Figure 4.1 contains all the system level blocks that are planned for the implementation of GSM-CommSense system. A wireless signal transmitted by a base station gets reflected off all the objects in its path carrying information about the path it travelled until it reaches the receiver. At the receiver this information is digitally sampled and converted to its equivalent I and Q components. This is done in the SDR receiver block. The sampled IQ data is then used as the digital representation of the analogue data received by the antenna and processed to extract the channel information.

The method by which the channel information is extracted is named channel

estimation in communication system. The estimated channel information is then used by the communication system to equalize the effects of channel that helps in increasing the SNR, decreasing Intersymbol Interference (ISI), mitigating the multipath effect, etc.. The GSM-CommSense system is designed to utilize these estimated channel values and perform statistical analysis on it without affecting the parent system.

Since the current plan is to implement the GSM-CommSense system to extract the channel information from GSM frames, a brief introduction to the GSM system is necessary.

4.3 Characteristics of GSM

GSM is a communication protocol that acts as a host to the current system. It uses FDMA and TDMA in order to accommodate multiple users. The major frequency bands used for operation of this technology are 800 MHz, 900 MHz, 1800 MHz and 1900 MHz. GSM works in frequency duplex mode thus a different band of frequency is used for uplink and downlink.

The 900 MHz GSM band is used here to implement the GSM-CommSense system and collect data. It is also referred to as GSM900 . It uses $880 - 915$ MHz for uplink and $925 - 960$ MHz for downlink operations. These bands are further divided into 200 kHz bands and separated by a number called ARFCN.

The wireless channel is divided into TDMA frames of 8 time slots of 577 μs each. Each time slot can be used either by the mobile or the base station to transmit a burst. There are many logical channels piggybacked on the physical channels, described in [59]. A normal burst of GSM carries information along with 26 bits of training sequence as shown in Figure 4.3. There are 114 bits of information transmitted in each burst separated with tail bits and a flag. The guard period is added to provide a window of error against distortions that occur due to the rise and fall time of the signal. More details about GSM system can be found in [60].

4.4 GSM Channel Equalization

Each frame in the GSM system transmits a known bit sequence called training sequence. This training sequence is used by the receiver to detect and reduce the possibility of transmission error. In this section, we show the calculations are necessary to understand the channel estimation technique and explain in detail about the algorithm that will be implemented to extract the channel information from the received signal.

Figure 4.2 contains a block representation of the proposed implementation. The dark shaded region belongs to the communication system and will not be affected by the GSM-CommSense system operation. This section focuses only on the region under the dotted box.

4.4.1 Channel equalization filter

4.4.1.1 Calculation

The received GSM signal can be mathematically represented as:

$$y(t) = h(t) * x(t) + n(t) \quad . \tag{4.1}$$

Here, the received signal, $y(t)$, is represented as a convolution of the transmitted signal, $x(t)$, and the Channel Impulse Response (CIR), $h(t)$, in the presence of AWGN, $n(t)$. The transmitter sends a known training sequence in each frame as shown in Figure 4.3, which is divided into reference length of P and guard period of L bits [51]. This equation can be represented in the matrix form and is shown in:

$$\mathbf{y} = \mathbf{Mh} + \mathbf{n} \quad . \tag{4.2}$$

Here, $\mathbf{M}$ represents a Toeplitz like matrix structure of the given training sequence

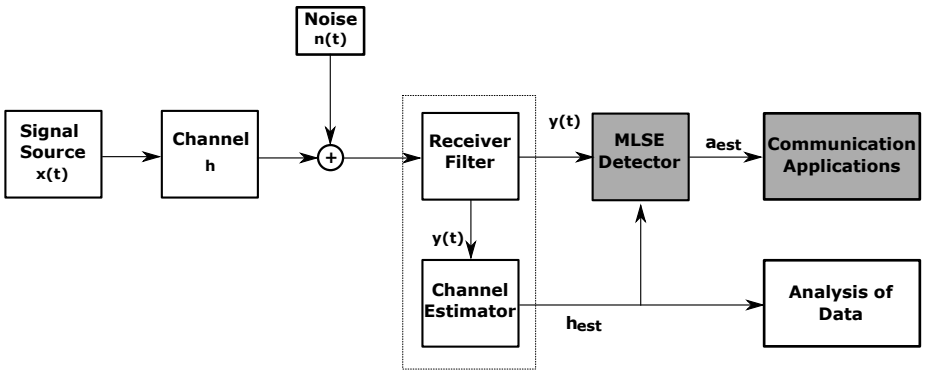


Figure 4.2: Channel estimation block diagram.

such as:

$$\mathbf{M}_{P+Cl-1\times Cl} = \begin{bmatrix} m_0 & 0 & 0 & \cdots & 0 \\ m_1 & m_0 & 0 & \cdots & 0 \\ m_2 & m_1 & m_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_P & m_{P-1} & m_{P-2} & \cdots & m_0 \\ 0 & m_P & m_{P-1} & \cdots & m_1 \\ 0 & 0 & m_P & \cdots & m_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & m_P \end{bmatrix} . \tag{4.3}$$

Here, $\mathbf{M}$ is a $P + Cl - 1 \times Cl$ matrix where, P is the reference length of training sequence and Cl is the length of the CIR. This matrix is made of the $\mathbf{m}$ array that is shown in Equation 4.4, which is actually the oversampled information from each received frame after filtering out the Gaussian noise.

$$\mathbf{m} = \begin{bmatrix} m_0 & m_1 & \cdots & m_{P+L-1} \end{bmatrix}^T \tag{4.4}$$

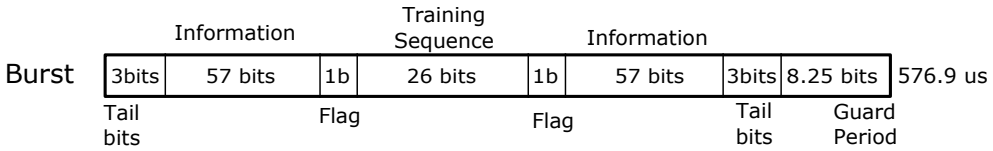


Figure 4.3: GSM normal channel frame structure.

$$h = [h_0 \quad h_1 \quad h_2 \quad \cdots \quad h_n]^T \tag{4.5}$$

Equation 4.5 shows the CIR array. Here, each value of the array represents the channel parameters of a single scattered path. Thus, the length of the array determines the number of multipath signal under analysis. The least squares algorithm is used to find the CIR by minimizing the squared error quantity which in the presence of white Gaussian noise is given as:

$$\mathbf{h}_{est} = (\mathbf{M}^H\mathbf{M})^{-1}\mathbf{M}^H\mathbf{y} \tag{4.6}$$

Here, $\mathbf{M}^H$ denotes Hermitian transpose matrix and $()^{-1}$ denotes matrix inverse. It can be observed here that the received signal matrix $\mathbf{y}$ is multiplied to a matrix also known as the pseudo inverse matrix of $\mathbf{M}$.

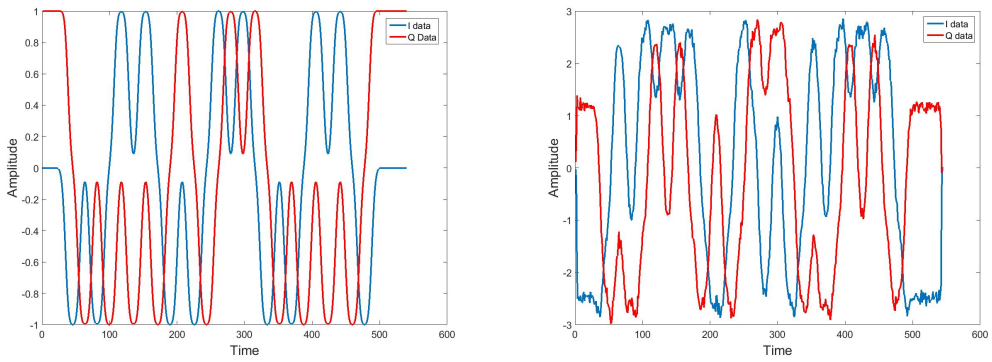
4.4.1.2 Simulation

The calculated channel estimation algorithm is simulated to check the accuracy of the calculations. One of the known training sequence in GSM standard is used to perform the current simulation. The training sequence used here is [0 0 1 0 0 1 0 1 1 1 0 0 0 0 1 0 0 0 1 0 0 1 0 1 1 1] [59]. The CIR used as interference is $[0.0007, 0.2605, 0.9268, 0.2605, 0.0007]$. The received signal will be a convolution of transmitted training sequence and the CIR along with AWGN, as shown in Equation 4.1. Since the central 16 bits of the training sequence is

used to estimate the channel effects the matrix, **M**, is represented as:

$$M_{20\times5} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.7}$$

Figure 4.4 contains the simulated I and Q data which are used to check the channel equalization filter as shown above. Since GSM uses GMSK modulation scheme, the modulated transmit signal is simulated as shown in Figure 4.4a. The received signal with the channel impairments is shown in Figure 4.4b.



(a) (b)

Figure 4.4: Simulated IQ data for GMSK modulated GSM training sequence. (a) contains the transmitted signal without impairments. (b) contains the received signal with channel impairments and AWGN.

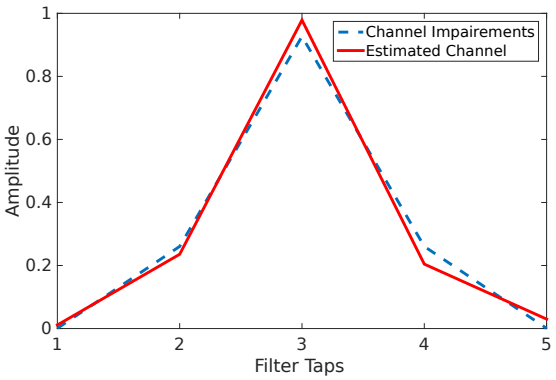


Figure 4.5: Simulated least square channel estimation in presence of AWGN.

The simulation result showing the least square estimated channel information is shown in Figure 4.5. This result is for a single GSM training sequence being transmitted over a known channel of length 5 taps. The estimated value is very similar to the original CIR value which means the algorithm working as intended and can be implemented in real time to obtain actual channel values.

4.4.2 Implemented algorithm description

The above mentioned calculations provided expected results in simulations. The next step here is to implement it on real data to extract the channel information from actual GSM frames. The details of the algorithm implemented in real time is presented here.

When a signal is received it consists of unknown noise components. A low pass filter cleans up the noise and the ADC samples a relatively less noisy signal. The signal is oversampled with an Oversampling Ratio (OSR) of 4. The oversampled signal is then available at the receiver to perform channel estimation.

Algorithm 1 Pseudo-code for real-time implementation

```
Search 2× Guard period
if Receive burst ← (Transmitted burst + 2× Guard period)× OSR then
  Train position←Tail bits+(Data bits + flag)+ first 5 bits of training sequence
  if Normal Burst then
    center ← Train position + guard period×OSR
    start ← center + 1
    stop ← CIR length×OSR 2×OSR
    for i=start; i <stop; i++ do
      correlate ←known training sequence, receive burst[i]
      correlation buffer [correlate]
      Apply Least Squares to estimate the CIR
    end for
    Plot CIR in real-time and repeat for next burst
  else
    wait for normal burst and repeat
  end if
else if Error finding the starting point of receive burst then
  Discard the current burst
  Repeat from top
end if
```

Algorithm 1 contains the pseudo-code which provides an overview of the implemented algorithm. Initially, the system waits to receive and detect one full burst. This is done by looking for two consecutive guard periods. The informa-

tion within those guard periods is considered as the received frame. Here the frame structure is checked to identify a GSM normal frame as shown in Figure 4.3. Once the normal frame structure is identified, the training sequence is extracted from the correct position. This training sequence is passed through the channel equalization filter to extract the CIR. The CIR is then saved to a file and the process is repeated until an external interrupt stops the algorithm.

The current algorithm is implemented to extract the CIR from the normal frame structure as it is the most transmitted burst which contains the information and the training sequence. A burst containing no information is called dummy burst. This burst has the same structure as the normal burst except the information bits are random. Since there is no need for the information bits here both the normal burst and the dummy burst are considered under the normal frame structure.

4.5 Real-Time Implementation of the Channel Estimation Algorithm

The system is implemented in real-time on open source hardware and software platform. GNU Radio is a software tool-kit that provides basic signal processing blocks for SDR applications. The GSM-CommSense system is implemented in GNU Radio, in which the channel estimation is performed. The estimated CIR is plotted in real time as shown in Figure 4.6.

The SDR hardware used for this implementation is BladeRF×40, which works in a frequency ranging from 300 MHz to 3.8 GHz with a maximum physical bandwidth of 28 MHz. An off the shelf, quad-band, omni-directional, GSM antenna is used to receive the signal.

In order to test the operation of the algorithm, the extracted CIR is used to decode the received GSM frames. The decoded bit-stream is analysed in Wireshark (a packet analyser tool) as shown in Figure 4.6. Wireshark extracts the basic information from the frames such as the location of the BTS, the name of the carrier service provider, etc..

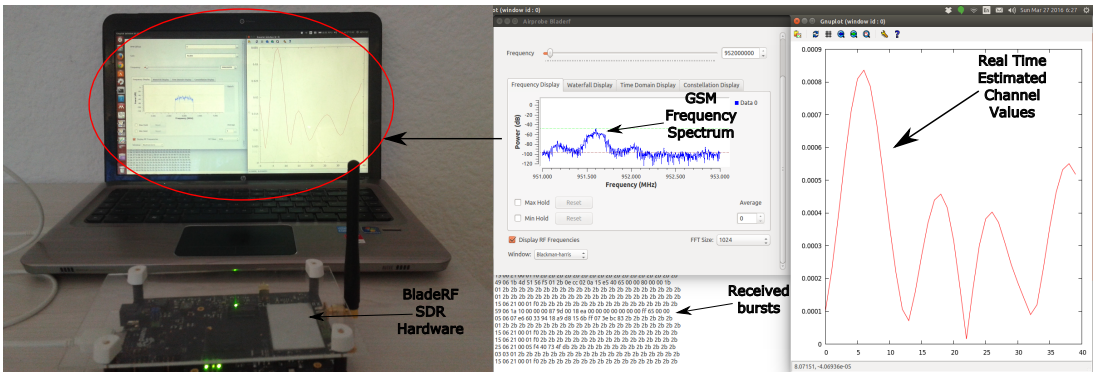


Figure 4.6: GNU radio implementation of channel equaliaztion with BladeRF×40.

4.5.1 Extracting the channel values

The steps taken to implement the channel equalizer in real-time are mentioned here. Correlation between the received signal and the known training sequence is used to extract the differences between the two signals. This information is saved in a file to perform further analysis.

A SDR hardware called BladeRF, is used to receive the GSM signal. This down-converts the signal according to its Phase Locked Loop (PLL) clock frequency and outputs the I and the Q components of the signal.

Initially the receiver has to be synchronized with the base station. This is performed with the help of a Synchronization Channel (SCH) burst. The SCH burst consists of a 64 bit long training sequence and transmits the same sequence in every timeslot allocated to it. This is a distinguishable feature that helps the receiver to get synchronized. Once the receiver is synchronized with the GSM base station, the frequency offset of the allocated physical channel is calculated with the help of a Frequency Correction Channel (FCCH) burst.

The received burst type is identified in two steps, first with the help of the carrier index of the burst and then with the burst number. Carrier Index Zero (C0) is a special case where only Broadcast Channel (BCCH) data is transmitted. BCCH carries a repeating pattern of system information such as identity, configuration

and available features for the BTS. In case of other carrier indices, the burst type is determined by the burst number. The most commonly transmitted burst is the normal burst thus in this implementation we are focusing on extracting the channel information from the normal burst.

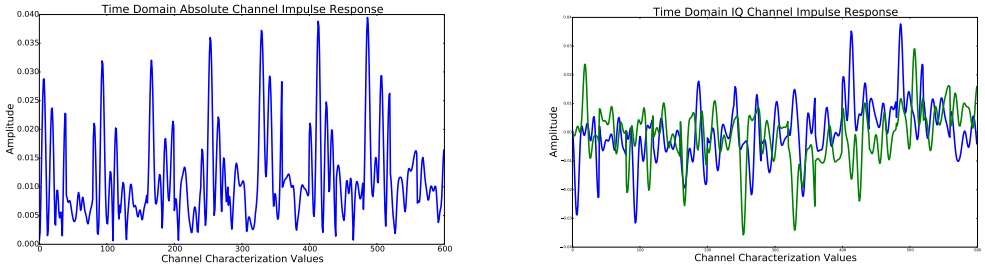
In order to receive a burst completely the receiver needs to wait for two consecutive guard periods. Although there will be an overlap between the guard periods of two consecutive bursts, this is necessary to ensure that full burst is received. Once a normal burst is identified the start and the stop position of the burst is determined by eliminating the guard period as it consists of 8.25 bits on each side and is distinguishable. The central 26 bits of the remaining bits are extracted as they consist of the received training sequence. The training sequence extracted here is a complex data so the already available sequence also needs to be converted into a complex number. The two complex training sequences are correlated with each other and the correlation information is saved in a buffer. A pointer is placed at the beginning of the buffer and based on predefined CIR length, it is moved from the beginning to the length of the CIR. The values corresponding from the beginning to the CIR length is saved in an array and then moves to a file where it can be saved for further analysis.

In case of communication systems the maximum absolute value of the CIR is chosen discarding the rest of the values and that information is used to equalize the channel and get better communication. For the purpose of this implementation the entire channel information is necessary. Thus, it is saved.

In order to check that the implemented receiver is working as intended the entire received structure is made referring to the "gr-gsm" libraries available for GNU Radio and the bursts are printed out using a message printer. Figure 4.6 contains the plot of the entire implementation including the GSM frequency spectrum, received bursts and the real time estimated channel values.

4.5.2 Real-time data capturing

Implementation of the algorithm mentioned above allowed us to capture data in real-time and save it for further processing. Figure 4.7 contains the plot showing the saved channel information. The individual I and Q component is shown in Figure 4.7b and the absolute value calculated by $h_{est} = \sqrt{I^2 + Q^2}$ is shown in Figure 4.7a.



(a) Calculated as $h_{est} = \sqrt{I^2 + Q^2}$. (b) The I and the Q components.

