Spring 2017

Intrusion Detection for Smart Grid Communication Systems

Brycent Chatfield

INTRUSION DETECTION FOR SMART GRID COMMUNICATION SYSTEMS

by

BRYCENT CHATFIELD

(Under the Direction of Rami J. Haddad)

ABSTRACT

Transformation of the traditional power grid into a smart grid hosts an array of vulnerabilities associated with communication networks. Furthermore, wireless mediums used throughout the smart grid promote an environment where Denial of Service (DoS) attacks are very effective. In wireless mediums, jamming and spoofing attack techniques diminish system operations thus affecting smart grid stability and posing an immediate threat to Confidentiality, Integrity, and Availability (CIA) of the smart g rid. Intrusion detection systems (IDS) serve as a primary defense in mitigating network vulnerabilities. In IDS, signatures created from historical data are compared to incoming network traffic to identify abnormalities. In this thesis, intrusion detection algorithms are proposed for attack detection in smart grid networks by means of physical, data link, network, and session layer analysis. Irregularities in these layers provide insight to whether the network is experiencing genuine or malicious activity.

INDEX WORDS: Smart grid, Intrusion detection, Moving Target Defense, Planar key, Received Signal Strength Indicator

INTRUSION DETECTION FOR SMART GRID COMMUNICATION SYSTEMS

by

BRYCENT CHATFIELD

B.S., Georgia Southern University, 2014

M.S., Georgia Southern University, 2017

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

INTRUSION DETECTION FOR SMART GRID COMMUNICATION SYSTEMS

by

BRYCENT CHATFIELD

Major Professor: Rami J. Haddad

Committee: Sungkyun Lim

Adel El Shahat

Electronic Version Approved:

May 2017

DEDICATION

This thesis is dedicated to my grandparents, Raymond and Rose M. Chatfield. Your unconditional encouragement of my endeavors has compelled me to go beyond any limitation. Thank you for all your support along the way.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere and heartfelt gratitude to my research advisor, Dr. Rami J. Haddad, for his valuable guidance, direction and encouragement throughout my research endeavor at Georgia Southern University. Without him, none of this would be possible. His success inspires me both academically and life in general to achieve heights greater than I have ever imagined. By introducing me to new ideas for smart grid security, my dedication to research within this area of study has increased significantly. He has also taught me the joys of conquering a challenge that others would deem impossible. I would like to express sincere recognition to my family and friends for their tremendous support of my ambitions during the whole of my collegiate career. A special thanks goes to my parents, Daryl and Shondrille Chatfield, for being the greatest inspiration anyone could ask for. I continue to strive to be the best in everything that I do in order to maintain the values you have instilled in me.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Smart Grid Solution

Traditional power grids play a critical role in the functioning of today's society. Without it, commonly enjoyed luxuries such as iphones, macbooks, television, music, and more would simply not exist. With that said, energy demands by consumers, industries and civilians alike, remains a daily challenge in terms of efficiency. In fact, according to statistics provided by [3] in their study on laboratory-based smart grid test beds demonstrated that energy production and consumption from 1950 to 2008 increased approximately two to three times, respectively. Consequently, energy demands of this magnitude have driven the current power grid towards its limitation. The development of the current power grid not being able to keep pace with industrial and social advancements is a direct cause.

