Master of Science in Electrical Engineering with Emphasis on
Radio Communication

PHYSICAL LAYER SECURITY USING PSEUDO-RANDOM SEQUENCE KEY GENERATION

NAGA VENKATA SAI TEJA GURRALA
SRIHARI AROLLA

Faculty of Engineering
Blekinge Institute of Technology
SE-371 79 Karlskrona Sweden

This thesis is submitted to the Faculty of Engineering at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering with Emphasis on Radio Communication. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:
Authors:

Naga Venkata Sai Teja Gurrala
e-Mail: Nagu16@student.bth.se

Srihari Arolla
e-Mail: Srar16@student.bth.se

Supervisor:
Hans-Jürgen Zepernick
Hans-jurgen.zepernick@bth.se
Department of Creative Technologies (DIKR)

University advisor:
Sven Johansson
sven.johansson@bth.se

Faculty of Engineering
Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

ABSTRACT

Nowadays, network security plays a major role in the field of wireless communications. Wired networks propagate electrical signals or pulses through cables. Whereas wireless signals propagate through the air. If wireless networks are left open and exposed to the outside world, there are high chances of being misused by others. The intruders take advantage of this, to intercept the wireless signals. This is the reason why an extra level of security is required for wireless networks.

The physical layer is one of the important layers of the Open System Interconnection (OSI) model which plays an important role in the network's physical connections like wireless transmission, cabling, connections etc. The physical layer supports the bit-level transmission between various devices by connecting to the physical medium for synchronized communication.

In this thesis, a method is studied for exchanging secret key [1] bits using a pseudo-random sequence generator based on Frequency Division Duplex (FDD) systems. The principle of this method is to generate a secret key in a manner that produces low correlation at the intruder. By uniquely relating the secret key bits to the channel in a private version of the universal codebook, a robust key exchange between the transmitter and the receiver is then performed.

Keywords: Physical Layer Security, Bit Error Rate, Key Error Rate, Pseudo Random Generator, Secret Key.

ACKNOWLEDGMENT

Firstly, we would like to thank God for giving us this wonderful opportunity to complete the project. Even though we are confronted with a lot of difficulties to complete the task, we still managed to finish the work without any complications.

Then it's time to show our gratitude towards our supervisor, Hans-Jürgen Zepernick, who helped us by guiding towards the right path by scheduling a week to week time plan. He had provided us with sufficient knowledge in our study and made sure that we understand the concepts without any hesitation.

Finally, a great thanks to our family who supported us and encouraged to complete the work. We are also thankful to our friends because they stood with us all time and corrected us in every step of our decision. Our friends, family and our supervisor provided us with a lot of benefits by giving ideas and comments on our project so that we can prove ourselves with good assessment in our project. They created us a passion to work hard and accept the challenge to complete the task.

TABLE OF CONTENTS

Abbreviations

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BS	Base Station
CDMA	Code Division Multiple Access
CP	Cyclic Prefix
CSI	Channel State Information
DoS	Denial of Service
FDD	Frequency Division Duplex
FTP	File Transfer Protocol
HTTP	Hypertext Transfer Protocol
I.I.D.	Independent and Identically Distributed
IP	Internet Protocol
ISI	Inter Symbol Interference
IoT	Internet Of Things
KER	Key Error Rate
LoS	Line of Sight
LTE	Long-Term Evolution
MIMO	Multi-Input Multi-Output
MITM	Man-in-the-Middle Attack
MS	Mobile Station
OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open System Interconnection
RF	Radio Frequency
SISO	Single-Input Single-Output
SMTP	Simple Mail Transfer Protocol
SNR	Signal-To-Noise Ratio
SQL	Structured Query Language
TCP/IP	Transfer Control Protocol/Internet Protocol
TDD	Time Division Duplex
UDP	User Datagram Protocol
WCDMA	Wideband Code Division Multiple Access
WWW	World Wide Web

1. INTRODUCTION

1.1 Motivation:

Wireless systems have been encountering an excessive development in the recent years because of its less expensive cost, more prominent scope and appealing adaptability. Versatility improvement is another remarkable element of wireless systems that give the clients "anywhere anytime" access, as well as the opportunity to utilize the network remotely.

Notwithstanding its promising highlights, security is yet a major problem in wireless systems since the wireless systems are defenseless to many types of attacks as the interference of information and eavesdropping, can be feasible for anybody with access to the systems [2]. Consequently, a security system is proposed by exchanging the secret key bits utilizing a pseudo-random sequence generator in view of a Time Division Duplex (TDD) system. The purpose of this technique is to generate a pseudo-random sequence in such a way that produces low correlation at the eavesdropper [3]. By particularly relating the secret key bits to the channel in a private version of the general codebook, a key exchange is done between the transmitter and the receiver.

1.2 Goals and Objectives:

The aim of this thesis is to study secret transmission of information over the physical layer using a pseudo-random key generator [4]. This objective is organized into three major phases:

- In the first phase, a random sequence generator is constructed to produce a series of sequences, out of which a random seed of a full-length secret key is generated and transmitted.
- In the second phase, a public codebook is produced by assuming all the obtained sequences that are generated in a single set.
- In the third phase, the receiver collects the secret key bits of each successful index detection and concatenates the assigned indexes after converting it to its private version.

1.3 Scope:

This thesis work mainly concentrates on the design and implementation of a pseudo-random key generator that produces a set of sequences called keys. These keys provide access to retrieve the confidential information by providing high secrecy transmission over the physical layer. The performance metrics like bit error rate and key error rate are evaluated for both Rayleigh and additive white Gaussian noise (AWGN) channels.

1.4 Report structure:

Chapter 2 provides the fundamentals of wireless communication, i.e., channels and fading, MIMO systems, Orthogonal Frequency Division Multiplexing (OFDM), TDD and Frequency Division Duplex (FDD) systems. Chapter 2 also provides some basic knowledge about secret key generation techniques. Chapter 3 gives a brief explanation regarding the methodology used in this research. Chapter 4 illustrates the results of the approach that are designed and implemented. Chapter 5 concludes our work and suggests the scope for future researchers.

2. FUNDAMENTALS OF WIRELESS COMMUNICATIONS

Wireless communication systems use electromagnetic signals which are produced by a transmitting section and received by a receiving section. The transmission of information over the channel does not include any kinds of wires or cables.