Figure 4.7: Estimated normal channel values in terms of the I and Q components.

One of the major challenges faced during this implementation is finding the starting point of the burst. This is done with the help of the guard period. The receiver waits for two consecutive guard periods which, when received, is considered as the starting point of the burst. In the case of an error in receiving the guard periods, the system drops the current set of data and waits for the next set and then repeats until it gets two consecutive guard periods, thereby eliminating the chances of any ambiguous data.

4.6 User Requirement Analysis for Hand-Held GSM-CommSense System

At this point a basic implementation of the GSM-CommSense system is completed. Currently the system is drawing power from a wall power socket. This brings to the next part where we want the system to be mobile and self-sufficient

without an external power source. Mobility of the system will help in taking captures at different locations and analysing the results.

In this section, we present the major requirements for the implementation of hand-held GSM-CommSense system.

4.6.1 Top level design specifications

4.6.1.1 Capture GSM signals

GSM works on FDMA and TDMA in order to allow effective communication to multiple users [43, 44]. Thus, it has different frequency bands for transmit and receive. The CommSense system is designed to receive GSM900, which has an uplink frequency of $890 - 915$ MHz and downlink frequency of $935 - 960$ MHz. These 25 MHz uplink and downlink bands are further divided into 200 kHz carrier bands which are classified by a number called ARFCN. More details about GSM can be found in the standards [59, 60, 58, 52]. The implemented receiver is required to receive the 200 kHz instantaneous band based on a particular ARFCN. The receiver hardware should have a minimum instantaneous bandwidth of 200 kHz across the band of interest at an OSR of 4 so that the data is received in minimizing loss.

4.6.1.2 Processing unit

The processing unit should be chosen based on the following requirements:

4.6.1.2.1 Pre-processing Unit: Once the signal is received by the RF front end and passes through the ADC, it needs to be processed and demodulated to confirm that the received signal is GSM. To implement this part a processing system is required. Since one of the objectives for this implementation is to have a hand-held system. The processing system should be portable.

4.6.1.2.2 Memory (to save data): The goal of implementing this system is to receive and save the estimated channel data. Thus, the designed unit should have enough memory space to be able to capture and save multiple datasets. It has been observed that the system can capture approximately 960 MB of data in an hour. Although eventually the plan is to stream the data through a post-processing unit, it is imperative to have on system memory in certain cases. All these captured data will be analysed in the future to observe the differences in the channels. The operating system, on the portable processing devices, takes about 4 GB of space. So, in order to have space for the data minimum required memory space is 8 GB preferably more.

4.6.1.2.3 GNU Radio: The processing unit should be capable of running a Linux operating system as the GSM-CommSense system will be implemented using GNU Radio.

4.6.1.3 Upward compatibility

Although the system is designed to be self-sufficient there should be scope for further development in the future. This means that the major components of the system should be chosen such that, if necessary, more hardware can be added to the system without changing any part of the basic implementation.

4.6.1.4 Stand-alone system

The system needs to be stand-alone with battery power to ensure usage in remote locations, away from any nearby power sources. This will give the system a mobility and a degree of freedom which will aid in getting better data captures.

4.7 Stand-Alone System Design

The system is designed to meet all the user requirements while also providing a buffer space for future developments.

The first step in the design of the hand-held GSM-CommSense system is to get the hardware suitable for the above mentioned requirements. To capture the GSM signal a RF front end receiver is required, which has become fairly easy to pick from off-the-shelf components with the recent advent of SDR technology. Different SDR hardware, that are investigated to find the best suited receiver for the current implementation, is shown in Table 4.1.

Devices Analysed				
SDR	**Frequency Range (MHz)**	**Receive Bandwidth (MHz)**	**ADC Resolution (bits)**	**Transmit Capability**
RTL-SDR (R820T)	$24 - 1766$	3.2	8	No
FUNcube Dongle Pro +	$0.15 - 260$ and $410 - 2050$	0.192	16	No
Airspy	$24 - 1800$	10	12	No
BladeRF	$300 - 3800$	40	12	Yes

Table 4.1: Different SDR hardware specifications.

The RTL-SDR covered most of the GSM operating frequencies except 1800 MHz. The FUNcube Dongle Pro + covered all the necessary frequencies but lacked in the receive bandwidth as the GSM signal has 200 kHz bandwidth. The Airspy also lacked the ability to scan at a frequency of 1800 MHz. The BladeRF covers all the GSM operating frequencies, with a high receive bandwidth and 12 bit ADC resolution. Additionally, it has transmission capabilities as well. Although the current implementation is on GSM900 and no transmitting is required, the plan here is to design an upward compatible system, which means having the ability to transmit, in case such a capability is required in the future. Another device that is commonly used is a USRP board. The cost of such a board is significantly higher than all the other boards investigated here. Another problem with the

USRP board is the requirement of daughter boards which further adds to the cost. Thus using BladeRF for the implementation of GSM-CommSense RF front end is a feasible choice.

During the initial system design a laptop was being used as a processing unit. This did not serve the purpose of portability of the system as it was bulky to carry. To make GSM-CommSense portable a different processing unit was necessary which can run GNU Radio on it. Out of all the portable processing devices available in the market Raspberry pi has an extensive user community and support. Additionally, due to the ease of implementation on a Raspberry pi, it was chosen as the processing unit. The basic operating system provided by Raspberry pi is called Raspbian, which is a Debian based operating system and can run GNU Radio comfortably.

Initially Raspberry pi 2 model B was chosen for the implementation. It contains a 900 MHz quad-core ARMv7 Central Processing Unit (CPU) with 1 GB Random Access Memory (RAM) and 4 Universal Serial Bus (USB) 2.0 ports. There is an Ethernet port on board that can be used to communicate with the device. Later, it was upgraded to Raspberry pi 3 model B that increased the CPU frequency to 1.2 GHz and provided on board bluetooth and Wireless Fidelity (Wi-Fi). This increased the processing speed and was helpful in transferring the data to another device wirelessly.

The BladeRF can be powered up from one of the USB ports as it draws approximately 500 mA of current at 5 V Direct Current (DC) input. Alternatively, with a basic change in hardware setting, the BladeRF can be powered using an external power bank, which can be used in the cases where the device requires higher power.

A compatible class 10 microSD card of size 32 GB is used to load the operating system and keeps all the data captured. The base operating system takes approximately 4 GB of space on the memory card and the GNU Radio takes about 1 GB thereby keeping approximately 27 GB extra space for other storage. It is very important to find a compatible microSD card as all memory cards do not perform efficiently on the Raspberry pi. The compatibility can be checked online

at [104].

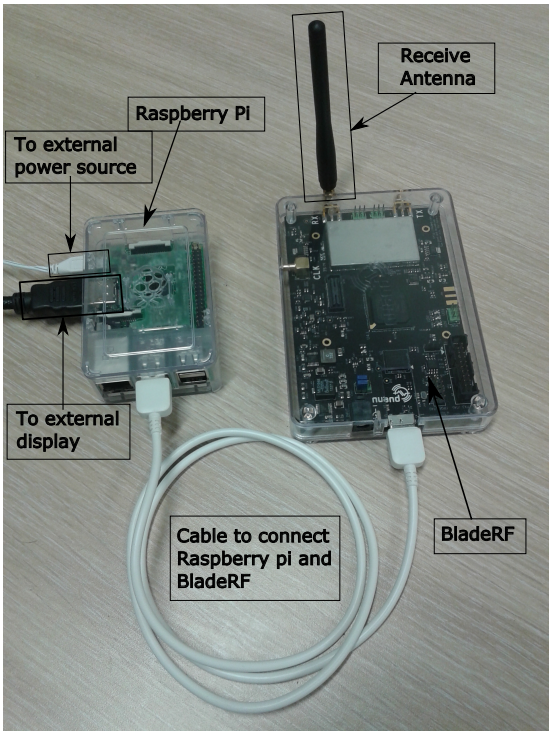


Figure 4.8: Implementation picture using Raspberry pi and BladeRF.

Upward compatibility is a key element for the current implementation of the system as it leaves room for further development. The BladeRF specifications are higher than the current necessity, along with transmit capabilities which can be used in the future. The Raspberry pi has extra USB ports which can be used to connect more peripherals. The information from the Raspberry pi can be transferred to other devices, such as a graphics processor unit, using either the Ethernet port or Wi-Fi.

The proposed stand-alone system should be able to sustain 2 A of current at 5 V DC for a minimum of 3 to 4 hours. The plan is to use an external power bank to provide the necessary power, but all the devices tested provided a maximum current of 1.1 A of current at 5 V DC. Thus, we decided to use two separate off-the-shelf power banks, one powering the Raspberry pi and the second one to power up the BladeRF.

Figure 4.8 contains the implementation picture showing the Raspberry pi and the BladeRF. Here, the BladeRF is connected to the Raspberry pi through a USB cable. In the picture the power cable for BladeRF is not shown, but while operation a power cable from one of the power banks is connected to it as power source. The antenna shown in this picture is off-the-shelf GSM quad band omnidirectional antenna which works on all GSM frequency bands (800 MHz, 900 MHz, 1800 MHz and 1900 MHz). The Raspberry pi has a built in High-Definition Multimedia Interface (HDMI) connector which can be connected to a monitor to use it as a screen for the device, if needed.

4.8 Summary

This chapter contains the details about the implementation of the GSM-CommSense system. The real-time system implementation will be useful in proving the original hypothesis.

We started with the top-level system architecture showing the multipath interference caused due to the wireless communication channel and the block representation of the planned implementation. Since the current GSM-CommSense system is implemented using the GSM system as the parent system, a brief introduction to the GSM protocol is presented.

The GSM channel equalization filter is explained in detail showing calculations and simulations. After getting proof that the implemented algorithm is operational on simulated data, it was implemented in real-time to receive the normal frames from the GSM broadcast channel. The channel values are then extracted and the plots of the received IQ data are presented.

The user requirements are analysed and the system specifications are defined in order to design a hand-held/portable stand-alone system. Multiple devices are analysed and finally Raspberry pi 3 model B is chosen as the processing device and a SDR receiver named BladeRF is used to receive the signal in real time. The details of the hardware specifications used for the implementation is shown

Table 4.2: Hand-held system design summary.

Specifications		**Comments**
Capture GSM signal		BladeRF with quad band omnidirectional GSM antenna.
Processing unit	Pre-processing unit	Raspberry pi to run GNU Radio and process the received GSM signal.
	Memory	
	GNU Radio	
Upward compatibility		Opportunity of improvement without changing the hardware. BladeRF and Raspberry pi has higher specifications than necessary at this moment.
Stand-alone system		Power bank that can provide around 2 A of current at 5 V DC if unavailable use separate power banks to power each device.

in Table 4.2.

Open source software development tool-kit named GNU Radio is used to implement the GSM-CommSense system in real-time. The major signal processing blocks are implemented in c++ and each of the blocks is linked with each other using python.

Chapter 5

Statistical Analysis of GSM-CommSense Data[3]

5.1 Introduction

At this point a basic implementation of GSM-CommSense system is complete. The next step in proving the original hypothesis is to statistically characterize the captured data based on each event and observe the differences to be able to predict certain behaviours.

In this chapter, the estimated channel information is analysed to show the changes in the channel state in different environmental condition. Here, we will perform three different analyses on the captured data, two in statistical domain and one in data/time domain.

In the first approach, we estimate the Probability Density Function (PDF) of the channel values captured at different scenarios. This shows that the estimated PDF's vary as the environmental conditions change which means that a "hypothesis test" algorithm can be used to distinguish different environment

[3] Based on Abhishek Bhatta and Amit Kumar Mishra, "GSM-based CommSense system to measure and estimate environmental changes," in *IEEE Aerospace and Electronic Systems Magazine* 32, no. 2 (2017): 54-67.

types from the estimated channel values.

Table 5.1: The different sets of captured data analysed in this chapter.

Captured Data			
Sl. No.	**Category**	**Name**	**Description**
1	**Corner Reflector (CR)**	No CR	The following captures are taken on the roof of a building. In this case the receiver is placed at a location and captures are taken without the presence of a CR.
		Vertical CR	The receiver is at the same location as above with a metal plate in vertical alignment with the receive antenna at a distance of 3 m from the antenna is placed. This is done for each of the metal plates used to construct the dihedral CR.
		Dihedral CR	A dihedral CR, made of placing two metal plates at right angle to each other, is placed at a horizontal distance of 3 m from the receive antenna.
2	**Environment**	Rain	Captures are taken while placing the receiver at the balcony of a house. This part is captured when there is heavy rain pouring down.
		Humid No Rain	Captures are taken at the same location as above when the rain stop but the climate is still humid.
		Hot Day	Keeping the location same as above the captures are taken on a different day when it is a bright hot sunny day.

Table 5.1: The different sets of captured data analysed in this chapter.

Captured Data			
Sl. No.	**Category**	**Name**	**Description**
3	**Parking**	Parking Full	Captures are taken at a parking space when the place was packed with cars.
		Parking Empty	The captures for this event is taken at the same parking space with the same type of environment if the parking space is empty.
4	**Train Station**	With Train	The data here is captured at a train station when there is a train, at a horizontal distance of 4 m from the receive antenna.
		No Train	The receiver location is kept the same as mentioned above but in this case the train is not present at the station.
5	**Jammie Shuttle**	With Jammie location 1 and 2	The internal bus taking students from one campus to the other at UCT is known as Jammie shuttle. The captures for this set is taken at two different locations when a Jammie shuttle is present at 3 m from the receiver.
		No Jammie location 1 and 2	Keeping the respective locations constant these captures are taken when the Jammie shuttle is not present in the vicinity of the receiver.

Table 5.1: The different sets of captured data analysed in this chapter.

Captured Data			
Sl. No.	**Category**	**Name**	**Description**
6	**Car**	Car day 1 and 2	The captures are taken in two days keeping the location of the car and the receiver same.
		No Car day 1 and 2	These captures are taken in two days right after the car captures are taken. To take this capture the cars are to move away from the vicinity of the receiver while maintaining other environmental conditions as much as possible.
7	**Miscellaneous**	Jameson Stairs Location 1 and 2	The captures are taken at the entrance of a building of approximate height 10 m.
		Sea Beach	The captures for this set is taken at the camps bay beach.
		Beach Rock	The captures for this set is taken at the same beach but near boulders of approximate height $8 - 10$ m.

In the second approach, we perform a Chi-square test to assess the goodness-of-fit between the observed data and the theoretically expected dataset.

In the third approach we check for the clustering of the estimated channel values for different capture scenarios. For easier visualization we use Principal Components Analysis (PCA) to reduce the dimension of the data. This shows that the data is highly clustered in the PC domain.

A list detailing the data analysed here is given in Table 5.1. The results are plotted and the plots are discussed.

5.2 Probability Density Function Analysis

5.2.1 Description

In this analysis, we observe the empirical PDF of the data collected from different scenarios as described in Table 5.1. In addition, we also compare the empirical PDF of the data with respect to four different theoretical PDF models, named Rayleigh, Gaussian, log-normal and gamma distributions. The individual parameters are extracted from the empirical PDF using the Maximum Likelihood Estimation (MLE) algorithm for each distribution.

A brief description of the distributions and MLE expression of their parameters are as follows.

5.2.1.1 Gaussian distribution

The most general form of distribution used in statistical analysis is the Gaussian distribution. The PDF for the distribution is given as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad . \tag{5.1}$$

Here, μ is the location parameter sometimes also referred as mean and σ is the scale parameter also known as standard deviation of the Gaussian distribution.

Standard Gaussian distribution where $\mu = 0$ and $\sigma = 1$ is given as:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad . \tag{5.2}$$

The MLE of the parameters for the Gaussian distribution is derived by minimizing the log-likelihood function of the equation (5.1) and is given as:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{N} x_i \tag{5.3}$$

$$\hat{\sigma}^2 = \frac{1}{(n-1)}\sum_{i=1}^{N}(x_i - \hat{\mu})^2 \quad . \tag{5.4}$$

5.2.1.2 Rayleigh distribution

One of the most common distributions used in wireless communication systems to model the channel is Rayleigh distribution. The PDF for this distribution is given as:

$$f(x) = \frac{1}{\sigma^2} x e^{\frac{-x^2}{2\sigma^2}} \quad . \tag{5.5}$$

Here, σ is the scale parameter, also known as mode of the Rayleigh distribution.

The MLE of the parameters for Rayleigh distribution is given as:

$$\hat{\sigma}^2 = \frac{1}{2n}\sum_{i=1}^{N} x_i^2 \qquad x > 0 \quad . \tag{5.6}$$

5.2.1.3 Gamma distribution

Gamma distribution is also a versatile distribution which can be manipulated in many ways to generate other distributions such as k-distribution. The general formula for the PDF of gamma distribution is given as:

$$f(x) = \frac{(\frac{x-\mu}{\beta})^{\gamma-1}\exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\Gamma(\gamma)} \qquad x \geq \mu; \gamma, \beta > 0 \quad . \tag{5.7}$$

Here, γ is the shape parameter, μ is the location parameter, β is the scale parameter and Γ is the gamma function given as:

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt \quad . \tag{5.8}$$

The case where $\mu = 0$ and $\beta = 1$ is called the standard gamma distribution and is given as:

$$f(x) = \frac{x^{\gamma-1}e^{-x}}{\Gamma(\gamma)} \qquad x \geq 0; \gamma > 0 \quad . \tag{5.9}$$

The MLE for the two parameter gamma distribution is calculated by solving the following equations simultaneously:

$$\hat{\beta} - \frac{\bar{x}}{\hat{\gamma}} = 0 \tag{5.10}$$

$$\log\hat{\gamma} - \psi(\hat{\gamma}) - \log\left(\frac{\bar{x}}{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}\right) = 0 \tag{5.11}$$

with ψ denoting the digamma function given by equation (5.12) is the mathematical derivative of the gamma function. These functions are solved using the stats function available in python scipy package.

$$\psi(z) \equiv \frac{d}{dz}\ln\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)} \tag{5.12}$$

5.2.1.4 Log-normal distribution

If x, is a random variable distributed log-normally then, $y = \ln(x)$, is Gaussian distributed where ln is natural log. The PDF is given as:

$$f(x) = \frac{e^{-((\ln((x-\theta)/m))^2/(2\sigma^2))}}{(x-\theta)\sigma\sqrt{2\pi}} \qquad x > \theta; m, \sigma > 0 \quad . \tag{5.13}$$

Here, σ is the shape parameter also known as standard deviation of log-normal

Table 5.2: Moments of different distributions shown in Figure 5.1.