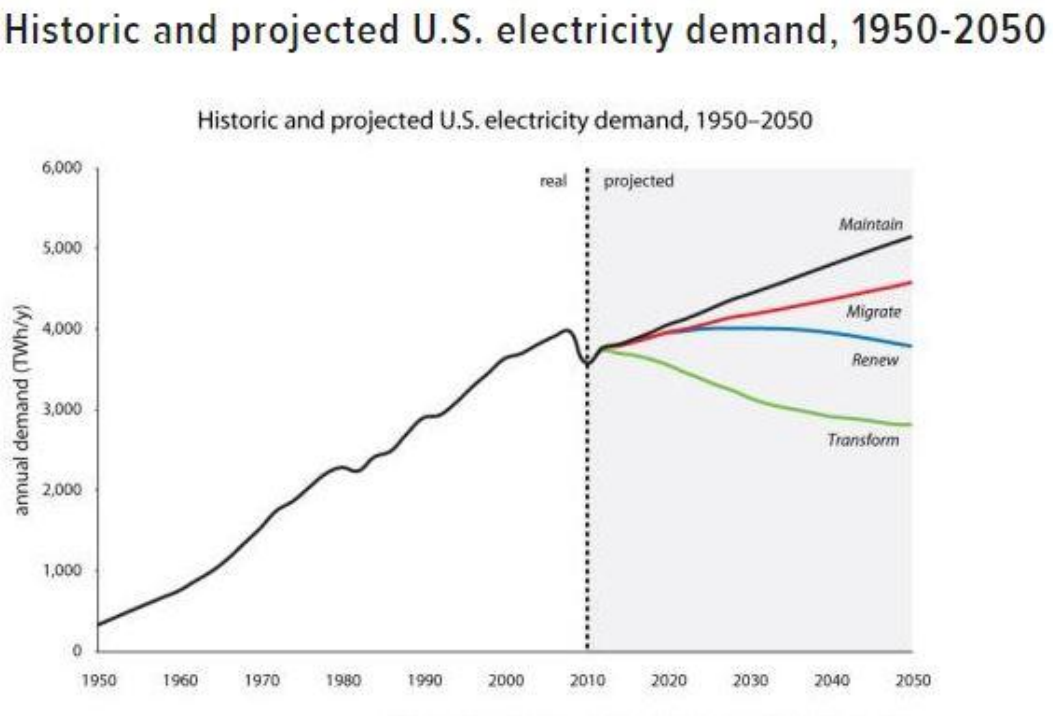


Figure 1.1: Historic and Projected U.S. Energy Demands [1]

The smart grid has been introduced as the next generation intelligent energy management infrastructure by integrating two-way communication technology into the traditional power grid. Added connectivity caused by the convergence of communication networks and energy systems enable consumers and energy suppliers to take advantage of convenience, reliability, and energy savings provided by real-time energy management [4].

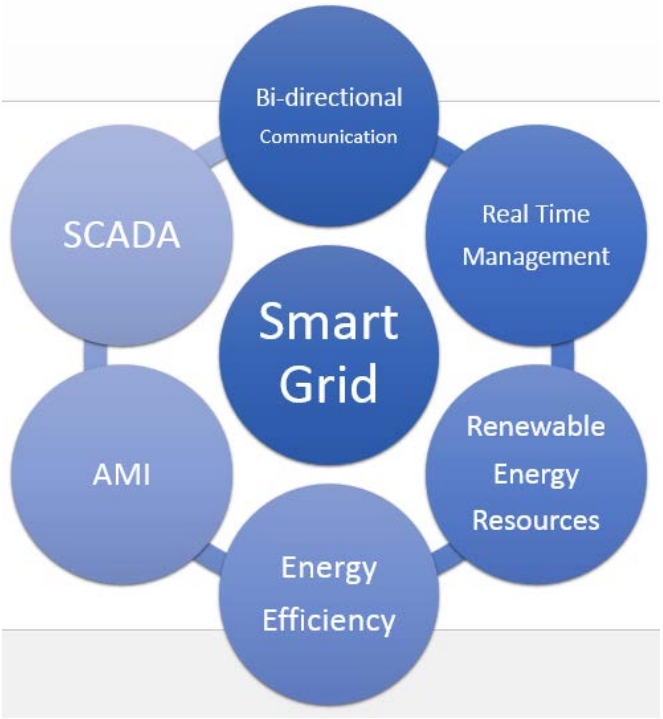


Figure 1.2: Smart Grid Overview

A smart grid incorporates millions of power equipment into a communication network. Furthermore, a smart grid features a dynamic and interactive infrastructure. An infrastructure this vast in magnitude is comparable to the size of the internet. Further, a smart grid provides better management capabilities. For example, advanced metering infrastructures are currently under research and employment in certain markets. With a smart grid enabled, advanced meters such as Itron's Centron Bridge Meter, become subject to better manageability. These improvements in turn enhance SCADA (Supervisory Control And Data Acquisition) operations for monitoring in the operations, transmission, and

distribution domains.