2.1 Fundamentals of MIMO-OFDM Systems

In general, modulation is a process of conveying the message by varying the frequency, phase or amplitude of the carrier signal. Multiplexing manages allocation/accommodation of users in each band (i.e. it deals with the allocation of available resources).

MIMO Systems

Systems using multiple transmission and reception techniques are usually known as MIMO systems [5]. This wireless system technology enhances both the range and capacity of wireless communications. For digital signal processing, a MIMO system provides a complex procedure to implement the schemes.

MIMO systems exploit multipath propagation by utilizing diverse transmission routing technique to the recipient. MIMO systems provide high capacity and low bit error rate (BER) when compared to single-input single-output (SISO) systems. But the difficulty faced by this system is the high fabrication cost and high energy consumption due to multiple radio frequency (RF) chains. The general block diagram of a MIMO system is shown in Figure 2.1.

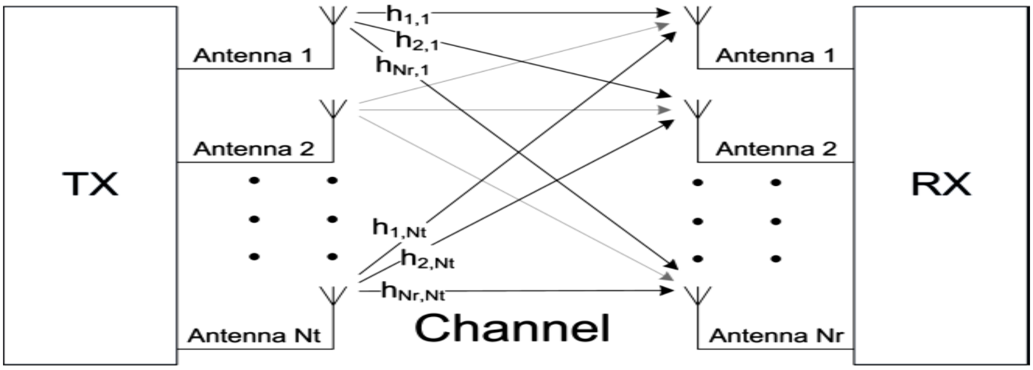

Figure 2.1 Block diagram of a general MIMO system.

Traditionally, multiple antennas (at one side of the wireless link) have been used to perform interference cancellation and to realize diversity and array gain through coherent signal combining. The use of multiple antennas on both sides of the link (MIMO, Fig. 2.1) offers an additional fundamental gain, i.e., spatial multiplexing gain, which results in increased spectral efficiency. A brief review of the gains available in a MIMO system is given in the following.

Spatial multiplexing provides high capacity, when compared to systems with a single antenna at one or both sides of the connection, with no additional usage of bandwidth. The gain is available if the channel exhibits high scattering.

Diversity gain is obtained by transmitting the information signal over different measurements in time, frequency, and space and by performing appropriate combining at the receiver. Spatial (i.e., antennas) diversity is obtained when compared with time or frequency, as it does not require any use of transmission time or bandwidth, individually. Space-time coding allows spatial diversity gain in the system with various transmission reception techniques without requiring channel information at the transmitter.

Array gain can be observed at both the transmitting and receiving sections. The channel provides the response for coherent signal combining, with an increase in the signal-to-noise ratio (SNR) to improve coverage. The capacity of a cellular system can be increased with the help of multiple antennae at both ends of transmission and reception by suppressing co-channel interference.

OFDM:

MIMO technology is utilized as a part of broadband networking systems that show frequency-selective fading and, therefore, inter-symbol interference (ISI). OFDM transforms the frequency-selective channel into an arrangement of parallel flat fading channels and is, afterward, a method for adapting to ISI. The fundamental rule that underlies OFDM is the inclusion of a guard band, called cyclic prefix (CP), which is a duplicate of the last section of the OFDM symbol and should be sufficiently long to delay the channel.

The utilization of the CP turns the activity of the channel on the transmitted signal from a direct convolution into a cyclic convolution with the goal that the exchange capacity of the information is diagonalized utilizing an inverse fast Fourier transform (IFFT) at the transmitter and fast Fourier transform (FFT) at the receiver. Subsequently, the overall frequency-selective channel is changed over into an arrangement of parallel level fading channels, which improves the equalization task. In any case, as the CP conveys repetitive data, it causes a loss in spectral efficiency, which is normally kept below 25 percent.

Binary Phase Shift Keying:

In binary phase shift keying (BPSK) modulation, the message is represented by the binary values 0 and 1 of the carrier signal. Because of the two different phases, it is also called as two-phase modulation scheme:

$$\text{Binary 1; } \theta = 0^0$$
$$\text{Binary 0; } \theta = 180^0$$

From the digital modulation techniques, a set of orthogonal basis functions are chosen using Gram Schmidt orthogonalization. Once the basic functions are picked, any vector in the signal space can be characterized as a linear combination of them. In BPSK, just a single sinusoid is used. Modulation is accomplished by changing the phase of the sinusoid relying upon the message bits. Specifically, the two diverse conditions of the carrier signal are represented as follows:

$$s_1(t) = A_c \cos(2\pi f_c t), \qquad 0 \le t \le T_b \, for \, binary \; 1 \qquad (1)$$

$$s_0(t) = A_c \cos(2\pi f_c t + \pi), \qquad 0 \le t \le T_b \, for \, binary \; 0 \qquad (2)$$

where A_C is the amplitude of the sinusoidal signal, f_c is the carrier frequency (Hz), t is the instantaneous time in seconds, and T_b is the bit period in seconds.

2.2 Security Vulnerabilities in Wireless Channels

In this section, we present the different security vulnerabilities in wireless networking systems [6]. The OSI model consists of the physical layer, the MAC layer, the network layer, the transport layer, and the application layer. The below Figure 2.2 shows the generic wireless OSI layered protocol architecture consisting of the different layers through which the data packets are transmitted from node A to node B through the wireless medium.