Distribution	Shape	Mean	Variance	Skew	Kurtosis
Gaussian	N/A	0.0	1.0	0.0	0.0
Rayleigh	N/A	1.25	0.43	0.63	0.25
Log-normal	0.2	1.02	0.04	0.61	0.68
	0.5	1.13	0.36	1.75	5.89
	1.0	1.64	4.67	6.18	110.93
	2.0	7.39	2926.35	414.36	9220556.98
Gamma	0.5	0.5	0.5	2.83	12.0
	2.0	2.0	2.0	1.41	3.0
	5.0	5.0	5.0	0.89	1.2
	9.0	9.0	9.0	0.67	0.67

distribution. The location parameter is given by θ and m is the scale parameter or the median.

The standard log-normal distribution is given when $\theta = 0$ and $m = 1$. It is denoted as:

$$f(x) = \frac{e^{-((\ln x)^2/2\sigma^2)}}{x\sigma\sqrt{2\pi}} \qquad x > 0; \sigma > 0 \quad . \tag{5.14}$$

The log-normal distribution is commonly characterized with its mean μ given as $\mu = \log m$. Using this parameter the density function can be given as:

$$f(x) = \frac{e^{-(\ln(x-\theta)-\mu)^2/(2\sigma^2)}}{(x-\theta)\sigma\sqrt{2\pi}} \qquad x > 0; \sigma > 0 \quad . \tag{5.15}$$

The MLE for the parameters of this distribution is given as:

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} \ln X_i \tag{5.16}$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} (\ln (X_i) - \hat{\mu})^2 \tag{5.17}$$

$$\hat{m} = \exp \hat{\mu} \quad . \tag{5.18}$$

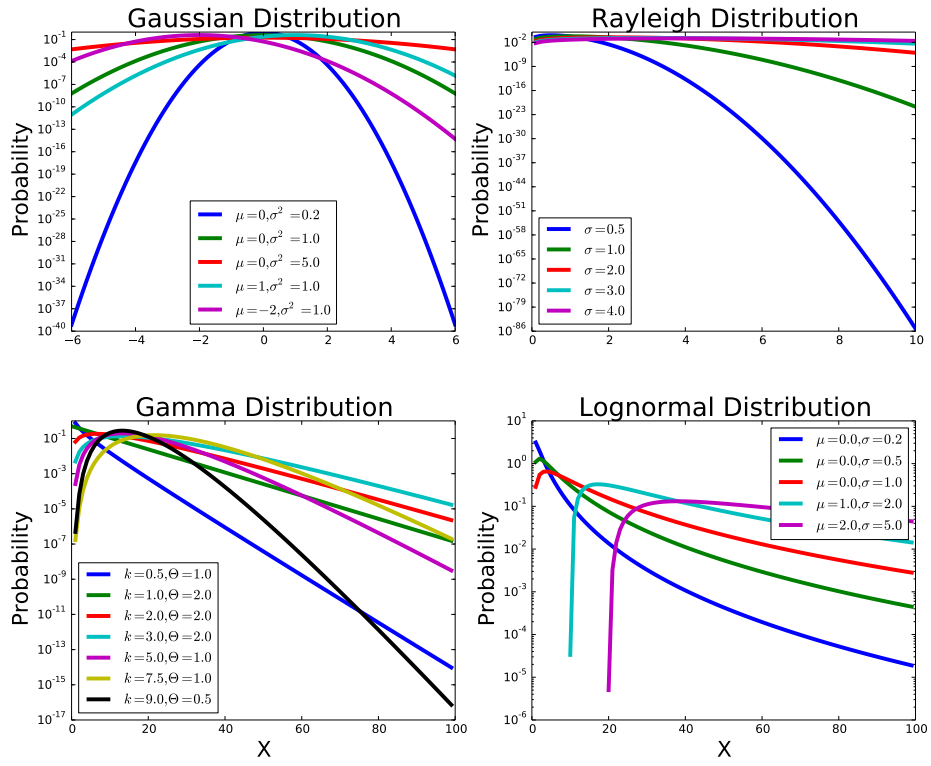


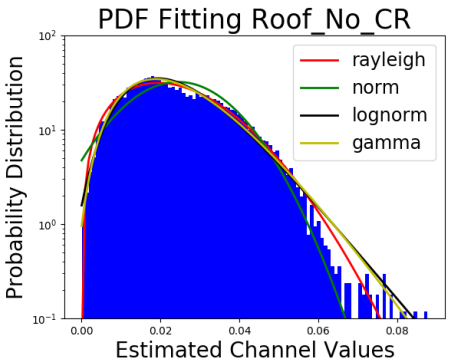
Figure 5.1: All the distributions mentioned above are plotted here (y-axis is in log scale).

Table 5.2 and Figure 5.1 contain the moments and graphs respectively of the above mentioned distribution.

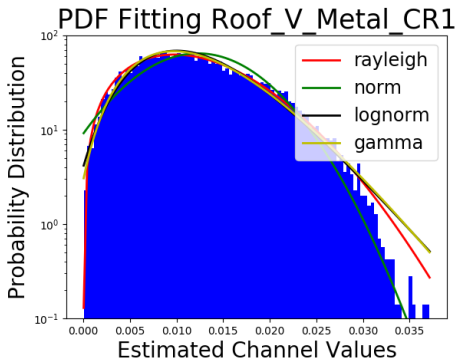
5.2.2 Analysis of captured data

The above mentioned distributions are widely used for analysing the captured multipath data. During the initial phases of comparison we matched to a multitude of different distributions and the best suited distribution was chosen for the analysis presented here.

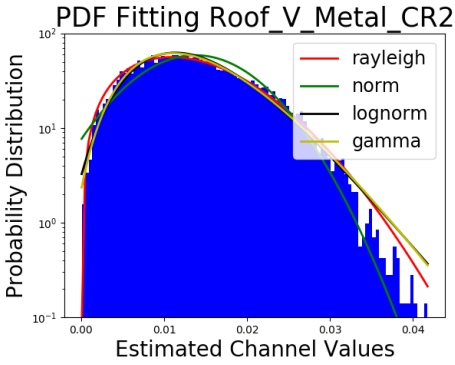
Figure 5.2 contains the empirical and fitted distributions in the presence and absence of different CR's. At first the receiver is placed at a particular location



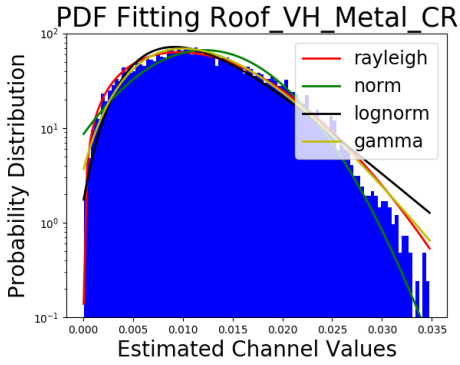
(a) Before placing a CR in the vicinity of the receiver.



(b) Metal plate CR1 in vertical orientation.



(c) Metal plate CR2 in vertical orientation.

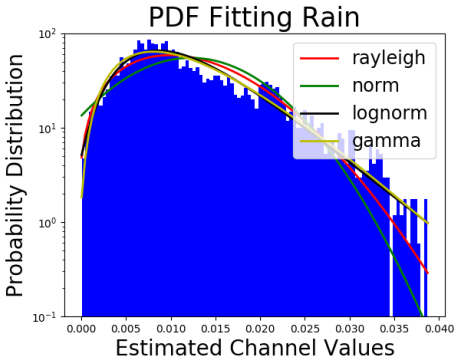


(d) Dihedral CR, with CR1 vertical and CR2 horizontal orientation.

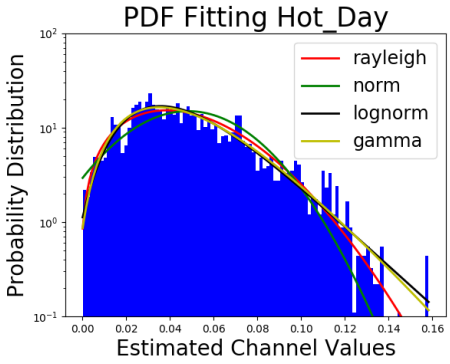
Figure 5.2: Empirical PDF is fitted to expected distributions showing the comparison of data captured by placing CR's at a horizontal distance of 3 m from the receive antenna.

on the roof of a building without any reflectors within the vicinity of the receiver. The data captured without a CR is shown in Figure 5.2a. Two different metal plates are used to construct a dihedral CR. Figures 5.2b and 5.2c contain the fitted distribution plots when the metal plates are placed at a horizontal distance of 3 m from the receiver antenna. One CR is placed at a time, in vertical orientation, in order to capture the data shown in the Figures 5.2b and 5.2c. Figure 5.2d contains the fitted distribution in the presence of a dihedral CR. Here, CR1 is vertically oriented and CR2 is horizontally oriented. The distribution shown in Figure 5.2d is visually similar to the distribution in Figure 5.2b with some differences. This is due to the fact that in both the cases the same

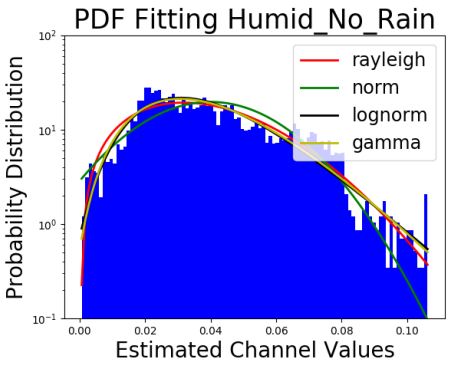
metal plate is kept in vertical orientation and since the antenna is also vertically oriented. Most of the data received are related.



(a) In the presence of heavy rain.



(b) On a hot day with low humidity.

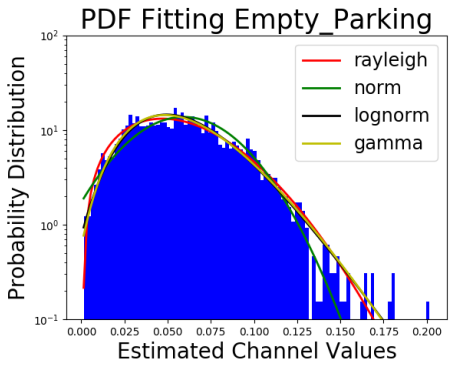


(c) When the rain stopped but was still humid.

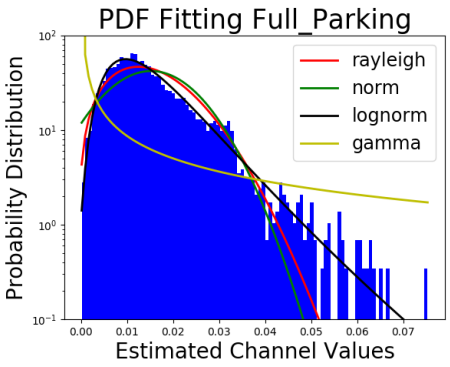
Figure 5.3: Empirical PDF comparison of weather information showing differences is due to different humidity conditions in the environment. The receiver location and orientation are fixed in order to take these captures.

Since the differences in data captured in the presence and absence of a corner reflector are validated, we move to analyse other conditions. The next set of analysis is done on weather information, particularly focusing on detection of rain or humidity with the help of the GSM-CommSense data. Figure 5.3 contains the empirical and fitted distribution of data captured in different humidity conditions. The receiver is placed at the window pane of a building in order to take the captures. Figure 5.3a is captured on a day with heavy rain pouring down. Figure 5.3c is captured on the same day as the rain capture. During this

capture the rain has stopped but the environment is still humid. Figure 5.3b is captured on the day after when the climate is hot and dry. The main idea for this analysis is to observe the difference in the distributions due to the different humidity conditions. The results show that the differences in the distribution is visually separable, this can be later translated into design of the system for weather monitoring.



(a) The parking space is completely empty.

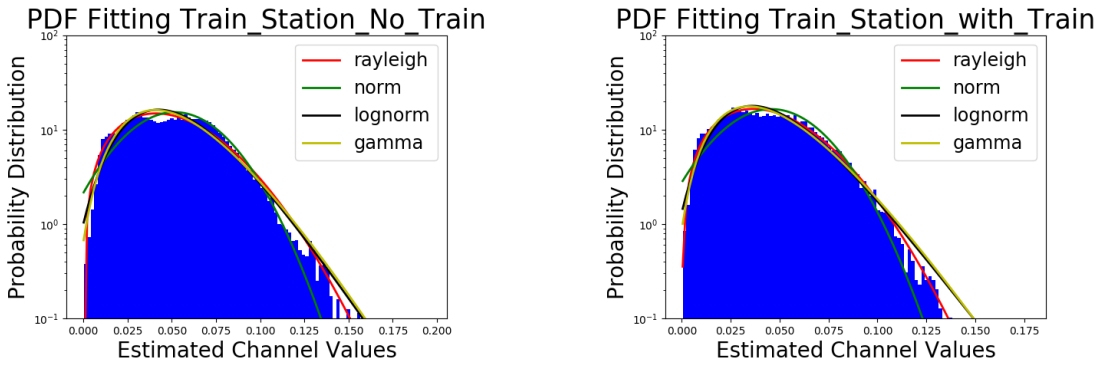


(b) The parking space is full with vehicles.

Figure 5.4: Empirical PDF comparison, when the data is captured at a parking space.

The next few sets of captures are shown to find the difference between the distributions in the presence or absence of different vehicles, such as, trains, cars, buses, etc.. Figure 5.4 contains the empirical and fitted distribution for data captured at an open air parking space. The captures are taken on two different days with similar weather conditions and keeping the receiver location fixed. Figure 5.4a contains the distribution for a condition when the parking space is completely empty. Figure 5.4b contains the distribution representing the parking space when it is full with cars. This test is performed to visualize changes in the distribution due to the presence of multiple vehicles in the vicinity of the receiver. The difference in the distributions is visible from the figures. This can be potentially applied for implementation of the system to monitor a parking space for empty parking slots.

The next analysis from here is to detect the presence or absence of a train in



(a) No train at the train station.

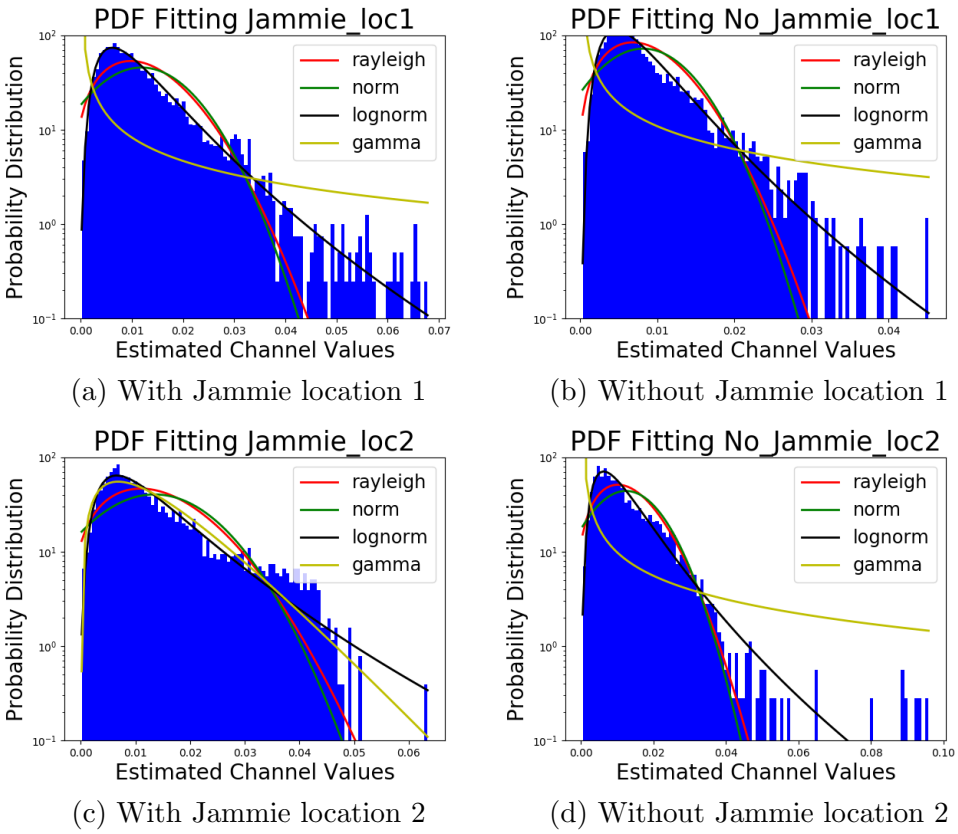
(b) A train is present.

Figure 5.5: Empirical PDF to compare the conditions when there is a train present 6 m from the receiver with no train in the station. The receiver location is fixed.

a train station. A train is a large vehicle thus moving from a parking space with many cars to the train is a logical choice. This set of capture is taken by keeping the receiver at a distance of 6 m from the train track on the platform. Figure 5.5 contains the empirical and fitted distribution related to the train captures. Figure 5.5a contains the distribution for the condition when the train is not present at the station. Figure 5.5b contains the distribution when a train is present at the station. The differences in the distribution, although available, are not clearly observable. These types of ambiguities require further analysis.

As the difference in the distribution with respect to train is not very clearly observable, we decide to move to other vehicles. Figure 5.6 contains the plots showing the difference in the distributions of data captured due to the presence and absence of the UCT shuttle service, named Jammie. The captures are taken in two different locations, Figures 5.6a and 5.6b are taken at location 1, with and without the presence of a Jammie. Figures 5.6c and 5.6d are taken at location 2, with and without the presence of a Jammie. The differences in the distributions for each location are clearly observable. This shows that even though the train dataset does not yield promising results the GSM-CommSense system can be used to differentiate from the presence or absence of vehicles.