The smart grid's vastness requires network devices that are able to assist in efficient operations. A favored solution is advanced metering infrastructure (AMI). AMI is a communication infrastructure that enables smart meters and utilities to exchange information such as power consumption, price update, or outage awareness [5]. In recent studies over the years, such as in [6], several efforts have been put forth in AMI deployment.

1.2 Inherited Vulnerabilities

Conversely, because of the integrated communication network, the smart grid inherits all related network vulnerabilities. Convergence creates a new realm of security issues ranging from larger attack surface to an abundance of critical information available to an intruder. Cyber threats, such as eavesdropping, informational leakage, denial of service and malicious code injection provide a potential intruder a tremendous amount of leverage over a network [7]. In essence, there are numerous vulnerabilities of a compromised network including theft of information via account details. Unauthorized access of consumers account leads to leakage of social security numbers, home address, lifestyle in terms of knowing when a consumer is home versus away from home, stealing of power, intended attacks to disrupt system operation, destruction of infrastructure, etc. Furthermore, due to strict latency requirements and the critical nature of power systems, the smart grid is very susceptible to Denial of Service (DoS) attacks in which blackouts and cataclysms can occur. DoS attacks attempt to delay, block, or corrupt communications and can severely degrade network performance. Additionally, DoS attacks can happen at a variety of communication layers.

1.3 Intrusion Detection Systems

Data in the smart grid is of critical nature due to the transmission of sensitive information between consumer and utility company. A compromised network grants an intruder access to bank accounts, contact information, consumption patterns, personal files, etc. Communication networks will be required to consistently perform profiling, testing and comparison monitoring for network traffic. In response to attacks, networks must have self healing capabilities to ensure continuation of network operatons. Resilience operations, as emphasized, are very critical in smart grid security. Identification, authentication, and access controls are necesary for access to devices due to the magnitude of the smart grid. Ideally, this process ensures appropriate communcation and access to the device.

As a result, advancements in intrusion detection system (IDS) algorithms have been encouraged for deployment. In general, IDS algorithms are to aid in mitigating such attack vectors in order to preserve Confidentiality, Integrity, and Availability (CIA) in the smart grid. Though there are intrusion detection algorithms in place, such as antivirus, firewalls, and malware detection, they ultimately symbolize a static nature of defense [8].

Development of the smart grid has incited a wave of research efforts in attempts to secure the grid in an effective manner. At the home area network level, any effort to secure the grid should have low computational complexity to be viable. This is due to limited computational resources available in the advanced metering infrastructure (AMI). The ramification of any cyber attack on the grid would be devastating. Therefore, intrusion detection systems (IDS) have gained much attraction for smart grid security. IDS for smart grid systems have been heavily emphasized by the National Institute of Standards and Technology (NIST) guidelines for cybersecurity [9].