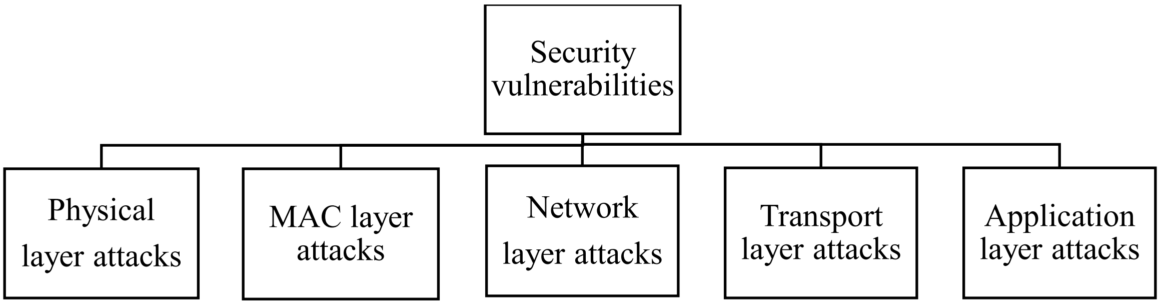

Figure 2.2 Security vulnerabilities.

Physical Layer Attacks: The physical layer is the bottommost layer in the OSI reference architecture through which the characteristics of the data transmission are defined. This layer is extremely defenseless to eavesdropping and other attacks. In wireless physical-layer attacks [7], the eavesdropping attack is nothing but the unauthorized usage of the wireless channel without providing any license. To provide security to the confidential information, this channel must be secured from third parties like an eavesdropper.

MAC Layer Attacks: The MAC layer allows multiple network nodes to access a shared medium using, e.g., CSMA/CA, OFDMA, CDMA, and so on. Each node has its own network interface controller, MAC address, which is used for authentication purpose.

One of the main attacks in MAC layer is the attempt to change its assigned MAC address with a malicious intention known as MAC spoofing. Although MAC address linked with network interface controller, it is possible to spoof a MAC address to hide the identity of attackers they can do any illegal activity. Also, MAC attackers steal the identity of other users to get confidential information which is known as an identity-theft attack.

In MAC layer, MITM attack and network injection are also included. MITM attack means man-in-the-middle, known as an attacker, which sniff the pair of network nodes without knowing attacker can control the network nodes. Network injection attack defines interrupting the operation of network devices such as routers, switches, etc. by injecting fake network reconfiguration commands. In this way, an attacker can initiate the commands, paralyzed the entire network, reprogramming of all networking devices.

Network Layer Attacks: In the network layer, Internet Protocol (IP) was designed to deliver packets from the source to the destination node using their IP addresses. Network layer attacks are mainly on IP spoofing, hijacking, and Smurf attack. IP spoofing is creating a fake IP address to hide the identity of the attackers for doing any illegal activities.

The packets are receiving from the fake source IP address will send a response back to a fake IP address. IP hijacking is another activity created by hackers for taking over another legitimate user's IP address. If they succeed, they can disconnect the user network and create a new connection for gaining the confidential information. There are other hijacking techniques such as prefix hijacking, route hijacking, and border gateway protocol hijacking.

A smurf attack is a Denial of Service (DoS) attack in the network layer. It can send a huge number of Internet Control Message Protocol (ICMP) packets to a user node or group of nodes' using IP address.

Transport Layer Attacks: In the transport layer, we have two kinds of attacks, one is TCP and the other is UDP. TCP is a connection-oriented transport protocol which is used for sending e-mails and for transferring the file from one network node to another. UDP is a connectionless transport protocol. It has a low price, it fails to deliver reliable data.
It is often used by delay sensitive applications like IP television, voice over IP and online games. Both TCP and UDP suffer from security vulnerabilities such as UDP and TCP flooding attacks and TCP sequence number prediction attacks.

TCP flooding is also known as ping flooding. It is a DoS attack in the transport layer where an attacker sends an enormous number of ping requests such as ICMP echo requests to a user node if it got a response for replies such as ICMP echo replies. This will flood both input and output buffers of the user node and even delay its connection to the network when the number of ping requests is high.

TCP sequence number prediction attack predicts the sequence index of TCP packets by the transmitting node and then fabricates. Specifically, a TCP sequence prediction attacker guesses the sequence index of victim transmitter, then fabricates packets and sends them to the victim receiver. UDP is prone to flooding attacks, which are imposed by sending the UDP packets instead of TCP flood attack. UDP flood attacker transmits many packets to a victim node, which will be forced to send numerous reply packets.

In this way, the victim node will be overwhelmed by the malicious UDP packets and becomes undetected by other legitimate nodes. Moreover, the UDP flooding attacker can hide from the legitimate nodes by using a spoofed IP address for generating malicious UDP packets.

Application Layer Attacks: The application layer has the HTTP attacks, FTP attacks, and SMTP attacks. HTTP attack is designed for exchanging hypertext across the World Wide Web (WWW), which as numerous security threats.

The HTTP attacks include the malware attack, structured query language (SQL) injection attack, and cross-site scripting attack. The malware refers to malicious software which is in the form of code, scripts, and active content programmed by attackers attempting to disrupt legitimate transmissions or to intercept confidential information.

The SQL injection is exploited to attack data-driven applications by inserting certain rogue SQL statements to gain unauthorized access to legitimate websites. The cross-site scripting attacks that occur in web applications and aim for bypassing some of the access control measures by injecting client-side scripts into web pages. FTP is used for large-file transfer from one network node to another, which also have certain security vulnerabilities.

The FTP bounce attacks and directory traversal attacks often occur. The FTP bounce attacks the PORT command to request access to ports through another victim node, acting as a middleman. The directory traversal attack attempts to gain unauthorized access to legitimate file systems by exploiting any potential security vulnerability during the validation of user-supplied input file names.

The SMTP is designed for transferring e-mails across the Internet, which, however, does not encrypt private information, such as the login username, the password, and the messages themselves transmitted between the SMTP servers and clients. E-mails are frequent carriers of viruses and worms. Thus, the SMTP attacks include the password sniffing, SMTP viruses, and worms as well as e-mail spoofing.

2.3 Wireless Channels and Fading

In wireless communications, fading is an attenuation of the signal with various variables such as time, geographical position, and radio frequency. The fading channel is a communication channel. In wireless systems, fading also occurs in multipath propagation, weather, a shadow from obstacles affecting the wave propagation also known as shadow fading.