Moving forward to another vehicle detection we decide to check the difference in



(a) With Jammie location 1

(b) Without Jammie location 1

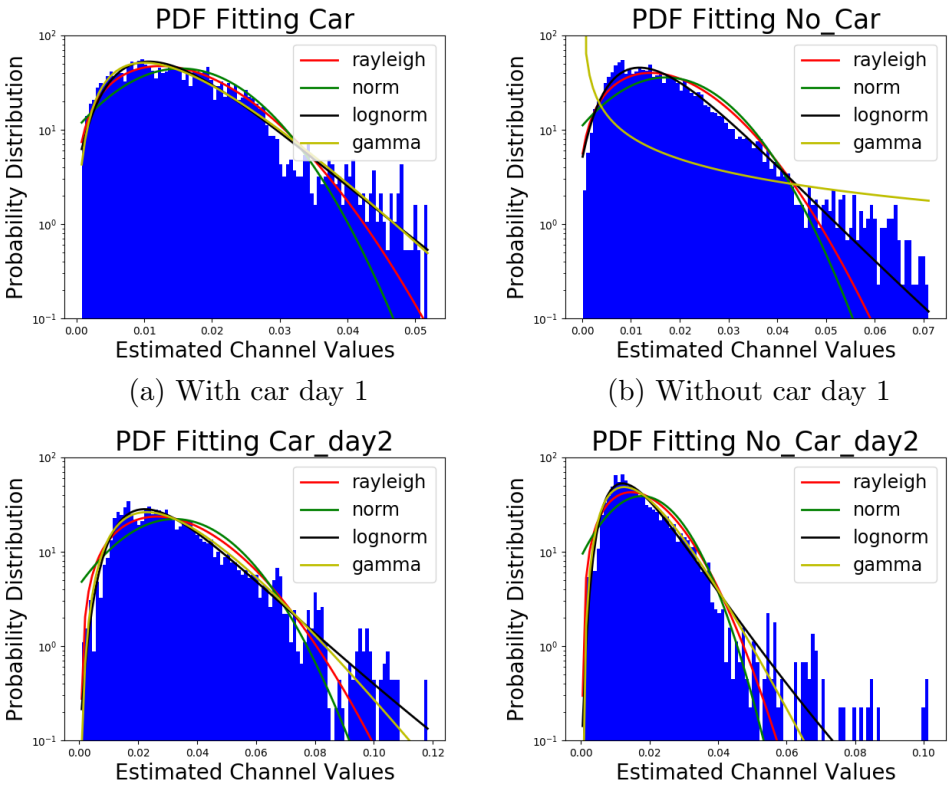
(c) With Jammie location 2

(d) Without Jammie location 2

Figure 5.6: Empirical PDF comparison observes the difference in dataset in the presence or absence of a bus in the vicinity of the receiver. The captures are taken in two different locations.

the distribution due to the presence or absence of one car in the vicinity of the receiver. Figure 5.7 contains the empirical and fitted distributions with respect to presence and absence of one car in the vicinity of the receiver. This data was captured in two different days to observe the differences in the distributions due to change in day. Figures 5.7a and 5.7b are captured on day 1 with and without a car in the vicinity of the receiver. Figures 5.7c and 5.7d are captured on day 2 with and without a car in the vicinity of the receiver. Although there are differences in the presence and absence of a car, the difference in the dataset captured in day 1 and day 2 is highly observable. This means that the environmental conditions play a major role in the captured distribution.

Some miscellaneous cases are also tested to observe the differences in the distri-



(a) With car day 1

(b) Without car day 1

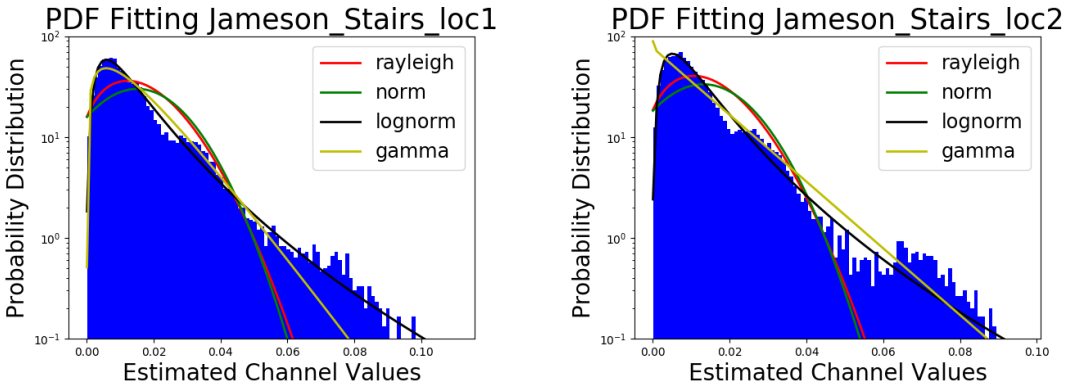
(c) With car day 2

(d) Without car day 2

Figure 5.7: Empirical PDF comparison between the conditions of presence and absence of one car in the vicinity of the receiver. These captures are taken in two days to observe the variation of the distribution due to change in day.

butions. One such case is shown in Figure 5.8. In this case the data is captured at the entrance stairway of a building of height 10 m. Figure 5.6a contains the distribution of captured data in slow wind conditions. Figure 5.6c contains the distribution of captured data in high wind conditions. Here, we can observe the similarities in the empirical distributions, although there are some differences due to the wind speed but the general pattern is similar.

This shows that the data captured from a single location will generally follow a similar pattern in a particular environmental condition (wind, humidity, temperature, etc.), thus aiding in providing enough information to map the surface and the environment at the time of the capture.



(a) Near a building of height 10 m in the presence of slow wind.

(b) Near a building of height 10 m high wind conditions.

Figure 5.8: Empirical PDF comparison of captured data near a building with different wind speeds.

At this point it is well-defined that each capture scenario can be defined by a particular distribution. Figures 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 and 5.8 showed the fitted distributions on the empirical data. Now, we extract the moments of the above mentioned empirical distributions. Since, log-normal and gamma distributions are also analysed the shape parameter specifically for these distributions are also extracted along with the moments.

Table 5.3: Moments of captured datasets.

Dataset	**Shape**		**Mean**	**Variance**	**Skew**	**Kurtosis**
	Log-normal	**Gamma**				
No CR	0.297	5.021	$2.45e^{-2}$	$1.55e^{-4}$	0.623	0.180
Vertical CR1	0.258	6.508	$1.23e^{-2}$	3.871	0.497	-0.242
Vertical CR2	0.250	6.493	$1.37e^{-2}$	$4.62e^{-5}$	0.503	-0.237
Dihedral CR	0.354	8.0	$1.21e^{-2}$	$3.61e^{-5}$	0.44	-0.237
Rain	0.382	3.036	$1.22e^{-2}$	$5.31e^{-5}$	0.921	0.386

Table 5.3: Moments of captured datasets.

Dataset	Shape		Mean	Variance	Skew	Kurtosis
	Log-normal	Gamma				
Humid No Rain	0.316	5.115	$3.99e^{-2}$	$4.12e^{-4}$	0.618	-0.188
Hot Day	0.343	4.185	$4.82e^{-2}$	$7.13e^{-4}$	0.921	0.386
Parking Full	0.486	0.235	$1.52e^{-2}$	$9.05e^{-5}$	1.503	3.368
Parking Empty	0.235	7.299	0.059	$8.44e^{-4}$	0.540	0.222
With Train	0.283	4.809	$4.57e^{-2}$	$5.85e^{-4}$	0.583	0.172
No Train	0.247	5.853	$5.16e^{-2}$	$6.80e^{-4}$	0.533	0.241
With Jammie location 1	0.631	1.999	$1.18e^{-2}$	$7.71e^{-5}$	1.949	5.516
With Jammie location 2	0.662	1.994	$1.34e^{-2}$	$9.79e^{-5}$	1.328	1.279
No Jammie location 1	0.595	0.245	$8.18e^{-3}$	$3.07e^{-5}$	1.848	4.700
No Jammie location 2	0.651	0.227	$1.24e^{-2}$	$8.25e^{-5}$	2.222	11.098
With Car day 1	0.395	2.912	$1.53e^{-2}$	$8.083e^{-5}$	0.970	1.199
With Car day 2	0.453	3.404	$3.23e^{-2}$	$3.5e^{-4}$	1.233	1.849
No Car day 1	0.389	0.234	$1.71e^{-2}$	$1.25e^{-4}$	1.442	2.823

Table 5.3: Moments of captured datasets.

Dataset	**Shape**		**Mean**	**Variance**	**Skew**	**Kurtosis**
	Log-normal	**Gamma**				
No Car day 2	0.466	3.348	$1.77e^{-2}$	$1.06e^{-4}$	2.031	7.994
Jameson Stairs location 1	0.762	1.612	$1.50e^{-2}$	$1.76e^{-4}$	2.150	6.232
Jameson Stairs location 2	0.768	0.996	$1.31e^{-2}$	$1.43e^{-4}$	2.378	7.830
Rhodes Memorial	0.296	5.546	$9.68e^{-3}$	$2.432e^{-5}$	0.951	1.757
Highland	0.903	1.277	$1.20e^{-2}$	$1.44e^{4}$	2.264	6.806

Table 5.3 contains the various moments extracted directly from the captured data. The shape parameters are extracted by fitting the empirical data to the expected distribution, thus there is a difference between the log-normal and gamma distribution shape parameters. These correlates with the information visible in the plots and gives us a better understanding of the changes in parameters due to the changes in environmental conditions.

From the figures and table provided, the differences in the distributions of the channel information due to the changes in the physical environment can be observed. This gives us an opportunity to utilize these differences and model a system that can differentiate from them. Although the differences are clear from the empirical PDF's, it is always beneficial to further analyse the received data and gain better understanding.

5.3 Chi-Square Test

The second set of analysis performed on the captured dataset is a type of goodness-of-fit test named the Chi-square test, more information regarding the test is given in [105, 106].

This is the test which takes two inputs, the observed data and the expected data and gives out a score following the Equation 5.19. The Degree of Freedom (DF) in this case is set to 100 since the length of the array is 101, because of the number of bins chosen to create the histogram, and the DF is defined as one less than the length of the array.

$$\chi^2 = \sum_{i=1}^{M} \frac{(O_i - E_i)^2}{E_i} \tag{5.19}$$

Here, O_i is the observed value and E_i is the expected value. In order to perform this analysis empirical PDF of the observed data and the distribution fitted expected data are passed through Equation 5.19 and the output of the test is recorded in Table 5.4.

A p-value is defined as the probability of obtaining a specific value equal to or higher than what actually is observed, maintaining the assumption that the model is correct. The points where the p-value is null can be rejected because this means the probability of getting a value similar or greater than that is null. All other cases in this test can be accepted. Log-normal distribution has the most consistent result with the test statistic and the p-value according to the data in Table 5.4.

5.4 Principal Components Analysis

In this analysis we investigate the clustering of the estimated channel values in different scenarios. In order to better visualize different sets of the captured data we use PCA.

Table 5.4: Chi-Square test values for the captured data compared to the fitted data for each of the expected distribution. Here, log-normal distribution provides the most consistent fit to all the datasets.

Chi-Square Value								
Dataset	**Rayleigh**		**Gaussian**		**Log-normal**		**Gamma**	
	Chi-Square	p-value	Chi-Square	p-value	Chi-Square	p-value	Chi-Square	p-value
No CR	568.67	$4.48e^{-89}$	726.25	$9.44e^{-121}$	565.77	$1.69e^{-88}$	566.71	$1.09e^{-88}$
Vertical CR1	1343.76	$1.92e^{-248}$	1316.92	$7.9e^{-243}$	1338.67	$2.23e^{-247}$	1341.42	$2.23e^{-248}$
Vertical CR2	1196.02	$1.41e^{-217}$	1174.87	$362e^{-213}$	1192.52	$7.61e^{-217}$	1194.58	$2.83e^{-217}$
Dihedral CR	1436.42	$7.16e^{-268}$	1411.50	$1.21e^{-262}$	1429.31	$2.23e^{-266}$	1433.02	$3.72e^{-267}$
Rain	1290.54	$2.63e^{-237}$	1236.88	$4.26e^{-237}$	1280.89	$2.72e^{-235}$	1285.16	$3.50e^{-236}$
Humid No Rain	467.01	$4.80e^{-69}$	457.44	$3.51e^{-67}$	464.22	$1.67e^{-68}$	465.41	$9.87e^{-69}$
Hot Day	310.75	$2.46e^{-39}$	318.90	$7.77e^{-41}$	307.33	$1.04e^{-38}$	308.27	$7.06e^{-39}$
Parking Full	1244.26	$1.22e^{-227}$	98209.53	0.0	661.44	$1.19e^{-107}$	$23657.61e^{10}$	0.0
Parking Empty	257.45	$1.05e^{-29}$	1218.27	$3.26e^{-222}$	246.60	$8.57e^{-28}$	247.76	$5.38e^{-28}$
With Train	352.02	$5.27e^{-47}$	9127.03	0.0	280.29	$8.73e^{-34}$	281.75	$4.75e^{-34}$
No Train	330.89	$4.67e^{-43}$	19388.17	0.0	260.57	$2.94e^{-30}$	261.10	$2.37e^{-30}$
With Jammie location 1	2710.11	0.0	$10402.59e^{1}$	0.0	734.88	$1.66e^{-122}$	$95601.90e^{10}$	0.0
With Jammie location 2	798.29	$1.41e^{-133}$	794.40	$1.28e^{-134}$	183.18	$2.49e^{-132}$	784.43	$1.38e^{-132}$
No Jammie location 1	2955.41	0.0	$19814.82e^{1}$	0.0	1112.68	$3.15e^{-200}$	$11318.45e^{10}$	0.0
No Jammie location 2	$85226.08e^{5}$	0.0	$48930.99e^{10}$	0.0	521.64	$9.29e^{-80}$	$12351.66e^{10}$	0.0
With Car day 1	975.44	$8.53e^{-172}$	931.99	$7.79e^{-163}$	968.93	$1.88e^{-170}$	973.73	$1.91e^{-170}$
With Car day 2	427.08	$2.67e^{-61}$	516.72	$8.67e^{-79}$	419.35	$8.22e^{-60}$	420.54	$4.85e^{-60}$
No Car day 1	704	$3.03e^{-116}$	694.84	$2.16e^{-114}$	697.70	$5.72e^{-115}$	$39635.69e^{10}$	0.0
No Car day 2	$27274.10e^{2}$	0.0	$71879.06e^{6}$	0.0	496.78	$7.23e^{-75}$	615.71	$1.8e^{-98}$
Jameson Stairs location 1	752438.53	0.0	87119627.61	0.0	447.26	$3.33e^{-65}$	452.96	$2.60e^{-66}$
Jameson Stairs location 2	199326.03	0.0	11569636.33	0.0	514.14	$2.79e^{-78}$	567.37	$8.13e^{-89}$
Rhodes Memorial	1502.05	$1.17e^{-281}$	1470.23	$5.70e^{-275}$	1499.21	$4.63e^{-281}$	1501.06	$1.89e^{-281}$
Highland	$5339.07e^{2}$	0.0	$3602.76e^{4}$	0.0	503.65	$3.23e^{-76}$	502.53	$5.36e^{-76}$

With this analysis we can reduce the dimensionality of the data and view the datasets from an angle that provides maximum information. Here we have observed different cluster formation of the datasets due to the changes in locations and environmental conditions. There are many ways to reduce the dimensionality of the data and calculate the principal components, for this implementation we have used Singular Value Decomposition (SVD) [30, 31, 107]. The major goals of PCA are to reduce the dimensionality of the data, extract the important information and analyse the structure of the data. The calculations for PCA are shown in [32].

In order to derive the principal components from the data we first need to generate the SVD equivalent of the dataset $\mathbf{A}$. The input matrix $\mathbf{A}$ has J sets of data explained by K variables, represented by $J \times K$. $\mathbf{A}$ has a rank L with $L \leq \min\{J, K\}$, then the SVD of $\mathbf{A}$ will be given by Equation 5.20.

$$\mathbf{A} = \mathbf{U}\Delta\mathbf{V}^{\mathrm{T}} \tag{5.20}$$

$$\mathbf{F} = \mathbf{U}\Delta \tag{5.21}$$

Here, $\mathbf{U}$ is $1 \times L$ matrix of left singular vectors, $\mathbf{V}$ is $K \times L$ matrix of right singular vectors and Δ is a diagonal matrix of singular values. The components for PCA are obtained from the data $\mathbf{A}$ using Equation 5.20. The principal component matrix $\mathbf{F}$ of dimension $J \times L$ is given by Equation 5.21. To get the coefficients of the linear combinations which are used to compute the factor scores the matrix $\mathbf{V}$ is used. The matrix can also be interpreted as the projection matrix because $\mathbf{A}$ times $\mathbf{V}$ gives the projection values of the observations on the principal components as:

$$\mathbf{F} = \mathbf{U}\Delta = \mathbf{U}\Delta\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{A}\mathbf{V} \quad . \tag{5.22}$$

Geometrically the components can also be represented by rotating the original axes and the matrix $\mathbf{A}$ can be interpreted as a product of the factor scores given

by Equation 5.23 as explained in [32]. Here $\mathbf{I}$ is the identity matrix.

$$\mathbf{A} = \mathbf{F}\mathbf{V}^{\mathrm{T}} \qquad \mathbf{F}^{\mathrm{T}}\mathbf{F} = \Delta^2 \quad \& \quad \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I} \tag{5.23}$$

In order to apply PCA the datasets for each observation are normalized and stacked together. Upon application of PCA on the empirical data the first two principal components are plotted to visualize the separation in the clusters.

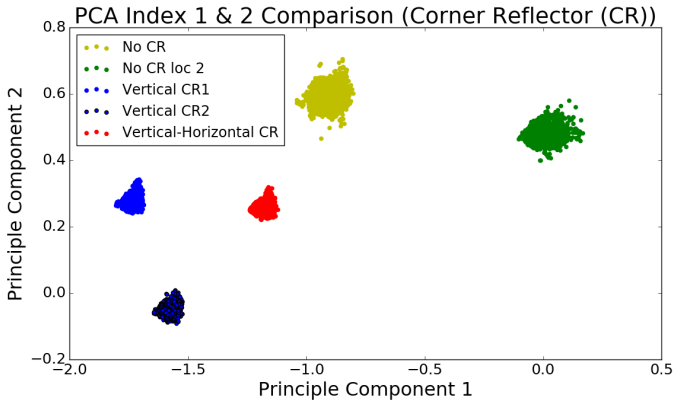


Figure 5.9: The clusters with different configuration of CR are placed at a distance of 3 m from the receiver.