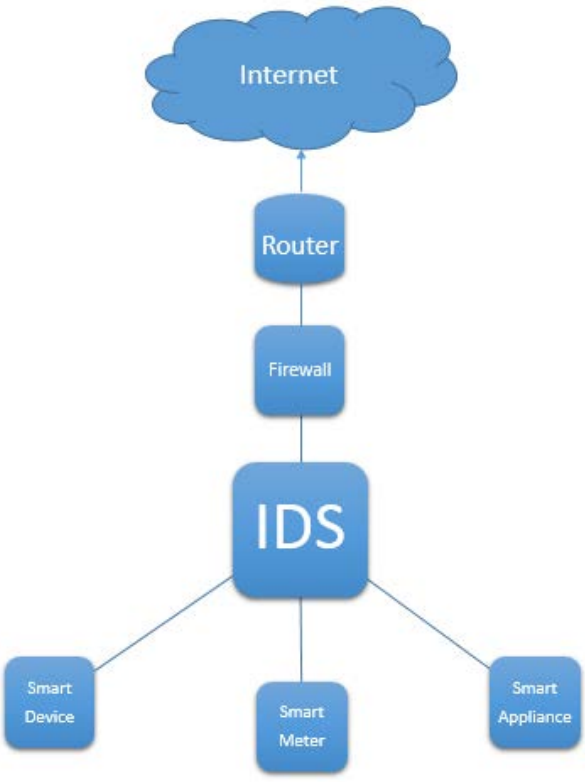


Figure 1.3: Intrusion Detectin System

1.4 Outline

The remaining chapters of this thesis is as follows. Chapter 2 provides an overview of smart grid architecture, area networks, and smart grid communications requirements. Additionally, wireless local area networking (WLAN) and long term evolution (LTE) protocols and their impact when integrated into the cyber-physical energy grid. Jamming, spoofing, and moving target attack vectors are described and their effectiveness on smart grid communications. Lastly, intrusion detection systems (IDS) along with attack detection and resiliency requirements are explored.

Chapter 3 introduces two intrusion detection algorithms for jamming attacks in WLAN and LTE networks. Theoretical analysis for both network types is exhibited along with experimental results. In the WLAN IDS, received signal strength indicator (RSSI) and

packet loss rate (PLR) are the two parameters utilized. The LTE IDS employs a similar concept where signal strength of synchronization signals are monitored to determine wether a smart device is able to establish a connection with the cellular network.

In chapter 4, a spoofing IDS algorithm is proposed to detect spoofing attacks in smart grid home area networks. Similar to jamming attacks, alterations in RSSI distinguishable from normal network traffic. In this approach, RSSI training data and sectoral based cosine similarity is used to create signatures that are unique to each smart device within a home area network. The advantage of the algorithm is creation of an attack surface that requires high precision to evade detection.

Chapter 5 presents Moving Target Defense Intrusion Detection System (MTDIDS) which is an algorithm for a new era of attacks known as moving target attacks (MTA). MTDIDS is comprised of an entropic nature that in turn generates a dynamic attack surface. By doing so, anomaly detection is possible. The primary advantage of anomaly detection is the ability to detect intrusions without the need for historical data of attack vectors otherwise known as zero day attacks. Therefore, securing the smart grid becomes more feasible to attacks that are yet to exist.

Chapter 6 concludes this thesis with an overview of all of the proposed algorithms. Implications of each is discussed. Moreover, a future works section is provided to expand on the possibilities of intrusion detection in smart grid communications.

CHAPTER 2

LITERATURE REVIEW

2.1 Smart Grid Architecture

Statistical results from a study provided by [10] has verified that there are over 2000 power distribution stations, 5600 distributed energy facilities, and more than 130 million customers within the U.S. The smart grid achieves manageability of an infrastructure of this magnitude by means of seven domains: bulk generation, transmission, distribution, customer, markets, service provider, and operations.