Fading Channels

Fading channels refers to attenuate the signal due to obstacles, such as houses, building, trees, and mountain causing reflection, diffraction, scattering, and shadowing of the transmitted signals and multipath propagation. Multipath means transmitted signals are arriving in different angles, amplitudes and time interval. Fading is an amplitude fluctuation of the received signal caused by frequency selective or time variant of the multipath channel. Fading may be characterized by the Rayleigh probability distribution or Rician probability distribution. It will depend upon the strength of scattering during the transmission. The Rician probability distribution will be considered for the strength of scattering in the line of sight but in Rayleigh probability distribution contains pure scattering without a line of sight.

Frequency Non-Selective Channel

Multipath channels with time delay spread can affect the transmitted wireless signal. It has two categories such as frequency non-selective channel or frequency selective channel. Frequency non-selective channel is obtained if the channel bandwidth is larger than the bandwidth of the transmitted signal. It is also called flat fading. In flat fading, the symbol period T_S is greater than the delay spread of channel τ without causing ISI.

Frequency Selective Channel

Frequency selective channel is obtained if the channel bandwidth is smaller than the bandwidth of the transmitted signal. It is also called wideband channel. In frequency selective fading, the symbol period T_S is smaller than the delay spread of channel τ therefore causing ISI.

Rayleigh fading

In Rayleigh fading, the magnitude of a signal has passed through a transmitted medium randomly, according to a Rayleigh distribution. It mostly effects of heavily built-up urban environments on radio signals. Rayleigh fading is applicable when there is no dominant propagation between the transmitter and receiver. If there is a dominant line of sight, Rician fading may be more applicable.

Additive white Gaussian noise

Additive white Gaussian noise (AWGN) is a model for thermal noise in communication channels. The noise is white, it means the power spectral density is flat, so the autocorrelation of the noise in the time domain is zero for any non-zero-time offset. The noise samples have a Gaussian distribution.

TYPES OF FADING

The received signal for the narrowband channel is found to exhibit three types of variations such as fast fading, slow fading, and range dependence.

Fast Fading occurs due to channel impulse response changes rapidly within the symbol duration. It means coherence time of the channel T_D is smaller than the symbol period of the transmitted signal T such as T_D<<T. In fast fading, the amplitude and phase change imposed by the channel varies considerably over the period of use. Fast fading is due to reflections and motion of the objects.

Slow Fading occurs when the coherence time of the channel T_D is larger than the symbol period of the transmitted signal T such as T_D>>T. In slow fading, amplitude and phase change imposed by the channel can be considered roughly constant over the period of use. Slow fading can be caused by events such as shadowing, large obstruction such as hill or large building.

Small-Scale Fading: The received pass-band signal without noise after transmitted unmodulated carrier signal cos($2\pi f_c$t) can be written as

$$x(t) = \sum_i \alpha_i(t) \cos\left(2\pi f_c\left(t - \tau_i(t)\right)\right) = R\left[\sum_i \alpha_i(t) e^{-j2\pi f_c \tau_i(t)} e^{j2\pi f_i t}\right] \qquad (2)$$

Here R represents the real part of the quantity within its brackets, $\alpha_i(t)$ is time–varying attenuation factor of the i^{th} propagation delay, $\tau_i(t)$ is the time-varying delay and fc is the carrier frequency. Assume that $\tau_i(t)$<<T where T is the symbol time. The equivalent baseband signal can then be expressed as

$$x(t) = \sum_i \alpha_i(t) e^{-j2\pi f_c \tau_i(t)} \qquad (3)$$

If the delays $\tau_i(t)$ change in a random manner, and when the number of prothe pagation path is large, the central limit theorem applies and h(t) can be modeled as a complex Gaussian process.

Rayleigh Distribution: When the components of h(t) are independent, the probability density function (PDF) of the amplitude r = h = γ has a Rayleigh pdf:

$$f(r) = \frac{r}{\sigma^2} e^{\frac{r^2}{2\sigma^2}} \tag{4}$$

where E $\{r^2\}$ =2σ^2 and r ≥ 0.

This represents a severe fading case because we do not consider having line-of-sight (LOS). The power is exponentially distributed. The phase is uniformly distributed and independent of the amplitude. This is the most used signal model in wireless communications.

Rician Distribution

In case the channel is complex Gaussian with a non-zero mean (there is LoS), the envelope r = | h | is Rician distributed. Here, we denote h = α$e^{j\varphi}$+ν$e^{j\theta}$ where α follows the Rayleigh distribution and ν >0 is a constant such that v^2 is the power of the LOS signal component. The angles φ and θ are assumed to be mutually independent and uniformly distributed on [-π, π). The Rician pdf can be written as

$$f(r) = \frac{r}{\sigma^2} e^{\frac{r^2+\gamma}{2\sigma^2}} I_0\left(\frac{\Gamma_\gamma}{\sigma^2}\right) \quad \text{and} \quad r \geq 0 \tag{5}$$

where Io (.) is the modified Bessel function of order zero and $2\sigma^2$= E $\{\lambda^2\}$. The Rice factor

$$\mathrm{K} = \frac{\gamma^2}{2\sigma^2} \tag{6}$$

is the relation between the power of the LOS component and the power of the Rayleigh component? When K → ∞, Rayleigh fading is obtained.

2.4 Fundamentals of FDD and TDD Systems

In telecommunications, duplexing is a two-way communication over a transmission channel. It has two types, one is half duplex and the other is full duplex. In half-duplex, communication takes place on the same channel but transmitting and receiving are not at the same time, this means, if transmitting takes place over the channel, the receiver must wait for its turn. Full duplex enables the transfer of information from both the sides at the same time, i.e., the communication can take place at transmitter and receiver at the same time.

Full duplex has two types, one is frequency division duplex (FDD) and the other is time division duplex (TDD).

Frequency Division Duplex:

In an FDD system, the communication between the transmitter and receiver takes place simultaneously through two different channels or frequency bands. In order to avoid interference, the transmitter and receiver are provided with guard bands which are placed between the frequency bands. It has good filtering or duplexers and shielding to ensure the transmitter and receiver not to become slow.

Time Division Duplex:

When compared to FDD, the TDD requires only a single channel for communication. A single frequency band is sufficient to carry the communication between the transmitter and receiver. The transmission and reception operations are carried by allotting alternating time slots in the band. The transmitted information can be a voice, video, or files but it must be in serial binary format. Each time slot may be a byte or multiple bytes.

LTE, Cable TV, mobile cellular systems and 4G systems are the examples of FDD systems. TDD uses wireless data transmissions such as WiMAX and Wi-Fi, mostly in digital cordless telephones.