The dimensions of the dataset is reduced from 40 to 2 major components and the relationship between the components are shown here. The clusters are clearly visible for each data. It can be noted here that taking more components we will get more information and hence will make scene classification even easier.

Figure 5.9 contains a particular set of captures taken with a CR. At first the captures are taken by placing the receiver in two different locations without the presence of any CR in the vicinity of the receiver. Then, a metal plate is placed near the receiver and data is captured, this acts as CR1, same is repeated with another metal plate named CR2. Then both the metal plates are placed at right angle to form a dihedral CR and data is captured for that. Each set of the captures here forms a different cluster thus proving that the data is not very correlated and can be separated for classification.

Figure 5.10 contains the clusters of the data captured at a sea beach near the

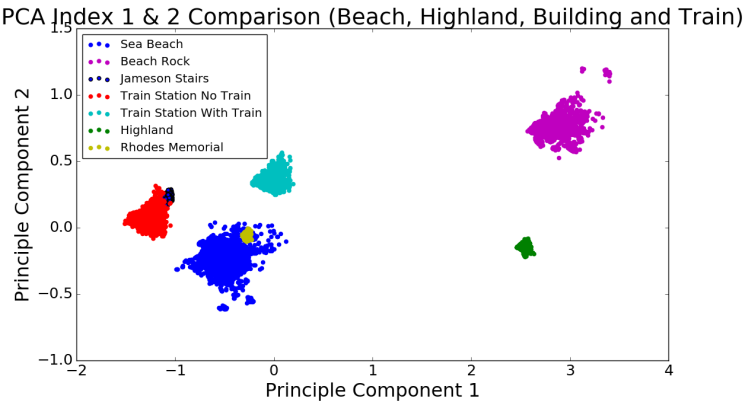


Figure 5.10: PCA clusters shown here include the captures taken near a beach, highland, building and train.

shore. The dataset named beach rock is captured at the shore while placing the receiver on the top of the rocks at a height of around 10 m from the sea level. The Jameson stairs dataset was captured at the entrance stairs of a building approximately 10 m tall, the building is called Jameson hall. Thus the dataset is named accordingly. The captures taken at a local train station in two different situations are given by train station no train and train station with train, the names are self-explanatory. The train station captures were taken by placing the receiver at a distance of 6 m away from the train tracks. The dataset named Highland, refers to a location where the captures are taken on the road next to a small hilly terrain and Rhodes memorial, is a location situated mid way up on to a mountain in Cape Town.

Figure 5.11 contains the clusters for the captures taken at a bus stop on the campus of UCT. These captures are taken at two different locations. The university bus service is known as "Jammie shuttle", as the name is given. The different scenarios for this capture are, in the presence of a Jammie approximately 4 m from the receiver antenna in location 1, 2. Keeping the location of the receiver same the captures are also taken in the absence of the Jammie.

Figure 5.12 contains the captures taken at a flat parking space when there are no cars and in the presence of cars. This set also shows the plots for a single car dataset, where the captures are taken with and without the presence of a car in

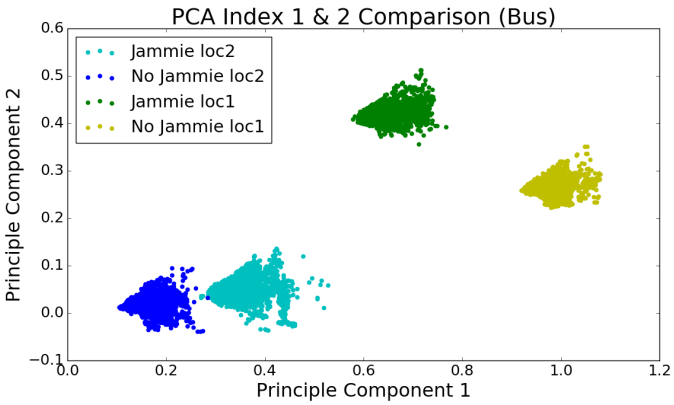


Figure 5.11: Different clusters formed by captures taken near a bus in different configurations and locations.

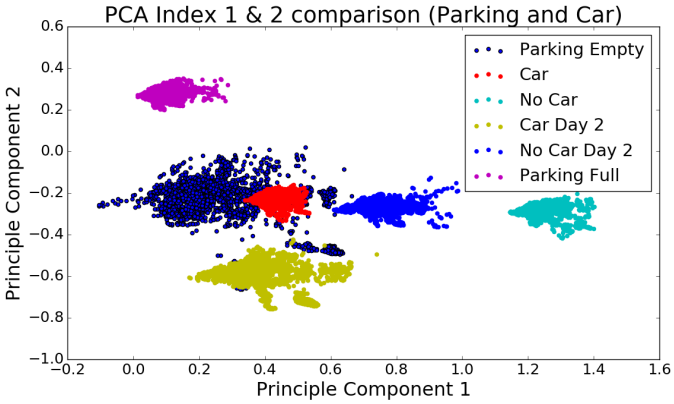


Figure 5.12: Clusters of captures taken at an open air parking lot when it is full/empty and with a car in proximity.

the same location on two different days (one at the morning 8 am and the other in the evening 6 pm) with different climatic conditions.

Figure 5.13 shows the difference in the clusters at the same location in different humidity conditions. The three conditions measured here are heavy rain, humid without rain and a hot day without rain. The receiver is placed at the same location to take the humidity captures.

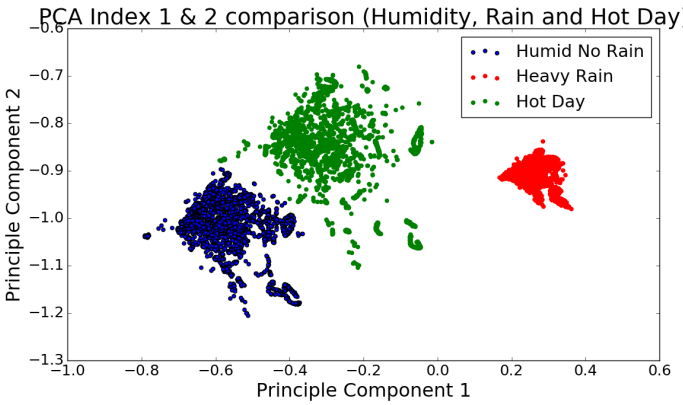


Figure 5.13: PCA of different climatic conditions at a single location. The conditions include high humidity without rain, heavy rain and hot day with very low humidity.

5.5 Summary

In this chapter, we have presented details of statistical analysis performed on the real-time data captured using the GSM-CommSense system. Table 5.1 contains a list detailing all the dataset analysed in this chapter. The analysis is performed in three steps.

In the first step of analysis, we plot the histogram of the empirical data and performed distribution fitting on it. The theoretical distribution used for the analysis were Rayleigh, Gaussian, log-normal and gamma distributions. These are well-known distributions and most widely used in the analysis of multipath data. In order to fit the distribution, we extract the parameters from the empirical data using MLE. The fitted distributions are then plotted on top of the histogram. On visual analysis of the fitted data, we observed that log-normal distribution had the most consistent fitting in all of the datasets. Although visual observation might be misled by the fact that the captured dataset were limited and the distribution might change when more data was captured.

Thus, we moved to the next step of the analysis. Here we took the empirical/observed data and the fitted/expected data and passed them through a chi-square goodness-of-fit test. This analysis presented us with a test statis-

tic and the probability of having this statistic within the expected distribution. The results from this test proved the visual observation correct that log-normal distribution had the most consistent fitting in all the datasets. Now that the distribution of the received data is following a pattern, we decided to visualize the data clusters.

In order to visualize the data clusters, we reduced the dimensions of the data and plotted the most prominent view angle. We did this using a well-known clustering algorithm called PCA. Here, we plotted the first two PCs of the transformed data. We found out that the data from each scenario was uniquely clustered.

Now that we have performed the statistical analysis on the dataset, the next step is to capture more data and classifying the data so that a prediction can be made relating the captured data to a particular set.

Chapter 6

Classification of GSM-CommSense data using learning algorithms[4]

6.1 Introduction

At this point, the GSM-CommSense system is designed and successfully implemented. The system receives the GSM broadcast information and extracts the channel information using the channel equalization techniques implemented in GSM system in real-time. The data is then statistically analysed and the variations due to change in the receiver's environment are observed. Now, we need to apply the GSM-CommSense system to detect and classify events of interest.

In this chapter, we implement two well-known supervised learning algorithms, viz. SVM and MLP, to classify the GSM-CommSense data. The classification accuracy is compared and the results are discussed. The learning algorithm with higher classification accuracy is used to detect events occurring across a wall.

[4] Based on Abhishek Bhatta, Amit Kumar Mishra and Jan Pidanic, "Classification of CommSense data using learning algorithms," in *International Conference on Radar Systems*, IET, 2017 and Abhishek Bhatta and Amit Kumar Mishra, "GSM–CommSense based through-the-wall sensing," in *Taylor and Francis Remote Sensing Letter*, 9, no. 3 (2017): 248-257.

6.2 Description of Terminologies in this Chapter

Terminologies used in the rest of the article.

- Event: An event is defined as a general test set-up in which data is captured and analysed.
- Test Scenario: A particular set-up under an event is defined here as a test scenario. Each event has multiple test scenarios.
- Set: A set represents data captured for 30 s, which is used for analysis. Each test scenario has multiple sets.
- Case: A case is used to categorise the particular set used for testing. Each case uses one set as test data and the other sets as training data.

6.3 Classifier Models

6.3.1 Support vector machine classifier

SVM is a supervised learning algorithm that takes a sample dataset and a predetermined kernel function as input, and generates a model for this sample. This model is then used to categorize the test dataset. The goal of SVM is to design a hyperplane or a set of hyperplanes in high-dimensional space that classifies all training vectors into different classes. Out of the multiple hyperplanes that can achieve the same task, the best choice will be the hyperplane that has the maximum separation from the nearest element of each class.

Let training data be represented as $D = \{(x_i, y_i)|i \in Z^+, 1 \leq i \leq n\}$, where $x_i \in R^d$ is training input points, $y_i \in \{1, -1\}$ are training labels, n is the size of the training data and d is the dimension of input data. In order to maximize the

geometric margin between two classes and minimize the error, the soft-margin SVM can be represented as:

$$\min_{\alpha \in R^F} \frac{1}{2}||\alpha||_2^2 + C\sum_{i=1}^{n} l(y_i, f_\alpha(x_i)) \quad . \tag{6.1}$$

Here, α is normal vector to the hyperplane separating the classes, $l(.)$ is a loss function, C is a regularization parameter weighing the smoothness and errors. $f_\alpha(x_i)$ is represented as:

$$f_\alpha(x_i) = \langle \phi(x_i), \alpha \rangle \quad . \tag{6.2}$$

Here, $\phi(x) : R^d \rightarrow R^F$ is a function mapping training data points from input space R^d to a new $F-$dimensional feature space R^F. For large F, the inner products of feature space can be calculated by a kernel function as shown in Equation 6.3, such as Radial Basis Function (RBF) as shown in Equation 6.4. Where x is the training input points and y is the training label.

$$k(x, y) = \langle \phi(x), \phi(y) \rangle \tag{6.3}$$

$$k(x, y) = \exp(-||x - y||_2^2/\sigma^2) \tag{6.4}$$

The representation of SVM shown here is referred from [108].

6.3.2 Multi-layer perceptron classifier

The network topology of a MLP with a single hidden layer is shown in Figure 6.1. Here, three layers are shown out of which the input and the output layers are visible to the users, whereas the hidden layer, as the name suggests, stays hidden. Each layer contains multiple nodes known as neurons. The nodes that are not a target of a connection are known as input neurons. Each neuron in the input

layer takes one feature from the input dataset. These features then act as the input for the subsequent hidden layers and this continues until it gets to the output layer.

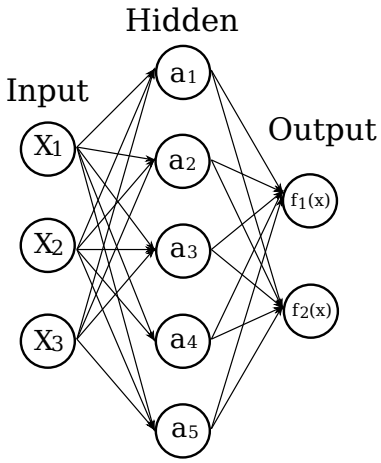


Figure 6.1: One hidden layer MLP classifier

In Figure 6.1, $\{X_1, X_2, X_3\}$ are the features of the input dataset, which then becomes the input information for the hidden layer thereby generating a bias $\{a_1, a_2, a_3, a_4, a_5\}$. The bias of the hidden layer then acts as the input features of the output layer. During the training period the weights of each neuron is set by a method called back-propagation. This process is used to adjust the weights of the input at the output layer.

$$\Delta w = \eta d X \tag{6.5}$$

During training, the weights Δw are set to get a particular output as shown in Equation 6.5. Here η is the learning rate that is usually less than 1, X is the input dataset and $d = Output_{predicted} - Output_{desired}$. The weight of each neuron is set individually by specific algorithms, such as gradient descent.

6.4 Comparison between SVM and MLP as the classifier Model

6.4.1 Experimental set-up

The different scenarios used to prove the claim of sensing the environment using the GSM-CommSense system is given below.

6.4.1.1 Environment classification

This scenario consists of four different environmental conditions named "Rain", "Medium Rain", "Humid No Rain" and "Hot Day". All the captures are taken with the receiver placed in the same location at different instances of time.

- Rain: This contains captures from a day when it rains heavily outside.
- Medium Rain: This contains captures from the same day as "Rain" captures but later in the day when the intensity of rain is considerably low.
- Humid No Rain: This contains captures from the day after the "Rain" and "Medium Rain" captures. This day is very humid and the clouds cover the region but there is no rain.
- Hot Day: This set of captures contains data from a very hot day without any humidity.

6.4.1.2 Vehicle detection (Train-Car)

This scenario consists of four different situations with and without the presence of a train or a car. To obtain one particular dataset the presence or absence of one vehicle is focused upon.

- With Train: This set of captures were taken at a train station with the receiver approximately 6 m from the train when the train was present in the platform.
- Without Train: This set of captures were taken just after the train left the station with the receiver at the same location as the "With Train" dataset. All other parameters were kept constant as far as possible.
- With Car: This set of capture was taken at an empty parking space with only one car in the direct vicinity of the receiver at about 3 m from the car. There were no other vehicles present around the receiver upto a distance of 100 m from the receiver. Although, there was a highway at about 105 m from the receiver in one direction.
- Without Car: This set of captures were taken at the same day and location as the dataset "With Car", when the car was not in the 100 m radius of the receiver.

Multiple sets of each of the above mentioned scenarios are captured for 30 s each with a gap of 20 s between consecutive sets. Each 30 s dataset contains 4500×40 points. The analysis presented here is performed on two of these sets, using one set to train the algorithm and the other set to test the prediction accuracy. In the case of "Environment" 4000×40 points from set one of each condition is used as training data and 1000×40 from capture set two of each condition is used as test data, thereby a total data used for training is of dimension 16000×40 and for testing is of dimension 4000×40. In case of "Train-Car" scenario the training data is made of 4000×40 points of set one from each situation and the test data contains 4000×40 points from set two of each situation, in total making 16000×40 points for training and 16000×40 points for testing.

6.4.2 Analysis of data

The data analysed in this section is captured using the GSM-CommSense system. The received signal is preprocessed to extract the channel impulse response.

Table 6.1: Confusion matrix showing classification of "Train-Car" dataset comparing SVM and MLP. Here the individual separation of a train and a car data is not great, but a general separation of vehicle and no vehicle is clearly visible.

Predicted Label	**SVM (in %)**				**MLP (in %)**			
With Train	58.95	38.40	1.00	1.65	55.30	41.775	0.675	2.25
With Car	57.075	37.575	1.85	3.50	53.975	38.95	1.70	5.375
Without Train	2.725	2.525	73.15	21.00	4.875	2.20	63.10	29.825
Without Car	11.075	10.30	43.475	35.15	13.225	11.90	39.65	35.225
	With Train	**With Car**	**Without Train**	**Without Car**	**With Train**	**With Car**	**Without Train**	**Without Car**
	True Label							

The channel estimation algorithm implemented in the GSM-CommSense system extracts 40 multipath channel state information from each frame of received signal. With the assumption that each of the multipath channel state information consists of a specific feature, is defined that the captured dataset consists of 40 features per frame.

The captured data is passed through SVM and MLP classifiers and the results are presented here. The kernel used for SVM classifier is RBF and the MLP has two hidden layers containing 10 neurons each. In case of MLP the layers and its size is chosen after performing multiple tests with different configuration and the one with optimum results are presented here.

Table 6.1 contain the confusion matrices for the classification between a presence and absence of a train or a car in the vicinity of the receiver. Although the train and car cannot be distinguished from the confusion matrix presented here, a general separation between the presence and absence of a vehicle is visible. The prediction of train is better than that of the car, mostly because the train is larger and reflects back stronger signal.

The MLP classifier provides a slightly poor classification between the train and a car, although even in this case the difference between the presence and absence of a vehicle is clear. In both the cases presented in Table 6.1 the prediction percentage of a car is lower and mixed with the features of the train, as both the observed objects have metallic surfaces having common features.

Table 6.2: Confusion matrix showing classification of "Environment" dataset comparing SVM and MLP. Here, the values distinguish between each of the events clearly.

Predicted Label	**SVM (in %)**				**MLP (in %)**			
Heavy Rain	53.90	38.60	7.50	0.00	53.00	43.40	3.30	0.30
Medium Rain	22.80	76.50	0.70	0.00	11.20	88.60	0.20	0.00
Humid No Rain	16.40	0.60	72.00	11.00	13.10	2.10	72.60	12.20
Hot Day	0.80	0.00	33.80	65.40	0.90	0.10	49.00	50.00
	Heavy Rain	**Medium Rain**	**Humid No Rain**	**Hot Day**	**Heavy Rain**	**Medium Rain**	**Humid No Rain**	**Hot Day**
	True Label							

Table 6.2 contains the confusion matrices to classify different environmental conditions as mentioned above. Predictions from the SVM and MLP classifier are shown here. In the SVM classifier, the clear separation between rain and no rain conditions is visible. There is some ambiguity between the medium rain and heavy rain conditions but that is expected as the water droplets having similar features. In case of MLP the prediction of hot day is mixed up with the prediction of humid no rain day. Although a similar prediction pattern is visible in SVM classification but the difference is almost double. It is observable that both the algorithms show ambiguity between the "Rain" and "Medium Rain" conditions which as explained above is due to the properties of water droplets.