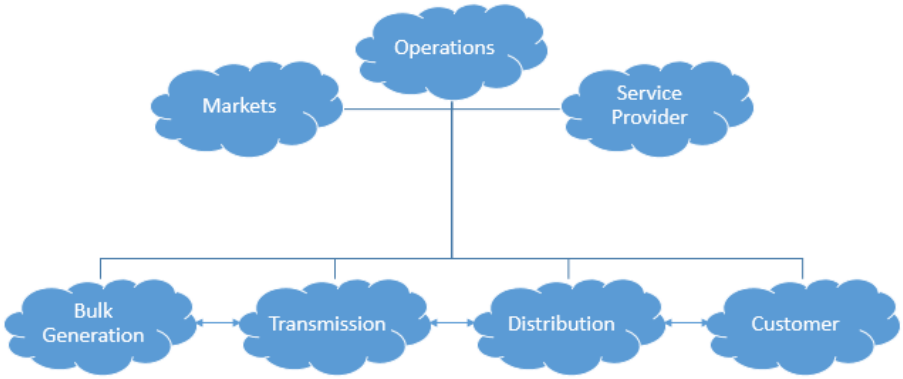


Figure 2.1: Seven Domains of Smart Grid

The first four domains, bulk generation, transmission, distribution, and customer, will feature two way power and information flow. Secondly, markets, service provider, and operations will feature information collection and power management. Due to the vast magnitude of a smart grid, communication networks are required to be highly distributed and hierarchal in nature.

2.1.1 Area Networks

For hierarchal purposes, a smart grid can be divided into three network tiers: Home Area Networks, Neighborhood Area Networks, and Wide Area Networks. In reference to a

survey in smart grid challenges and perspectives conducted in [11], we are able to state that home appliances of consumers are connected to Home Area Networks (HAN) in which they report their usage pattern of electricity in real-time to control and monitor power consumption. Neighborhood area networks (NAN) cover home area networks, substations, and distribution centers. Wide area networks cover power generation to transmission. Respectively, HANs are connected to Neighborhood Area Networks through Home Area Network gateways. NANs are connected to Wide Area Networks (WAN) through NAN gateways [12].

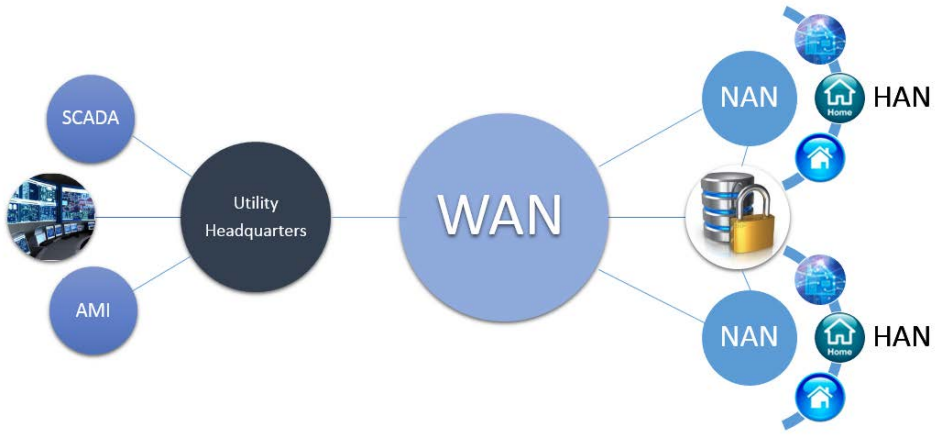


Figure 2.2: Smart Grid Area Networks

Reluctantly, communication networks employed today for cellular technologies, internet, and other commodities produces infrastructure mediums that are smart grid ready. Within the already existent infrastructure, fiber optic networks may be used to achieve high speed data and bulk and information transmission. Moreover, a smart grid features both wired and wireless technologies, ranging from fiber optics as mentioned to wireless means such as Wi-Fi, Zigbee, Ad-hoc, and 4G/LTE. An important asset associated with an already existent communication infrastructure is the reduction in costs for implementation of a smart grid.

2.2 Smart Grid Communications Requirement

Drastic differences in latency requirements of a smart grid, in comparison to the internet, demonstrates the time critical nature of the smart grid. Performance wise, internet focuses on high throughput and fairness amongst users whereas power communication networks focus to ensure reliable, secure, real-time message delivery. For example, time critical messages in substations are necessary for protection against faults that can lead to large scale blackouts. In general, the internet features many traffic flow protocols with World Wide Web (WWW) standing as the major protocol. In power networks, traffic flow is periodic thus allowing for consistent monitoring via SCADA systems.