Due to the assignment of the frequency spectrum, FDD is most preferred over TDD.

2.5 Random Sequence Generation Techniques

Pseudo Noise (PN) sequences play an important role in mobile communication systems [8]. A PN sequence is determined into ones and zeros with the help of specific mathematical rules. Well known PN sequences are:

1. M-Sequence.
2. Gold Sequence.
3. Kasami Sequence.

Code Division Multiple Access (CDMA) is utilized by the Second Generation (2G) and Third Generation (3G) telephones where every user that uses the network is provided one of a kind PN sequence which recognizes the users. The IS-95 CDMA framework utilizes m-sequences while WCDMA (3G) utilizes Gold sequences. Kasami sequence has been proposed for 4G systems.

a. M-Sequences

Maximal length sequences are simply called as m-sequence. They are generated with the help of linear shift registers which are defined by a polynomial of order m. An m-sequence is generated periodically, with a period of 2^{m-1}.

For CDMA systems, each user is allotted a PN sequence to allow multiple access in a band of frequencies. It is known that cross-correlation function between m-sequences can have a relatively large peak, so m-sequences are not suitable for CDMA system due to mutual interference. Although it is likely to select a small subset of m-sequences that have smaller cross-correlation peak values, the number of sequences is usually too small for CDMA applications.

Preferred m-sequences

Two m-sequences with a length L having periodic cross-correlation to obtain the possible values of {*-1, t(m), t(m)-2*} are called preferred pair m-sequences, where

$$t(m) = \begin{cases} 2^{\left(\frac{m+1}{2}\right)} & ,odd\ m \\ 2^{\left(\frac{m+2}{2}\right)} & ,even\ m \end{cases} \tag{7}$$

where m is the order of the sequence generator polynomial.

Here, to generate the preferred m-sequence, two different shift registers are used with the help of polynomials which describe these registers and are called preferred polynomial having the degree of m.

b. Gold Sequences

Gold sequences are preferred because a large number of sequences can be provided and the cross-correlation between the sequences are uniformly bounded [1]. A Gold sequence can be generated by logical addition (XOR) of two preferred m-sequences, so they have the same period of 2^{m-1} as for m-sequences.

Gold-Like Sequence has parameters very similar to those of Gold sequences. Number of sequences in a Gold-like sequence set is N = L + 1 = 2^m where L is the length of one sequence (number of bits).

c. Kasami Sequences

Kasami sequences exhibit good cross-correlation properties when compared to others. The set of Kasami sequences is divided into two types: the small set and the large set. The large set is simply the subset of the small set. The period for Kasami sequences is defined by $N = 2^n$-1, where n is a nonnegative even integer.

Small Set of Kasami Sequences for n Even

The small set of Kasami sequences is defined by the following formulas. Let u be a binary sequence of length N, and let w be the sequence obtained by decimating u by $2^{n/2} + 1$. Here, T denotes the left shift operator, m is the shift parameter for w, and ⊕ denotes addition modulo 2:

$$K_s(u,n,m) = \begin{cases} u & m = -1 \\ u \oplus T^m w & m = 0, \dots, 2^{n/2} - 2 \end{cases} \quad (8)$$

Note that the small set contains $2^{n/2}$ sequences.

Large Set of Kasami Sequences for mod (n, 4) = 2

The large set of Kasami sequences is defined as follows. Let v be the sequence formed by decimating the sequence u by $2^{n/2+1}$+ 1. The large set is defined by the following formulas, in which k and m are the shift parameters for the sequences v and w, respectively:

$$\boldsymbol{K_L(u,n,k,m)} = \begin{cases} u & k = -2;\ m = -1 \\ v & k = -1;\ m = -1 \\ \boldsymbol{u} \oplus T^k v & k = 0, \dots, 2^n - 2; m = -1 \\ \boldsymbol{u} \oplus T^m w & k = -2; m = 0, \dots, 2^{n/2} - 2 \\ \boldsymbol{v} \oplus T^m w & k = -1; m = 0, \dots, 2^{n/2} - 2 \\ \boldsymbol{u} \oplus T^k v \oplus T^m w & k = 0, \dots, 2^n - 2; m = 0, \dots, 2^{n/2} - 2 \end{cases} \quad (9)$$

However, the Kasami sequences form a larger set than the Gold sequences. The correlation functions for the sequences take on the following values:

$$\{-t(n), -s(n), -1, s(n) - 2\,, t(n) -\ 2\} \quad (10)$$

where $$t(n) =\ 1 + 2^{(n+2)/2}\ + 2, n\ even \quad (11)$$

$$s(n) = \frac{1}{2}(\, t(n) + 1) \qquad (12)$$

The peak value of the cross-correlation function $|R_c|_{Max}$= t(m), is higher for the small set of Kasami sequences and the same as for the set of Gold sequences.

2.6 Performance Metrics

Bit error rate is the measure of performance of a digital communication system, which characterizes the reliability of the radio system through bits "in to out". The concept of bit error rate (BER) is simply given by

BER = Errors/Total Number of Bits

It is the ratio of bit errors that are received in the data stream per unit time over the communication channel.

Key error rate: If a basic communication between Alice and Bob is considered, the information is shared between each node to generate a secret key. If there is any mismatch of single bit key, then it would disturb the communication. Therefore, BER solely becomes insufficient and the key error rate (KER) needs to be considered to measure the system performance.

3. System Model

3.1 Basic Principles and Adversary Model

Let us assume that there are M antennas that are separated with equidistant from each other having half of the carrier wavelength so that there is no effect of mutual coupling between the transmitting and receiving antennas.

Now the transmitter of the system is designed in such a way that the signal pattern is adjusted in definite multiple paths to improve the performance of the system through taking advantage of the strongest channel mode. By assuming the number of antennas as MT, the precoding codeword matrix is mapped as MT ×1 using the modulation symbol 's' at the transmitter. The received Rayleigh faded signal is denoted as y = Hx + n where H is an $M_R \times M_T$ channel matrix of independent identically distributed (i.i.d.) random variables and n is the additive white Gaussian noise vector of length M_R, distributed as i.i.d. random variables.

Basically, selecting the precise precoding codeword requires the downlink channel state information (CSI) being available at the transmitter. In FDD systems, providing the CSI through the receiver is in many instances viewed impractical because it truly degrades the spectrum efficiency.