Table 6.3: Simulation parameters.

	Train-Car Dataset		**Environment Dataset**	
	SVM	**MLP**	**SVM**	**MLP**
Error Rate (in %)	48.79	51.86	33.05	33.95
kernel	RBF	-	RBF	-
hidden layer size	-	$(10, 10)$	-	$(10, 10)$

The confusion matrices show details of the prediction errors and Table 6.3 contains the simulation parameters used to perform the analysis presented here and the error rate. The definition of error rate used here, for a particular set of output prediction, does not match the true value. This is calculated by $error(\text{in } \%) = \frac{100}{N}\sum(predicted - true)$, where N is the total number of er-

rors. The overall percentage of error in each of the algorithms for each scenario is not good. This is because the algorithms are predicting certain aspects of the scenario such as in "Train-Car" scenario it can separate the situation of presence and absence of vehicles and in "Environment" scenario it can separate rain and no rain conditions better than the others.

6.5 Through-the-wall Sensing Using SVM

6.5.1 Experimental set-up

The objective in the current work is to demonstrate the ability of the GSM-CommSense system to sense events across a brick wall of thickness 222 mm. There are in total two distinct events analysed in this work as described below.

(a) Presence (in motion/stationary) and absence of a person in 'Room 2' with the receiver placed in 'Room 1'

(b) Presence of two persons in 'Room 2' one of whom is carrying a weapon with the receiver placed in 'Room 1'.

The test is performed in a house with the layout as shown in Figure 6.2. The separations show different rooms in the house and the dotted lines are the location of the doors. The weapons used for the purpose of this test are toy gun and a knife covered with aluminium foil to increase the RCS. The weapons are shown in Figure 6.3. The aluminium foil is placed as the toy weapons are made of plastic and the idea is to mimic metallic concealed weapons. All the datasets of event (a) are captured within the duration of an hour keeping the external parameters such as location of the objects within the room, temperature of the room, etc., constant for the duration of the test. The experiments for event (b) were conducted on a different day from event (a). This test is also performed within the duration of an hour and the external parameters, as mentioned above, are kept constant for the duration of the test.

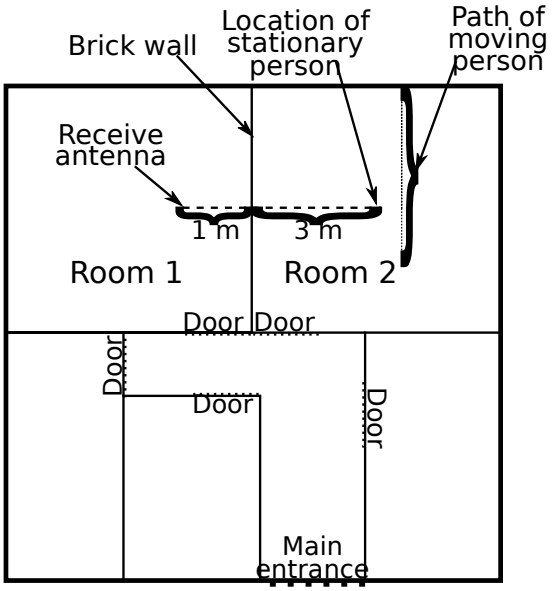


Figure 6.2: Layout of the experimental set-up.

Figure 6.3: Plastic toy weapons covered with aluminium foil to increase the reflectivity is used for testing.

The receiver is placed in 'Room 1' at a distance of 1 m from the wall and at a height of 1.12 m from the floor. The different test scenarios are described below.

6.5.2 Scenarios for event (a)

6.5.2.1 Person stationary

Person is standing in 'Room 2' at a distance of 3 m from the wall. Three sets of data are captured and analysed for this scenario. Each set contains data received

and sampled for a 30 s interval with a break of 20 s in between consecutive sets. This time interval is chosen to create statistical rigour to the received data along with enough samples per set to train the system.

6.5.2.2 No person

In this scenario there is no person in 'Room 2' maintaining all the other parameters of the test. This scenario also contains the same amount of datasets with the same duration as the previous case.

6.5.2.3 Person moving

In this scenario the person is walking in a loop in 'Room 2' at a distance of approximately 3 m from the wall, as shown in Figure 6.2 for the entire duration of the test. This scenario also contains same amount of datasets as in the other two test cases.

6.5.3 Scenarios for event (b)

6.5.3.1 Concealed weapon

There are two persons in 'Room 2', one of them concealing a weapon under the jacket and the other is not carrying a weapon. Location of the person concealing the weapon is 3 m from the brick wall and the location of the second person is randomized in the room, keeping their random position fixed for each capture. Each set is captured in the same way as the person-stationary scenario defined above.

6.5.3.2 No weapon

In this scenario there are two persons in 'Room 2' neither of them carrying weapon. Location of person-1 is still at 3 m from the wall and person-2 is ran-

domly located in the room remaining stationary for the duration of the capture.

6.5.3.3 Visible weapon

This scenario is similar to the concealed-weapon scenario but the only difference is that the weapons are visible and not concealed under jacket.

6.5.4 Analysis of data

Real-time data is captured using the GSM-CommSense system for each of the scenarios defined above. The wireless GSM signal is received and digitally sampled using an analogue-to-digital converter. The channel information is extracted from these samples and saved into a binary file for further analysis.

The analysis is performed in two steps, first the data is passed through PCA [32] to check if it contains any visually separable information. The limitations of this method is that we can only visually separate the information in three dimensions. This limits the percentage of information available for observation. Therefore, a supervised learning algorithm, SVM [109, 110] is used to predict each of the cases based on learning from the others. The major limitation of SVM is that it can only predict an outcome for an event that has been trained on, so for each new scenario the algorithm has to be trained separately. Each of these analysis is performed directly on the dataset captured by the GSM-CommSense system.

6.5.4.1 Results

Details of the datasets used in SVM for the purpose of training and testing is given in Table 6.4. After careful consideration and rigorous testing by varying the SVM parameters the penalty parameter is set to 1.0, a linear kernel is used and, the influence of the training example, given by γ, is set to 0.025 for event (a) and 0.0208 for event (b) to yield the best results. In case of event (a) there are in total 39000×40 data points used for training and 12000×40 data points

used for test. The training dataset is evenly distributed into the three scenarios defined above with 13000×40 points for each scenario. Since two sets are used for training, each set consists of 6500×40 points. In the case of event (b) the number of sample points from each frame changes from 40 to 48 thus in total 39000×48 data points are used for training and 12000×48 data points are used as test dataset. The additional 8 points per dataset create a noticeable difference in the prediction accuracy. The test dataset is evenly distributed among the three scenario and there are 4000×40 points in event (a) and 4000×48 points in event (b) per scenario.

Table 6.4: Description of the test set and training set to generate Table 6.5 and 6.6 along with the number of data points are used for training and testing. Each of the data point contains 39000 samples as training data and 12000 samples as test data captured in time with each sample containing 40 or 48 features depending on the case.

Case #	**Training**		**Test Set**	
	Set #	Data Points	Set #	Data Points
Case 1	Set $2a$ and $3a$	39000×40	Set $1a$	12000×40
Case 2	Set $1a$ and $3a$	39000×40	Set $2a$	12000×40
Case 3	Set $1a$ and $2a$	39000×40	Set $3a$	12000×40
Case 4	Set $2b$ and $3b$	39000×48	Set $1b$	12000×48
Case 5	Set $1b$ and $3b$	39000×48	Set $2b$	12000×48
Case 6	Set $1b$ and $2b$	39000×48	Set $3b$	12000×48

Table 6.5 contains the prediction output for event (a). The datasets captured for this event is represented by set $1a$, $2a$ and $3a$ in Table 6.4. The prediction results are shown in terms of percentage of correct classification. In case 1, set $1a$ is used as test set when the classifier is trained using sets $2a$ and $3a$. Similarly, in case 2, set $2a$ is used as test set when the classifier is trained using sets $1a$ and $3a$. Similarly, in case 3, set $3a$ is used as test set when the classifier is trained using sets $1a$ and $2a$. Since the number of datasets used for training and testing in each of the cases is the same, the average correct classification is calculated by summing the correct classification for each set and dividing by three. The results show that the lowest classification occurs (77.458%) in the scenario where no person is present in the room. Although, the difference in prediction is not

Table 6.5: Confusion matrix for detection of a person through a brick wall. The average correct classification provides an overview of the prediction output.

Predicted Label	**Case 1 (in %)**			**Case 2 (in %)**			**Case 3 (in %)**			**Average correct classification (in %)**
Person Stationary	86.00	4.825	9.175	85.725	3.525	10.75	86.125	4.975	8.90	85.95
No Person	16.925	71.925	11.15	8.25	82.50	9.25	12.85	77.95	9.2	77.458
Person Moving	9.775	6.675	83.55	5.775	4.275	89.95	10.55	6.475	82.975	85.491
	Person Stationary	**No Person**	**Person Moving**	**Person Stationary**	**No Person**	**Person Moving**	**Person Stationary**	**No Person**	**Person Moving**	
	True Label									

Table 6.6: Confusion matrix for detection of a person carrying weapon through a brick wall. The average correct classification provides an overview of the prediction output.

Predicted Label	**Case 4 (in %)**			**Case 5 (in %)**			**Case 6 (in %)**			**Average correct classification (in %)**
Concealed Weapon	91.675	6.75	1.575	95.85	2.525	1.625	98.10	1.50	0.40	95.208
No Weapon	6.55	92.475	0.975	1.35	97.875	0.775	0.925	97.675	1.40	96.008
Visible Weapon	0.575	0.875	98.55	0.20	0.775	99.025	0.525	1.275	98.20	98.591
	Concealed Weapon	**No Weapon**	**Visible Weapon**	**Concealed Weapon**	**No Weapon**	**Visible Weapon**	**Concealed Weapon**	**No Weapon**	**Visible Weapon**	
	True Label									

by a huge amount it can be explained by the fact that the reflections from a person's clothing might have similar features as the reflections from the curtains or the bed which the algorithm might be confusing as features of a person. Since there are only 40 feature points per frame, all the details are not captured and some similar features can confuse the algorithm to give ambiguous results.

Table 6.6 contains the prediction output for event (b). The datasets captured for this event are represented by set $1b$, $2b$ and $3b$ in Table 6.4. The number of correct classifications in percentage for each case is shown in Table 6.6. In case 4, set $1b$ is used as test data when the classifier is trained using sets $2b$ and $3b$. Similarly, in case 5, set $2b$ is used as test data when the classifier is trained using sets $1b$ and $3b$. Finally, in case 6, set $3b$ is used as test data

when the classifier is trained using sets $1b$ and $2b$. The number of feature points for this test scenario is 48 per frame and that shows a significant improvement in the prediction percentage. The scenario with the minimum average correct prediction percentage is the scenario with a concealed weapon (95.208%).

6.6 Summary

In this chapter, we presented a comparison between two supervised learning algorithms named SVM and MLP. The algorithms were used to detect two separate scenarios, environment classification to differentiate between different humidity conditions and vehicle detection. Both the algorithms classified the environment data with relatively high accuracy but the performance of SVM was better among the two. In the vehicle detection the classification of individual vehicle, a train or a car, was not clearly separable. Although if the classification accuracy was analysed in detection of the presence or absence of a particular vehicle, SVM outperformed MLP. Therefore, in both scenarios SVM had a better prediction accuracy.

Due to the accuracy of SVM, we decided to use it for detection of events occurring across a wall. The scenarios had been explained in detail and the classification accuracy had been presented. Since in this case we wanted to check the average accuracy of the classification, we decided to run the test on three sets of data each and calculated the average of the correct classifications. The results encouraging as a minimum average classification result obtained for detection of a person across a wall was 77.458% and in the case of detection of a person carrying a weapon across a wall was 95.208%.

The detailed analysis of the results had been presented thereby showing definitive proof that a GSM-CommSense system could be implemented in the real world.

Chapter 7

Conclusions

In this thesis, we introduced a novel architecture to sense the environment using the information within the telecommunication data-frames. We designed a system based on the proposed architecture and called it GSM-CommSense. We then used the system to capture data in real-time and performed statistical analysis on that data. We also explored some possible applications of the system and tested the performance in real world conditions.

In the following, we summarise the key contributions of our work and then, discuss some possible directions for future work.

7.1 Key Contributions

The scope of our work is as follows.

- After introducing the basic concepts related of GSM, SDR and commensal radars in Chapter 2, we designed some analytical tools necessary to understand the feasibility of the system in Chapter 3. Here, we performed the link budget analysis, calculated and simulated the ambiguity function and CRLB for a two path received signal model.

- In Chapter 4, we presented details for implementation of the GSM-CommSense system.
- In Chapter 5, we performed statistical analysis on the data captured by the GSM-CommSense system. We obtained encouraging results in distinguishing between different environmentat conditions.
- We then moved towards classification of the data using different learning algorithms, called SVM and MLP. We presented the comparison results between the two algorithms and explored an application of the GSM-CommSense system in Chapter 6.

Key contributions can be summarised as follows.

We started with a hypothesis that GSM based channel equalization modules and algorithms can be used to design and implement a commensal radar system to monitor the environment in real-time. In the quest to prove the hypothesis, we designed a set of analytical tools. Beginning with the link budget analysis, we performed all the power related calculations for the system, including the gains and losses. This, gave us an initial idea of feasibility of the system.

With this information we moved forward to calculate and simulate the ambiguity function for a two path ground reflection receive signal model. With the ambiguity function we analysed the effects of the scattered path signal on the received data in the presence of direct path information. Here, we also studied the effects of different transmit signal types along with different modulation schemes. We observed that the differences in the ambiguity function pattern, due to the scattered signal, were observable irrespective of the transmit signal type.

We then performed the CRLB calculations and simulations for the two path ground reflection receive signal model. We calculated the minimum variance of the time delay for the scattered path in the presence of direct path information. We observed that the minimum variance of time delay for the scattered path was dependent only on the scattered path parameters regardless of the presence of direct path. Then, we plotted the CRLB with respect to the distance of the point

of reflection from the receiver. We found the minimum variance of the order of 10^{-9} μs, it was because during the calculations, we considered idealistic scenario, and we believed in the real world, the value obtained would be measurable by available SDR hardware.

Having these analytical tools, we started designing the GSM-CommSense system. The initial design was made to be able to capture real world GSM base station broadcast signals. During the design phase, we tested multiple SDR hardware available in the market and found the BladeRF to be the best suited for this implementation. In the preliminary proof of concept design, we used a laptop to process the captured data and extracted the estimated channel values from the GSM frames. We used the least square estimation algorithm to perform channel estimation. Here, we presented the algorithm pseudocode that was finally implemented on the system. We captured data using this initial system and observed the change in the estimated channel values when the environmental condition changed.

Upon getting some initial positive results from the system we decided to build a hand-held system that could be carried to remote locations in order to capture more data. To build such a system, we decided to use a Raspberry Pi 3 as the processing device along with the BladeRF as the receiver. The software used for this final implementation was GNU Radio, which is an open source, free, software development tool-kit that provides signal processing blocks for SDR implementations. The final hand-held system was built to capture the GSM frames and extracted the channel information using the training sequence. Then, the estimated channel values were time-stamped and sent through Wi-Fi to a post processing device. If a post processing device was not attached, a provision was made to save the data on the Raspberry Pi memory card.

After capturing data from multiple locations, we performed statistical analysis on it to check the separability of the captured information. Here, we started with plotting the histogram of the empirical data on fitting it with known distributions. We observed, visually using the plots, that log-normal distribution had the most consistent fit on the data. To validate our observation, we performed

the chi-square goodness-of-fit test on the empirical and estimated distributions. This provided us with certainty that the best fit distribution for all the data analysed was log-normal distribution.

Next, we moved towards checking how clustered the captured data is for different environmental conditions. We performed PCA to find the best visualisation angle for the data. Here, we observed that the differences in the clusters for the first two PCs were clearly observable for data from different environmental conditions. At this point, we were certain that our original hypothesis was validated.

The validation of the hypothesis, gave us an opportunity for finding new applications of this newly formed system. We decided to explore some well-known machine learning algorithms to identify different environmental conditions. At this point, we realised that the system is extremely under-determined as the amount of channel information extracted per frame was not nearly enough to get a proper map of the environment. Then we decided to utilise a framework that was based on determining a particular event, we named it ASIN. With ASIN, we could train the GSM-CommSense system to detect a particular event and trigger an alarm based on that information. With this framework, we tested two widely accepted supervised learning algorithms, SVM and MLP. We found that SVM provided better prediction accuracy as compared to MLP in the available processing hardware. Thus, we decided to use SVM for further applications.

We used this final system for through-the-wall sensing application. Two different events were created to test the performance of the system as follows.

- Event (a): Detection of a person (stationary/ in motion) across a wall.
- Event (b): Detection of a person carrying a concealed weapon across a wall.

We found some encouraging results with minimum prediction accuracy for event (a) as 77.458% and for event (b) as 95.208%. We concluded that the change in performance result was mainly due to the high RCS of the concealed weapon.

All the source code used for the implementation of GSM-CommSense is available at [111]. The name of the repositories related to GSM-CommSense are "GSM_CE", "gr-gsm-CE", "Statistical_Analysis" and "gr-SVM".

7.2 Future Work

In this thesis, we designed a commensal radar system and checked its performance for a particular application. Further investigation of the system is necessary before it can be deployed in the real world. A list of suggestions for future research is provided hereafter.