In addition to traffic, the internet traffic delay requirement is typically 100-150 milliseconds, a necessity for internet services today [13]. Smart grids possess a wide range of delay requirements. Delay requirements, similar to substation messages as previously mentioned, can range from milliseconds to minutes. In short, smart grid traffic differs entirely from internet traffic.

The main objective of internet applications is to provide peer to peer communication between devices. Dynamically speaking, devices communicating with one another could be located within the same vicinity or reside in two completely different geographical locations. In smart grid communications, there are two types of communication protocols: top-down (center to device) and bottom-up (device to center).

2.3 Wireless LAN & Address Space

The smart grid features an array of communication infrastructures with wireless technologies serving as the predominant medium. For this reason, IEEE 802.11 protocols play a major role in the cyber-physical system. IEEE 802.11 is a set of media access control (MAC) and physical layer requirements to inact a wireless local area networking environment. In

general, an access point serves as the interface to convert data from a wired to wireless medium.

802.11 b,g and n are most commonly used since they represent the industrial, scientific, and medial radio bands (ISM). ISM operates in the 2.4 to 2.5 GHz spectrum and consists of 14 channels spaced 5 Mhz apart. Of the 14, channels 1, 6, and 11 are the most prevalent since they do not overlap with one another.

Standard internet communications operate under Internet Protocol version 4 (IPv4). Overall, IPv4 hosts a maximum of 2^{32} addresses available for network devices. Regrettably, the number of devices the smart grid contains in addition to current internet operations makes IPv4 not feasible for allocating addresses in an efficient manner. To correct this issue, Internet Protocol version 6 has come about permitting an astounding 2^{128} possible addresses. The total number of addresses in the IPv6 address space equates to nearly 5×10^{28} addresses for every one of the 6.8 billion people in the world [14, 15].

2.4 Long Term Evolution

Long Term Evolution is one of the most advanced wireless networks deployed today and is diligently paving its way as the primary cellular standard. Additionally, LTE's ease of access, ubiquity, high data rates, flexibility, and mobility signify major application in the realm of smart grids.

The smart grid concept presents itself as a viable solution to the currently limited power grid in a revolutionary manner. Enhancements are expected to be noticed in every domain in addition to SCADA operations. Incorporating a two-way communication infrastructure grants utility companies an upper hand in combating the issues of efficiency and manageability.

The advancement of the smart grid concept has brought attention to LTE networks as a communication infrastructure of choice. LTE's robust nature and ease of implementation

in devices with limited computational abilities is ideal for control of smart meters and other network components within the home, neighborhood, and wide area networks that are essential for operation. The usage of LTE in smart grid communications is extensively covered in the literature [16, 17, 18].

Smart metering utilized within the smart grid is related to the measurement of the power consumption of a building or apartment, and provides information about the Quality Electronic Report as well as feedback for the users of the power grid [17]. Furthermore, a utility company would be able to access instantaneous power readings of a given meter without having to dispatch technicians. Another improvement to note would be faster response time to issues that may arise. Due to computational limits of smart meters, incorporation of LTE protocols is ideal to obtain twoway communication.

2.5 Confidentiality, Integrity, Availability

Cyber security measures are subject to follow the CIA triad. The CIA acronym constitutes Confidentiality, Integrity, and Availability respectively. Confidentiality in the smart grid is needed to make sure that access to information is restricted to only authorized personnel while preventing unauthorized access by malicious users. In smart grid systems where home appliances are connected to power grids for real time bi-directional data communication and electricity flow, privacy is one of the important issues for the customers. If the information falls in the wrong hands, a malicious user could keep track of the life style of the victim, what appliances they use, whether the consumer is home or away, etc. Integrity of information in the smart grid is needed to maintain and assure the accuracy and consistency of data/information. The information should not be modified in an unauthorized or undetected manner. This feature helps the smart grid to provide robust real-time monitoring systems. Availability in the smart grid implies that the information must be available to authorized parties when it is needed without any security compromise. Power systems

are expected to be available 100 percent of the time, thus data availability also involves preventing denial-of-service attacks leading to blackouts [11].