Firstly, the non-coherent detection of BPSK for Rayleigh Fading channel is investigated and then followed by coherent detection. For both scenarios, we have assumed flat fading. The AWGN is also added to the signal. The received signal y can be represented as

$$y = h\,x + n \qquad (14)$$

where 'n' is the noise contributed by AWGN which is Gaussian distributed with zero mean and unit variance and 'h' is the Rayleigh fading response with zero mean and unit variance. For a simple AWGN channel without Rayleigh Fading the received signal is represented as y= x + n.

3.2 Proposed Procedure

In this section, a brief description of our proposed procedure is explained. The procedure includes three major phases.

Phase 1

In this phase, we considered a pseudo-random sequence generator that generates a series of random bits called codewords. We used Kasami generator as the sequence generator in this project.

This Kasami generator produces a set of sequences with a fixed length. These sequences are collectively placed together in a set called codebook. This codebook contains all the sequences that are produced by the generator. This codebook is privately accessed between the BS and MS without providing any access to the eavesdropper (Eve).

Now, each codeword in the codebook is considered as a key, that must be transmitted to the receiver section, i.e., a mobile station. Out of the remaining keys from the codeword, a random key is chosen by the generator and transmitted over the channel.

Phase 2

In the second phase, fading is introduced into the channel to consider a real-time scenario. There are two types of channels that are considered. They are Rayleigh channel and AWGN. Now, the secret key is transmitted over the channel individually and fading occurs in it.

Phase 3

In this phase, the received signal is mixed up with noise and fading has to be retrieved in order to obtain the original key. To ensure this process, the system uses correlation technique in which the received key is compared with the transmitted one so that the errors in the key are found. After the removal of noise in the signal, graphs are plotted to analyze the BER and KER to the signal-to-noise ratio (SNR) for both the Rayleigh and AWGN channels. The block diagram of the system model is drawn in Figure 3.1.

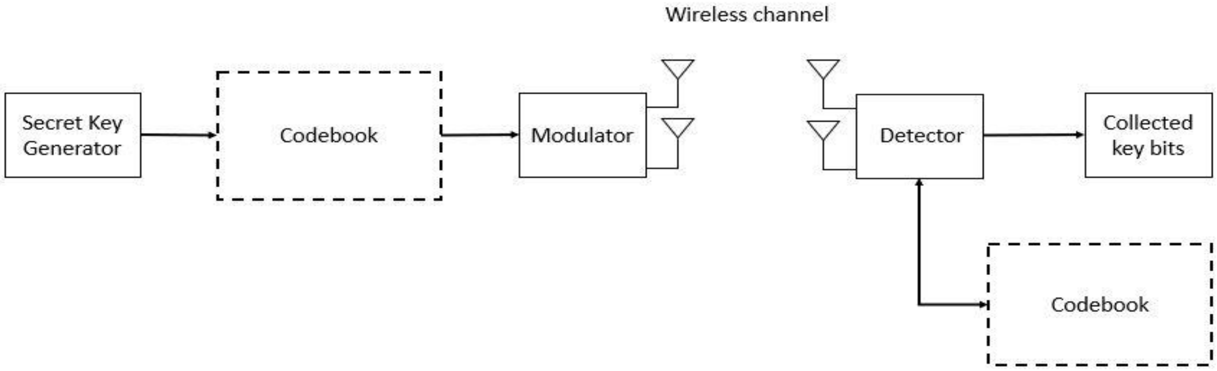

Figure 3.1 Block diagram of the considered model.

The generator produces a set of random sequences which are collectively stored in a codebook. Using BPSK modulation, the signal is transmitted through a wireless channel such as Rayleigh fading and AWGN. At the receiver, using CSI, the signal is filtered from the noise and the secret key is retrieved from the original signal.

4 NUMERICAL AND SIMULATION

We accessed the technique using MATLAB simulation with a practical setup to evaluate the performance of the system. With the help of the theoretical equations, we have plotted the of BER for the Rayleigh and AWGN channels in MATLAB environment for BPSK modulation.

The theoretical BER for BPSK modulation scheme over Rayleigh fading channel (with AWGN noise) is given by

$$p_b = \frac{1}{2}\left(1 - \sqrt{\frac{E_b/N_0}{1+E_b/N_0}}\right) \tag{15}$$

The theoretical BER for BPSK modulation scheme over an AWGN channel is given here for comparison

$$p_b = \frac{1}{2} erfc\left(\sqrt{E_b\ /\ N_0}\right) \tag{16}$$

where p_b is the probability of error and E_b/No is the ratio of received power to the interfered noise, which is simply the SNR.

Let us consider that the channel impulse response estimate at receiver is known and is perfect and accurate. The transmitted symbols x can be obtained from the received signal y by the process of equalization as given below:

$$\hat{y} = \frac{y}{h} = \frac{hx+n}{h} = x + z \tag{17}$$

Here, z is an AWGN except for the scaling factor 1/h. Now the detection of x can be performed in a manner like the detection in AWGN channels.

The binary input bits to the BPSK modulation system are detected as

$$\begin{aligned} r = real(\hat{y}) &= real(x+z) \\ \hat{d} = 1, &\quad if\, r > 0 \\ \hat{d} = 0, &\quad if\, r < 0 \end{aligned} \tag{18}$$

In the current evaluation, the Kasami generator produced 256 sequences out of which a random sequence is selected and transmitted by the BPSK transmitter. The graphs are plotted for both Rayleigh and AWGN channels and a discussion is given for each graph.

We assess the proposed technique performance simulation with the setup summarized in Table I.

Table I. Simulation Setup.

Channel Model	SCME
SISO System	single user
Modulation	BPSK
Codebook	DFT
Key Length	15 Bits

SCME is the Extended Spatial Channel model which is adapted for the testing and development of the Long-Term Evolution (LTE) standard. The channel bandwidth for SCME model is extended up to 100 MHz. When a single input is provided to the BPSK modulator, it produces a set of sequences. These sequences are stored in a codebook.

In this setup, Discrete Fourier Transform (DFT) is used, as it converts the finite sequence of samples with equal spacing into a complex valued function of frequency.

4.1 BER and KER Performance for AWGN channel

The program simulates binary phase shift keying at baseband through Monte Carlo method. The goal is to simulate BER over an AWGN channel. The effect of noise is represented in the baseband with Gaussian random samples added to each signal sample.