- There is a need to define resolution for the system, as any current definition of resolution in the context of radar systems is not applicable for this system.
- At this point, the system consists of a single node. The changes in the performance of the system with addition of nodes can be investigated.
- The system is inherently incoherent, where any two receivers cannot be synchronized. Different possibilities will be explored to increase the detection capabilities. Also, if necessary, ways to make the system coherent could also be explored.
- This system is currently defined for GSM systems, it will be interesting to observe the behaviour of the system when operating under different communication protocols such as Long-term Evolution (LTE), DVB-T, etc..
- Some potential applications of the current system, that can be studied are:
  - Crop monitoring, such as soil moisture, temperature differences, detection of pests, etc..
  - Seismic data analysis, such as effect of earthquake in individual building.
  - To provide an aid for the visually impaired or elderly people.

  - Monitoring biological diversity within a region.

- Adding element of perception to the system can be a way to upgrade this system, where, it can detect an event on its own, without prior information. A system that can be called Cognitive-CommSense.

Appendix A

Details about the modulation scheme used in GSM system

A.1 Gaussian Minimum Shift Key Modulation

GMSK is a modulation scheme used in various digital communication systems [112]. It is most widely used in GSM cellular technology to carry information. It carries the digitally modulated signal through a channel while using the spectrum efficiently. GMSK is a derivative of Minimum Shift Key (MSK) modulation scheme. In MSK a signal is modulated by two half cycle sine waves.

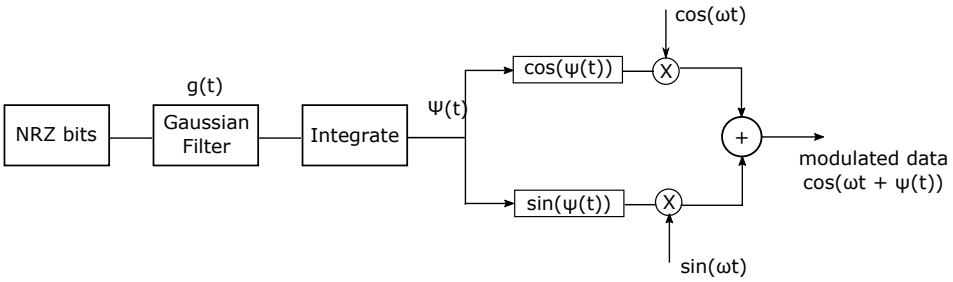


Figure A.1: GMSK Modulator

Figure A.1 contains the block diagram of a GMSK modulation performed in the GSM technology. As shown here, a non-return-to-zero (NRZ) sequence is passed through a Gaussian filter. The Gaussian low pass filter has an impulse response of:

$$g(t) = \frac{1}{2T}\left[(2\pi B\frac{t-\frac{T}{2}}{\sqrt{\ln 2}}) - Q(2\pi B\frac{t+\frac{T}{2}}{\sqrt{\ln 2}})\right] \quad , \tag{A.1}$$

where:
$g(t)$ = Gaussian Impulse Response
$Q(t)$ = Q-function
T = Time Period (bit duration)
B = Bandwidth

The Q-function is given as:

$$Q(t) = \frac{1}{\sqrt{2\pi}}\int_t^{\infty} e^{\frac{-\tau^2}{2}}\mathrm{d}\tau \quad . \tag{A.2}$$

The response of a Gaussian filter is shown in Figure A.2.

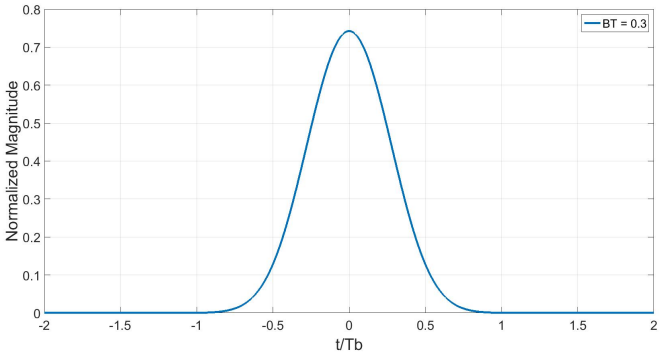


Figure A.2: Gaussian Filter Response

After passing the signal through a Gaussian filter, I and Q components are

separated out. These I and Q components are then multiplied by a carrier frequency as shown in A.1. The modulated signals are combined together and transmitted through a channel. In this case the medium of transfer is air, hence it is referred to as wireless channel.

A.2 Wireless Channel

As a signal travels through a medium, it is effected by information about the channel, which is generally referred to as impairments by telecommunication systems. A wireless channel is usually described as a filter with an impulse response and AWGN as shown in Figure A.3.

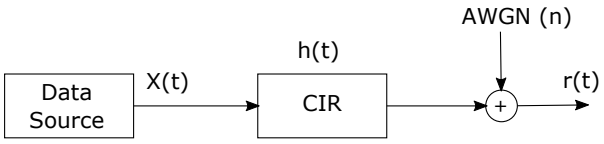


Figure A.3: Channel Model

For the purpose of this simulation, we assume the channel behaves as a 5^{th} order Finite Impulse Response (FIR) filter with AWGN. We have chosen the coefficients of the channel FIR filter, also known as CIR, from [113]. An OSR, $\eta = 1$, is chosen for simplicity. The CIR is shown in Table A.1

Table A.1: Coefficients of CIR $h(t)$, for GMSK with Time Bandwidth product (BT)=0.3

η	$h_{0,0}$	$h_{0,1}$	$h_{0,2}$	$h_{0,3}$	$h_{0,4}$
1	0.0007	0.2605	0.9268	0.2605	0.0007

A.3 Gaussian Minimum Shift Key Demodulation

A signal that travels through a wireless channel is received by a receiver. The received signal is represented as a convolution of the transmit signal and the channel noise, in this case AWGN as:

$$y(t) = x(t) * h(t) + n \quad , \tag{A.3}$$

where:
$y(t)$ = Received signal
$x(t)$ = Transmitted signal
$h(t)$ = Channel impairments
n = AWGN

The channel impairments are removed by a method called channel equalization. The process of channel equalization is explained in detail in Chapter 4. After the channel equalization is performed, the signal is passed through a GMSK demodulator as shown in Figure A.4.

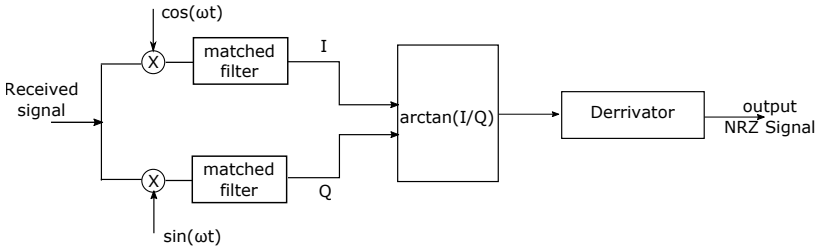


Figure A.4: GMSK Demodulator

After splitting the signal into its I and Q components and passing through a matched filter, the phase of the signal is equal to $\arctan(I/Q)$. Then finally the derrivator block extracts the NRZ sequence.

Appendix B

Basic Definitions

B.1 Transmit Power

The signal strength can be transmitted by a transmitter antenna. So that the received signal can be analysed and processed without any loss of information known as the transmit power. This value at the base station side depends on the required area or 'cell' to be covered. Typically, the transmit power from an outdoor base station ranges from a few Watts to 100 Watts, while for indoor base station is lower.

When looking at the mobile phones, its functionality is a bit different from the base station. It searches for the nearest base station and latches to it by responding to specific control signals at regular intervals of time. This is done in such a way that the power consumption of the phone is optimized keeping a good call quality. Typically, the power levels emitted by a mobile phone is 0.6 Watts up to the maximum level which is less than 3 Watts.

B.2 Directivity of Antenna

The simplest form of an antenna is the antenna that radiates in all directions uniformly known as isotropic antenna [114]. Irrespective of the transmitted power, the radio wave intensity decreases rapidly as it travels away from the antenna. In free space the intensity of the signal decreases at a rate of square of the distance travelled by it. In real-time it decreases much more quickly because of the presence of various obstacles in its path. Directivity of an antenna is the measure of how directional an antenna is. The directivity of an isotropic antenna is given as 1 (or 0 dB)

B.3 Free Space Loss or Path Loss

It is based on a very trivial concept that the intensity of a signal decreases with the distance travelled. Understanding path loss of an isotropic radiator will give an insight to the concept. Imagine a point source which radiates in all directions equally, making the radiation pattern as a sphere. As the distance from the source increases the surface area of the radiation also increases, which means the finite amount of energy radiated by the isotropic source will be spread on a larger surface area. Thus, at any given point on the surface the strength of the signal will be less than the point before that.

If we visualize a radiation pattern as shown in Figure B.1, we can write the power received at a surface A on the sphere with radius R is:

$$P_A = \left[\frac{A}{(4\pi R^2)}\right] P_{tx} \quad . \tag{B.1}$$

Since the cone of radiation increases with distance, while the area of cone's base intercepted by the received antenna decreases with the frequency, we can infer that the path loss is both a factor of distance and the frequency of operation.

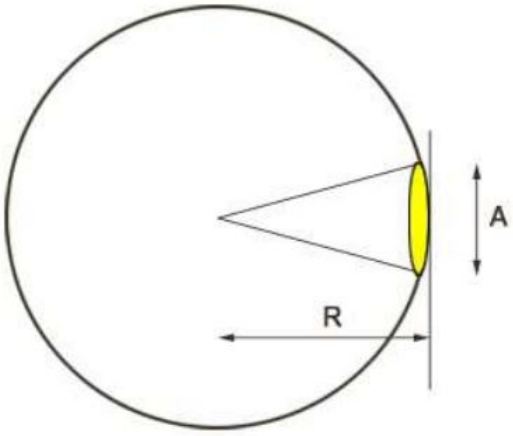


Figure B.1: Isotropic antenna radiation.

Thus, a simplified formula for free space path loss is derived as:

$$\text{Free Space Path Loss (dB)} = 32.44 + 20log_{10}(f) + 20log_{10}(D) \quad . \tag{B.2}$$

Here, f represents frequency of the signal in MHz, D is the distance between the transmitter and the receiver in km.

B.4 Range

Range can be defined as the maximum distance that a signal can travel with the desired signal strength. As a signal travels through free space, its power dissipates as a function of range. This effect occurs because of the spreading of the waves as it propagates, and can be calculated as:

$$L = 20log_{10}\left(\frac{4\pi D}{\lambda}\right) \quad . \tag{B.3}$$

Here, range is given by L, D is the distance between receiver and transmitter, λ is free space wavelength calculated as $\frac{c}{f}$. The speed of light is given by c and is

approximately $(3 * 10^8 m/s)$. f is frequency in Hz.

Equation B.3 is for the line of sight propagation, as losses due to other obstructions such as buildings, walls or indoor propagation can be significantly higher.

B.5 Bit Rate

It is the rate at which data transfers in a communication system. The upper limit of bit rate is given by Shannon's channel capacity theorem and is calculated as:

$$C = B * log_2(1 + S/N) \tag{B.4}$$

Here, C is the channel capacity measured in bits/s, B is the channel bandwidth measured in Hz, S is the signal strength in Watts and N is the noise power in Watts.

LTE theoretically has a net bit rate capacity of 100 Mbps downlink and 50 Mbps uplink if a 20 MHz channel is used and more if Multiple Input Multiple Output (MIMO) is used. The symbol rate varies with respect to the modulation scheme used. In case of downlink it is possible to select three types of modulation for LTE signals:

1. **QPSK** 2 bits per symbol
2. **16QAM** 4 bits per symbol
3. **64QAM** 6 bits per symbol

Each 15 kHz subcarrier in LTE is capable of transmitting 15 ksps, providing a raw symbol rate of 18 Msps at its 20 MHz system bandwidth (1200 subcarriers, 18 MHz) [115] [116].

B.6 Bandwidth and Resource Blocks

In the case of LTE it uses multiple bandwidths depending on the requirement of the users and the number of resource blocks used also vary with the bandwidth[117]. The relation is:

Channel Bandwidth (MHz)	1.4	3	5	10	15	20
Number of Resource blocks	6	15	25	50	75	100

B.7 Channel Noise

For all communication systems, channel noise depends on the bandwidth. All object that has heat emit RF signals in terms of random noise. The amount of radiation emitted can be calculated by

$$N = kTB \tag{B.5}$$

Here, N is the noise power in Watts, K is the Boltzmann Constant ($1.38 * 10^{-23}$J/K). The system temperature represented by T is usually assumed to be at 290 K and the channel bandwidth is given by B measured in Hz.

B.8 Receiver Noise

- Noise picked up by antenna: The antenna will pick up many signals other than the desired signal which can be termed as noise because it is unwanted in our case. Some examples of these types of noises are wireless communication noise, switching noise, etc..
- Noise generated by the receiver: The receiver has many devices which are capable of generating noises such as amplifier, filter, detector, etc..
- Noise power: it is similar to channel noise in the form of calculation, this also happens due to the temperature in the receiver.

B.9 Receiver Sensitivity

Receiver sensitivity is defined as the minimum signal strength required by the receiver to meet the Bit Error Rate (BER) requirement [118].

Example of reference sensitivity: −102 dBm for GSM, −117 dBm for Wideband Code Division Multiple Access (WCDMA).

B.10 Signal to Noise Ratio

It is basically the power difference between the received signal and the received noise. Normally it is calculated by:

$$\text{SNR (dBm)} = \text{Received signal power (dBm)} - \text{Received noise power (dBm)} \tag{B.6}$$

B.11 Gains

Transmitting antenna gain increases mainly by the focus of the transmitted power in a particular direction [119]. The receiver antenna gain depends on the ability of the receiver antenna to capture signals at particular directions as compared to others. Some other types of gains can also affect the signal such as using some specific type of algorithms that will reduce the SNR required by the receiver. The antenna gain is mainly dependent on the design of the antenna.

B.12 Losses

There are many types of losses in a communication system and these have to be anticipated in order to design a working system some of them are

- Transmission Loss: This includes the losses due to any impedance mismatches between the transmitter and the antenna or any loss occurring due to cabling issues.
- Propagation Loss: When a signal travels through a medium for some distance it encounters losses because it radiates out some energy for travelling a certain distance.
- Shadowing/Fading Loss: Shadowing occurs if an obstacle comes in between the transmitter and the receiver such as a hill. Due to the movement of both/either of the transmitter and receiver if signal power is lost it is termed as slow fading
- Reception Loss: It is generally due to the losses in cabling and other factors at the receiver. Typically, a mobile receiver does not have any cabling but still some losses occur due to the orientation of buildings or other environmental conditions.

Bibliography

[1] W. contributors, "Software-defined radio," 2018. [Online; accessed 20-Jan-2018]. xii, 21

[2] H. Griffiths, "Passive bistatic radar," *RTO Educational Notes-Lecture Series RTO-EN-SET-133 Waveform Diversity for Advanced Radar Systems, Brno*, 2009. 1, 30

[3] P. Krysik, K. Kulpa, M. Baczyk, Ł. Maślikowski, and P. Samczynski, "Ground moving vehicles velocity monitoring using a GSM based passive bistatic radar," in *Radar (Radar), 2011 IEEE CIE International Conference on*, vol. 1, pp. 781–784, IEEE, 2011. 1, 28, 30

[4] P. Samczynski, K. Kulpa, M. Malanowski, P. Krysik, and L. Maślikowski, "A concept of GSM-based passive radar for vehicle traffic monitoring," in *Microwaves, Radar and Remote Sensing Symposium (MRRS), 2011*, pp. 271–274, IEEE, 2011. 1, 28, 30

[5] H. Griffiths, I. Darwazeh, and M. Inggs, "Waveform design for commensal radar," in *Radar Conference (RadarCon), 2015 IEEE*, pp. 1456–1460, IEEE, 2015. 1, 16, 27

[6] M. Inggs, C. Tong, R. Nadjiasngar, G. Lange, A. Mishra, and F. Maasdorp, "Planning and design phases of a commensal radar system in the FM broadcast band," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 29, no. 7, pp. 50–63, 2014. 1, 26

[7] M. R. Inggs and C. Tong, "Commensal radar using separated reference and surveillance channel configuration," *Electronics Letters*, vol. 48, no. 18, pp. 1158–1160, 2012. 1, 6, 16, 25, 30

[8] S. Haykin, "Cognitive radar: a way of the future," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 30–40, 2006. 1

[9] J. R. Guerci, "Cognitive radar: A knowledge-aided fully adaptive approach," in *Radar Conference, 2010 IEEE*, pp. 1365–1370, IEEE, 2010. 1

[10] W. A. Holm, "Continuous wave radar," in *Principles of Modern Radar*, pp. 397–421, Springer, 1987. 1

[11] J. L. Geisheimer, W. S. Marshall, and E. Greneker, "A continuous-wave (CW) radar for gait analysis," in *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, vol. 1, pp. 834–838, IEEE, 2001. 1

[12] M. Inggs, A. van der Byl, and C. Tong, "Commensal radar: Range-doppler processing using a recursive DFT," in *Radar (Radar), 2013 International Conference on*, pp. 292–297, IEEE, 2013. 1, 16, 26

[13] M. R. Inggs, C. Tong, A. K. Mishra, and F. Maasdorp, "Modelling and simulation in commensal radar system design," in *Radar Systems (Radar 2012), IET International Conference on*, pp. 1–5, IET, 2012. 1, 26

[14] C. Coleman and H. Yardley, "Passive bistatic radar based on target illuminations by digital audio broadcasting," *Radar, Sonar & Navigation, IET*, vol. 2, no. 5, pp. 366–375, 2008. 1, 30

[15] J. Thomas, H. Griffiths, and C. Baker, "Ambiguity function analysis of digital radio mondiale signals for HF passive bistatic radar," *Electronics Letters*, vol. 42, no. 25, pp. 1482–1483, 2006. 1

[16] J. Homer, K. Kubik, B. Mojarrabi, I. Longstaff, E. Donskoi, and M. Cherniakov, "Passive bistatic radar sensing with LEOS based transmitters,"

in *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 1, pp. 438–440, IEEE, 2002. 1

[17] D. V. Gupta, "Software-Defined Radio," July 9 2013. US Patent 8,483,626. 2, 20

[18] W. H. Tuttlebee, *Software Defined Radio: enabling technologies*. John Wiley & Sons, 2003. 2, 20

[19] F. K. Jondral, "Software-Defined Radio: basics and evolution to cognitive radio," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 3, pp. 275–283, 2005. 2, 20

[20] H. Arslan, *Cognitive radio, software defined radio, and adaptive wireless systems*. Springer, 2007. 2, 20