The SNR is varied to show the effect of SNR on BER. Finally, the theoretical BER, i.e. Q(sqrt(2*SNR)) is also plotted to show the correctness of simulation.

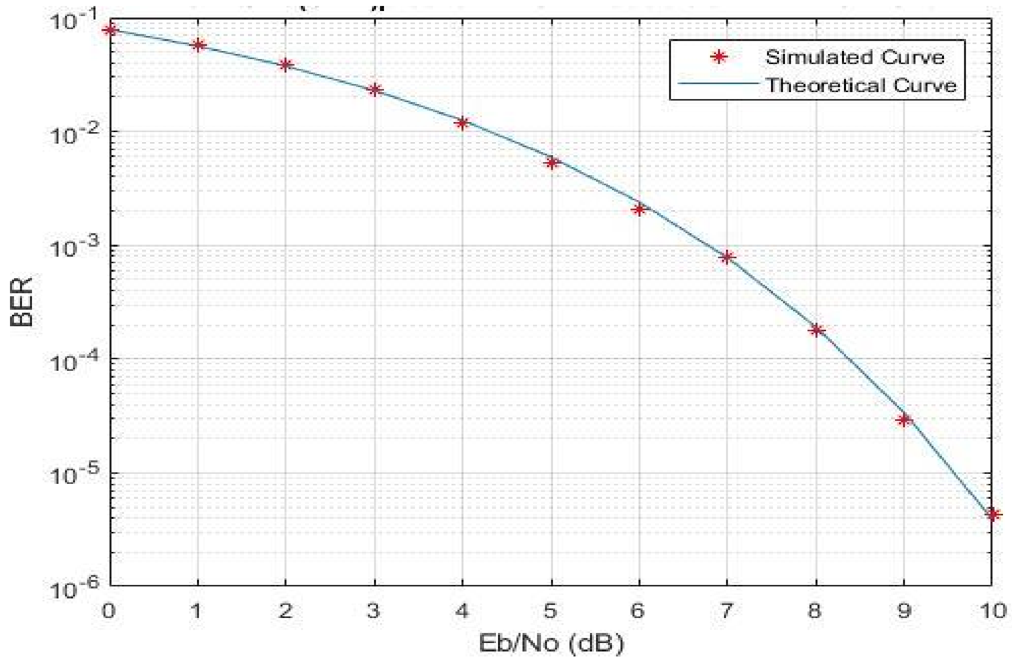

Figure 4.1.1: BER versus SNR plot for BPSK in AWGN.

From the above graph, it is observed that the theoretical and practical curves fit well.

The key error rate is the number of errors that are present in the key.

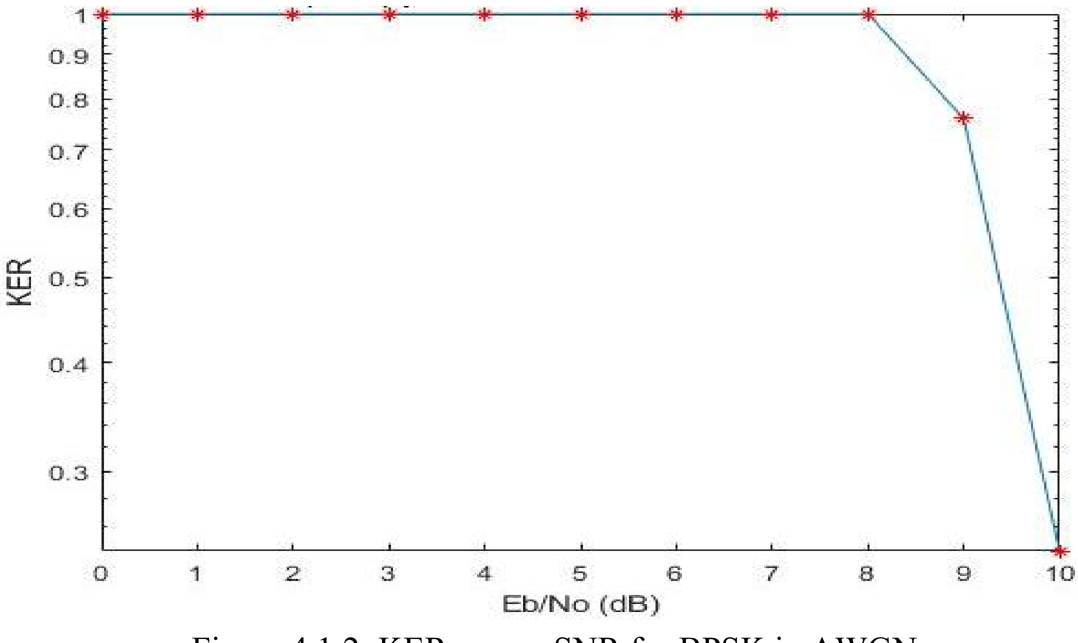

Figure 4.1.2: KER versus SNR for BPSK in AWGN.

4.2 BER and KER Performance for Rayleigh channel

The BER versus SNR is given in Figure 4.1.3.

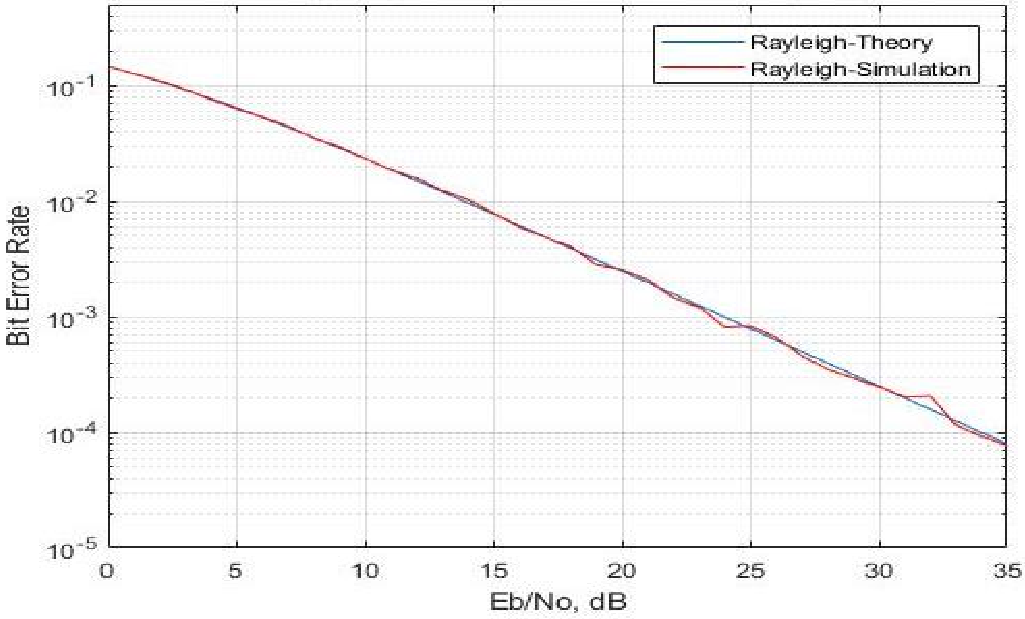

Figure 4.1.3: BER versus SNR plot in Rayleigh channel for BPSK.

From the above graph, it is observed that till E_b / N_0=15 dB, the simulated curve and the practical curve is the same and there are only small variations detected in the simulated curve. The KER versus SNR for Rayleigh fading is given in Figure 4.1.4.

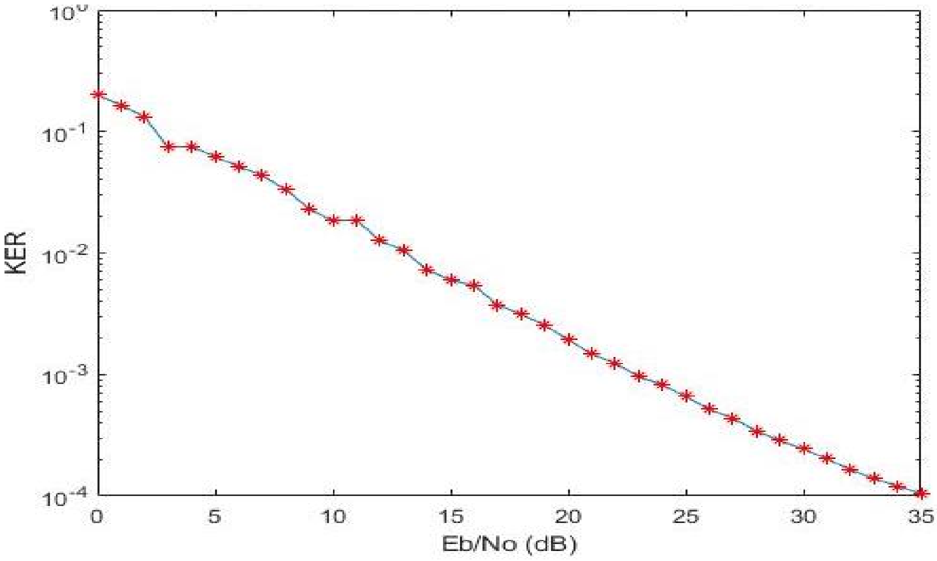

Figure 4.1.4: KER versus SNR in Rayleigh fading for BPSK.

BPSK modulation in AWGN provides good performance compared to that of Rayleigh channel. Also, the performance of BPSK degrades when the channel is subjected to fading with increasing value of Doppler shift (Hz). In other words, it performs poorly as the speed of mobile terminal is increased.

Table II. Set of Kasami sequences in the codebook for different sizes 4,8,16.

Generator input	Size of codebook	Number of sequences	Length of the sequences
Kasami (4)	[4,15]	4	15
Kasami (8)	[16,255]	16	255
Kasami (16)	[256,65535]	256	65535

From Table II, we can observe that, when different sizes of the codebook are given as the input to the generator, random kasami sequences are generated depending on the given input. For example, if we consider Kasami (4), it is generating 4 random sequences with a fixed length of 15.

5. CONCLUSION AND FUTURE WORK

This project mainly focused on an approach to increase the secrecy of data transmission over the physical layer. The concept of exchanging the key between the transmitter and receiver is completely based on a private random precoding. In this thesis, a comparison of key error rate is done for both AWGN and Rayleigh fading channels with different codebook sizes. Furthermore, from the evaluation, it is observed that the proposed model is providing a good key error rate for both channels.

The future work for this thesis includes a MIMO transmission and reception by involving multiple antennas. By this concept in MIMO systems, the information will be transmitted more securely without any interception of an eavesdropper. In the future, to implement the idea of secure transmission, the length of the key must be increased. The bigger the length of the key, the more complicated to intercept.

REFERENCES

[1] S. M. Bellovin and M. Merritt, “Encrypted key exchange: password-based protocols secure against dictionary attacks,” in *Proceedings 1992 IEEE Computer Society Symposium on Research in Security and Privacy*, 1992, pp. 72–84.

[2] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, “Improving Wireless Physical Layer Security via Cooperating Relays,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.

[3] Y. Zou, J. Zhu, X. Wang, and V. C. M. Leung, “Improving physical-layer security in wireless communications using diversity techniques,” *IEEE Netw.*, vol. 29, no. 1, pp. 42–48, Jan. 2015.

[4] “Secret key exchange under physical layer security using MIMO private random precoding in FDD systems - IEEE Conference Publication.” [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7511622/. [Accessed: 17-Nov-2017].

[5] C. Ehrenborg and M. Gustafsson, “Fundamental Bounds on MIMO Antennas,” *IEEE Antennas Wirel. Propag. Lett.*, vol. 17, no. 1, pp. 21–24, Jan. 2018.

[6] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, “A Survey on Wireless Security: Technical Challenges, Recent Advances, and Future Trends,” *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[7] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, “Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey,” *IEEE Commun. Surv. Tutor.*, vol. 16, no. 3, pp. 1550–1573, Third 2014.

[8] F. Sinnesbichler, A. Ebberg, A. Felder, and R. Weigel, “Generation of high-speed pseudo-random sequences using multiplex-techniques,” in *1996 IEEE MTT-S International Microwave Symposium Digest*, 1996, vol. 3, pp. 1351–1354 vol.3.

[9] A. A. Giordano and A. H. Levesque, “Digital Communications BER Performance in AWGN (BPSK in Fading),” in *Modeling of Digital Communication Systems Using SIMULINK*, Wiley Telecom, 2015, p. 416-.