[21] "CellC coverage map." http://www.cellc.co.za/network-coverage-map, accessed December 2014. 3

[22] "MTN coverage map." https://www.mtn.co.za/Pages/Coverage-map.aspx, accessed December 2014. 3

[23] "Telkom coverage map." http://www.telkom.co.za/coverage/, accessed December 2014. 3

[24] "Vodacom coverage map." http://www.vodacom.co.za/vodacom/coverage-map, accessed December 2014. 3

[25] M. Drutarovsky, "GSM channel equalization algorithm-modern DSP co-processor approach," *Radioengineering*, vol. 8, no. 4, 1999. 4, 17

[26] R. O. E *et al.*, "Communication by radar beams," Aug. 5 1969. US Patent 3,460,139. 6

[27] M. Tsuruno, "Wireless communication device and radar detection method therefor," May 11 2010. US Patent 7,715,801. 6, 30

[28] A. Bhatta and A. K. Mishra, "Implementation of GSM channel estimation using open-source SDR environment," in *2015 International Conference on Microwave, Optical and Communication Engineering (ICMOCE)*, pp. 322–325, IEEE, 2015. 6, 13

[29] A. Bhatta and A. K. Mishra, "GSM-based CommSense system to measure and estimate environmental changes," *IEEE Aerospace and Electronic Systems Magazine*, vol. 32, no. 2, pp. 54–67, 2017. 6, 12

[30] H. Abdi, "The eigen-decomposition: Eigenvalues and eigenvectors," *Encyclopedia of measurement and statistics*, pp. 304–308, 2007. 7, 103

[31] H. Abdi, "Singular value decomposition (SVD) and generalized singular value decomposition," *Encyclopedia of measurement and statistics. Thousand Oaks (CA): Sage*, pp. 907–912, 2007. 7, 103

[32] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010. 7, 103, 104, 120

[33] A. K. Mishra, "Application Specific Instrumentation (ASIN): A Bio-inspired Paradigm to Instrumentation using recognition before detection," *ArXiv e-prints*, Oct. 2016. 7, 62

[34] A. K. Mishra and S. Sardar, "Application specific instrumentation and its feasibility for UWB sensor based breast cancer diagnosis," in *Power, Control and Embedded Systems (ICPCES), 2010 International Conference on*, pp. 1–4, IEEE, 2010. 7

[35] S. Sardar and A. K. Mishra, "ASIN-based UWB radar for sludge monitoring," *Access, IEEE*, vol. 2, pp. 290–300, 2014. 7

[36] A. K. Mishra, "Monitoring changes in an environment by means of communication devices," Oct. 27 2016. WO Patent App. PCT/IB2016/052,235. 8, 30

[37] A. Bhatta and A. K. Mishra, "GSM-CommSense: A GSM based environment sensing system using software defined radio." https://youtu.be/AGdIa7L56KY. 12

[38] A. Bhatta and A. K. Mishra, "GSM–CommSense-based through-the-wall sensing," *Remote Sensing Letters*, vol. 9, no. 3, pp. 248–257, 2018. 13

[39] A. Bhatta and A. K. Mishra, "Ambiguity function and Cramér-Rao Lower Bound calculation for CommSense system," *IEEE Transactions on Aerospace and Electronic Systems, (Accepted)*, 2018. 13

[40] A. Bhatta and A. K. Mishra, "GSM based hand-held CommSense-sensor for environment monitoring," in *2016 11th International Conference on Industrial and Information Systems (ICIIS)*, pp. 360–364, IEEE, Dec 2016. 13

[41] A. Bhatta, A. K. Mishra, and J. Pidanic, "Classification of CommSense data using learning algorithms," 2017. IET Radar Conference. 13

[42] European Telecommunications Standards Institute, "ETSI website." http://www.etsi.org/, accessed January 2017. 16

[43] F. Hillebrand, *GSM and UMTS: the creation of global mobile communication*. John Wiley & Sons, Inc., 2002. 16, 76

[44] J. Lempiäinen and M. Manninen, *Radio interface system planning for GSM/GPRS/UMTS*. Springer Science & Business Media, 2001. 16, 18, 76

[45] T. Kos, M. Grgic, and G. Sisul, "Mobile user positioning in GSM/UMTS cellular networks," in *Multimedia Signal Processing and Communications, 48th International Symposium ELMAR-2006 focused on*, pp. 185–188, IEEE, 2006. 17

[46] M. S. Braasch, "Multipath," in *Springer Handbook of Global Navigation Satellite Systems*, pp. 443–468, Springer, 2017. 17

[47] V. M. Baronkin, Y. V. Zakharov, and T. C. Tozer, "Cramer-rao lower bound for frequency estimation in multipath rayleigh fading channels," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 4, pp. 2557–2560, IEEE, 2001. 17, 42

[48] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 240–254, 2004. 17

[49] T. Petermann and K.-D. Kammeyer, "Blind channel estimation in GSM receivers: a comparison of HOS and SOCS based approaches," in *Higher-Order Statistics, 1999. Proceedings of the IEEE Signal Processing Workshop on*, pp. 80–84, IEEE, 1999. 17

[50] L. Pu, J. Liu, Y. Fang, W. Li, and Z. Wang, "Channel estimation in mobile wireless communication," in *Communications and Mobile Computing (CMC), 2010 International Conference on*, vol. 2, pp. 77–80, IEEE, 2010. 17

[51] M. Pukkila, "Channel estimation modeling," *Nokia Research Center*, 2000. 17, 66

[52] "Digital cellular telecommunications system (phase 2+); general packet radio service (GPRS); mobile station (ms) - base station system (bss) interface;radio link control / medium access control (rlc/mac) protocol." ETSI TS 145 001 V10.1.0, (2010-03). (3GPP TS 44.060 version 9.3.0 Release 9). 18, 20, 76

[53] T. Halonen, J. Romero, and J. Melero, *GSM, GPRS and EDGE performance: evolution towards 3G/UMTS*. John Wiley & Sons, 2004. 18

[54] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson, "EDGE: Enhanced data rates for GSM and TDMA/136 evolution," *IEEE Personal Communications*, vol. 6, no. 3, pp. 56–66, 1999. 18

[55] O. Gabbard and P. Kaul, "Time-division multiple access," in *EASCON'74; Electronics and Aerospace Systems Convention*, pp. 179–184, 1974. 18

[56] T. S. Chan, "Time-division multiple access," *Handbook of Computer Networks: LANs, MANs, WANs, the Internet, and Global, Cellular, and Wireless Networks, Volume 2*, pp. 769–778, 2007. 18

[57] R. M. Gagliardi, "Frequency-division multiple access," in *Satellite Communications*, pp. 215–250, Springer, 1991. 18

[58] "Digital cellular telecommunications system (phase 2+); physical layer on the radio path." ETSI TS 145 001 V10.1.0, (2012-01). (3GPP TS 45.001 version 10.1.0 Release 10). 20, 76

[59] "Digital cellular telecommunications system (phase 2+); multiplexing and multiple access on the radio path.." ETSI TS 145 002 V12.3.0, (2015-01). (3GPP TS 45.002 version 12.3.0 Release 12). 20, 65, 68, 76

[60] "Digital cellular telecommunications system (phase 2+); radio transmission and reception." ETSI TS 145 005 V12.4.0, (2015-01). (3GPP TS 45.005 version 12.4.0 Release 12). 20, 65, 76

[61] D. C. Tucker, G. Tagliarini, *et al.*, "Prototyping with GNU radio and the USRP-where to begin," in *Southeastcon, 2009. SOUTHEASTCON'09. IEEE*, pp. 50–54, IEEE, 2009. 22

[62] GNU Radio Website, "GNU radio website." http://www.gnuradio.org, accessed October 2015. 22

[63] M. Afendykiw, J. Boyle, and C. Hendrix, "Bistatic passive radar," May 21 1974. US Patent 3,812,493. 23

[64] J. D. Sahr and F. D. Lind, "The Manastash Ridge radar: A passive bistatic radar for upper atmospheric radio science," *Radio Science*, vol. 32, no. 6, pp. 2345–2358, 1997. 23

[65] F. D. Lind, J. D. Sahr, and D. M. Gidner, "First passive radar observations of auroral e-region irregularities," *Geophysical research letters*, vol. 26, no. 14, pp. 2155–2158, 1999. 23

[66] M. A. Ringer and G. J. Frazer, "Waveform analysis of transmissions of opportunity for passive radar," in *Signal Processing and Its Applications, 1999. ISSPA'99. Proceedings of the Fifth International Symposium on*, vol. 2, pp. 511–514, IEEE, 1999. 23

[67] C. L. Zoeller, M. Budge, and M. J. Moody, "Passive coherent location radar demonstration," in *System Theory, 2002. Proceedings of the Thirty-Fourth Southeastern Symposium on*, pp. 358–362, IEEE, 2002. 24

[68] H. Griffiths, C. Baker, H. Ghaleb, R. Ramakrishnan, and E. Willman, "Measurement and analysis of ambiguity functions of off-air signals for passive coherent location," *Electronics Letters*, vol. 39, no. 13, pp. 1005–1007, 2003. 24

[69] H. Griffiths and C. Baker, "Passive coherent location radar systems. part 1: Performance prediction," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 152, no. 3, pp. 153–159, 2005. 24

[70] C. Baker, H. Griffiths, and I. Papoutsis, "Passive coherent location radar systems. part 2: Waveform properties," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 152, no. 3, pp. 160–168, 2005. 24

[71] D. Gould, R. Pollard, C. Sarno, and P. Tittensor, "A multiband passive radar demonstrator," in *Radar Symposium, 2006. IRS 2006. International*, pp. 1–4, IEEE, 2006. 24

[72] D. Gould, R. Pollard, C. Sarno, and P. Tittensor, "Developments to a multiband passive radar demonstrator system," 2007. IET. 24

[73] H. Griffiths and C. Baker, "The signal and interference environment in passive bistatic radar," in *Information, Decision and Control, 2007. IDC'07*, pp. 1–10, IEEE, 2007. 24

[74] D. Shephard and G. Richards, "Passive target tracking using transmitters of opportunity," 2008. IET. 24

[75] C. Coleman and H. Yardley, "DAB based passive radar: Performance calculations and trials," in *Radar, 2008 International Conference on*, pp. 691–694, IEEE, 2008. 24

[76] M. Malanowski, K. Kulpa, and J. Misiurewicz, "PaRaDe-passive radar demonstrator family development at warsaw university of technology," in *Microwaves, Radar and Remote Sensing Symposium, 2008. MRRS 2008*, pp. 75–78, IEEE, 2008. 25

[77] M. Malanowski and K. Kulpa, "Digital beamforming for passive coherent location radar," in *Radar Conference, 2008. RADAR'08. IEEE*, pp. 1–6, IEEE, 2008. 25

[78] F. Colone, D. O'hagan, P. Lombardo, and C. Baker, "A multistage processing algorithm for disturbance removal and target detection in passive bistatic radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, 2009. 25

[79] F. Berizzi, M. Martorella, D. Petri, M. Conti, and A. Capria, "USRP technology for multiband passive radar," in *Radar Conference, 2010 IEEE*, pp. 225–229, IEEE, 2010. 25

[80] D. K. P. Tan, M. Lesturgie, H. Sun, and Y. Lu, "Target detection performance analysis for airborne passive bistatic radar," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pp. 3553–3556, IEEE, 2010. 25

[81] M. Conti, F. Berizzi, M. Martorella, E. Dalle Mese, D. Petri, and A. Capria, "High range resolution multichannel DVB-T passive radar," *IEEE Aerospace and Electronic Systems Magazine*, vol. 27, no. 10, pp. 37–42, 2012. 25

[82] D. Petri, A. Capria, M. Conti, F. Berizzi, M. Martorella, and E. Dalle Mese, "High-range resolution multichannel DVB-T passive radar: aerial target

detection," *International Journal of Microwave and Wireless Technologies*, vol. 4, no. 2, pp. 147–153, 2012. 25

[83] J. W. Brown, *FM airborne passive radar*. PhD thesis, UCL (University College London), 2013. 26

[84] F. Maasdorp, J. Cilliers, M. Inggs, and C. Tong, "Simulation and measurement of propeller modulation using FM broadcast band commensal radar," *Electronics Letters*, vol. 49, no. 23, pp. 1481–1482, 2013. 26

[85] F. Maasdorp, R. Nadjiasngar, and M. Inggs, "A cramer rao analysis on receiver placement in a FM band commensal radar system based on doppler only measurements," in *Radar Conference (Radar), 2014 International*, pp. 1–6, IEEE, 2014. 26

[86] A. K. Mishra and M. Inggs, "FOPEN capabilities of commensal radars based on whitespace communication systems," in *Electronics, Computing and Communication Technologies (IEEE CONECCT), 2014 IEEE International Conference on*, pp. 1–5, IEEE, 2014. 26

[87] F. Maasdorp, J. Cilliers, M. Inggs, and C. Tong, "FM band commensal radar technology used for the detection of small aircraft and the measurement of propeller modulation," in *Radar Conference (RadarCon), 2015 IEEE*, pp. 0664–0668, IEEE, 2015. 27

[88] A. Nicol, M. Inggs, and D. O'Hagan, "Evaluating commensal sensors for detecting objects of interest in the low earth orbit," in *Radar Conference (RadarConf), 2016 IEEE*, pp. 1–4, IEEE, 2016. 27

[89] D. K. Tan, H. Sun, Y. Lu, and W. Liu, "Feasibility analysis of GSM signal for passive radar," in *Radar Conference, 2003. Proceedings of the 2003 IEEE*, pp. 425–430, IEEE, 2003. 27

[90] D. K. Tan, H. Sun, Y. Lu, M. Lesturgie, and H. Chan, "Passive radar using global system for mobile communication signal: theory, implementation and measurements," in *Radar, Sonar and Navigation, IEE Proceedings-*, vol. 152, pp. 116–123, IET, 2005. 27

[91] H. Sun, D. K. Tan, and Y. Lu, "Aircraft target measurements using a GSM-based passive radar," in *Radar Conference, 2008. RADAR'08. IEEE*, pp. 1–6, IEEE, 2008. 28

[92] P. Krysik, P. Samczynski, M. Malanowski, L. Maslikowski, and K. Kulpa, "Velocity measurement and traffic monitoring using a GSM passive radar demonstrator," *IEEE Aerospace and Electronic Systems Magazine*, vol. 27, no. 10, pp. 43–51, 2012. 28

[93] K. Chetty, Q. Chen, and K. Woodbridge, "Train monitoring using GSM-R based passive radar," in *Radar Conference (RadarConf), 2016 IEEE*, pp. 1–4, IEEE, 2016. 28

[94] J. Zyren and A. Petrick, "Tutorial on basic link budget analysis," *Application Note AN9804, Harris Semiconductor*, 1998. 31

[95] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elec. Commun. Lab*, vol. 16, no. 9, pp. 825–73, 1968. 32, 35

[96] M. Hatay, "Empirical formula for propagation loss in land mobile radio services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980. 32, 35, 36, 37

[97] A. P. Barsis, "Radio wave propagation over irregular terrain in the 76-to 9200-MHz frequency range," *IEEE Transactions on Vehicular Technology*, vol. 20, no. 3, pp. 41–62, 1971. 32

[98] W. R. Young, "Comparison of mobile radio transmission at 150, 450, 900, and 3700 mc," *Bell Labs Technical Journal*, vol. 31, no. 6, pp. 1068–1085, 1952. 32

[99] "LTE radio link budget @ https://sites.google.com/site/lteencyclopedia/lte-radio-link-budgeting-and-rf-planning," 39

[100] J. G. Proakis, "Digital communications fourth edition," 2001. McGraw-Hill Companies, Inc., New York, NY. 42

[101] N. Levanon and E. Mozeson, *Radar signals*. John Wiley & Sons, 2004. 43

[102] S. M. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*, vol. 3. Pearson Education, 2013. 54

[103] G. R. Curry, *Radar essentials: a concise handbook for radar design and performance*. The Institution of Engineering and Technology, 2012. 58

[104] "Raspberry pi compatible memory cards." http://elinux.org/RPi_SD_cards, accessed April 2016. 80

[105] A. Vardasbi, M. Salmasizadeh, and J. Mohajeri, "Multiple-chi-square tests and their application on distinguishing attacks," in *Information Security and Cryptology (ISCISC), 2011 8th International ISC Conference on*, pp. 55–60, IEEE, 2011. 101

[106] M. Kendall and A. Stuart, "The advanced theory of statistics vol. 2, inference and relationship. charles griffin and co., ltd," 1961. 101

[107] Y. Takane, "Relationships among various kinds of eigenvalue and singular value decompositions," in *New developments in psychometrics*, pp. 45–56, Springer, 2003. 103

[108] H. Li, L. Xiong, L. Ohno-Machado, and X. Jiang, "Privacy preserving RBF kernel support vector machine," *BioMed research international*, vol. 2014, 2014. 111

[109] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013. 120

[110] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine learning*, vol. 46, no. 1-3, pp. 131–159, 2002. 120

[111] A. Bhatta, "GSM-CommSense: Source code." https://github.com/bhattaomatic. 128

[112] T. Turletti, "GMSK in a nutshell," *Telemedia Networks and Systems Group*, 1996. 130

[113] N. Al-Dhahir and G. Saulnier, "A high-performance reduced-complexity GMSK demodulator," in *Signals, Systems and Computers, 1996. Conference Record of the Thirtieth Asilomar Conference on*, vol. 1, pp. 612–616, IEEE, 1996. 132

[114] J. Voogt, "Introduction to antenna types and their applications," Researchgate. 135

[115] C. Zhu, X. Li, and F. Li, "Nonlinear analysis of SC-FDMA spectrum for LTE up-link," *International Journal of Electronics Letters*, vol. 2, no. 1, pp. 30–36, 2014. 137

[116] A. Family, "3GPP long term evolution," 137

[117] S. Sesia, I. Toufik, and M. Baker, *LTE: the UMTS long term evolution*. Wiley Online Library, 2009. 138

[118] H. Holma, A. Toskala, *et al.*, *WCDMA for UMTS*, vol. 2006. Wiley Online Library, 2000. 139

[119] P. Dhande, "Antennas and its applications," *DRDO Science Spectrum*, pp. 66–78, 2009. 139