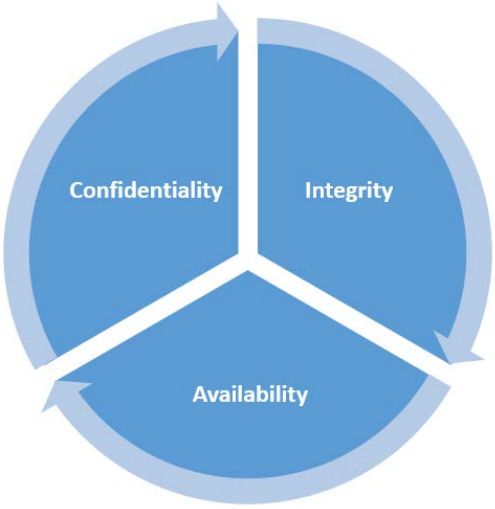


Figure 2.3: CIA Triad for Cyber Physical Smart Energy Grid

2.6 Attack Detection and Resiliency

A smart grid features a relatively open communication network over large geographical areas. Therefore, ensuring invulnerability of every node within the network becomes a near impossible task. Communication networks need to consistently perform profiling, testing, and comparison monitoring of network traffic. In response to attacks, networks must have self healing capabilities to ensure continuation of network operations. Resilience operations, as seen, are very critical in smart grid security.

Identification, authentication, and access controls are necessary for access to devices due to the magnitude of the smart grid. Ideally, this process ensures that appropriate personnel are able to access devices. If accessed by an unauthorized user, sensitive information could be leaked as well as control of the infrastructure. Thereon, cryptography measures are needed for each device whether symmetric or asymmetric.

2.7 Smart Grid Security Threats

Denial of Service attacks target availability. In short, they attempt to delay, block, or corrupt communications and can severely degrade network performance. Denial of Service attacks can also happen at a variety of communication layers. Furthermore, the smart grid does not need to be completely shut down to feel the effects of a Denial of Service attack. Simply causing a weak Denial of Service attack can cause catastrophic damage to the infrastructure in relation to time critical messages.

Physical layer attacks typically present themselves in the form of channel jamming. Channel jamming is the most efficient when aimed at physical layer links. The smart grid features a multitude of wireless technologies in which channel jamming becomes very effective.

MAC layer attacks aim to disrupt device to device communication. An attacker could change MAC parameters that could potentially cause performance issues in other devices that share the same communication channel. Secondly, spoofing is a major issue in MAC layer vulnerabilities. An attack can mask himself as another device to send false information.

The network and transport layers are responsible for multi-hop communications. A Denial of Service attack can degrade end to end communications via traffic flooding and worm propagation.

Application layer attacks focus predominately on transmission bandwidth. An attacker intends to exhaust a respective device by flooding computationally intensive requests into the system. Application layer attacks pose a hug threat to smart grid devices. Millions of devices within the smart grid feature limited bandwidth and computational abilities.

Integrity attacks also take place within the application layer. They are less brute and more sophisticated than Denial of Service attacks as they attempt to modify information in a stealthy fashion. Studies have shown that SCADA is very susceptible to integrity attacks that include false data injection.

Attacks targeting confidentiality intend to acquire unauthorized information from network resources. Eavesdropping on a communication channel allows an attacker to access customer information and electricity usage. Eavesdroppers also have the advantage of being undetectable due to inactivity within the network.

2.7.1 Jamming Attack Vectors

Jamming attack vectors are critical in nature and are presented in four primary forms: Constant, Random, Deceptive, and Reactive. Constant jammers emits a continous high powered noise signal to inject random bits into the channel. Random jammers achieve success by operating in intervals. During the off state, the jammer is considered to be "sleeping". In the on state, the random jammer operates as a constant jammer. In essence, constant and random jammers do not follow MAC protocols.

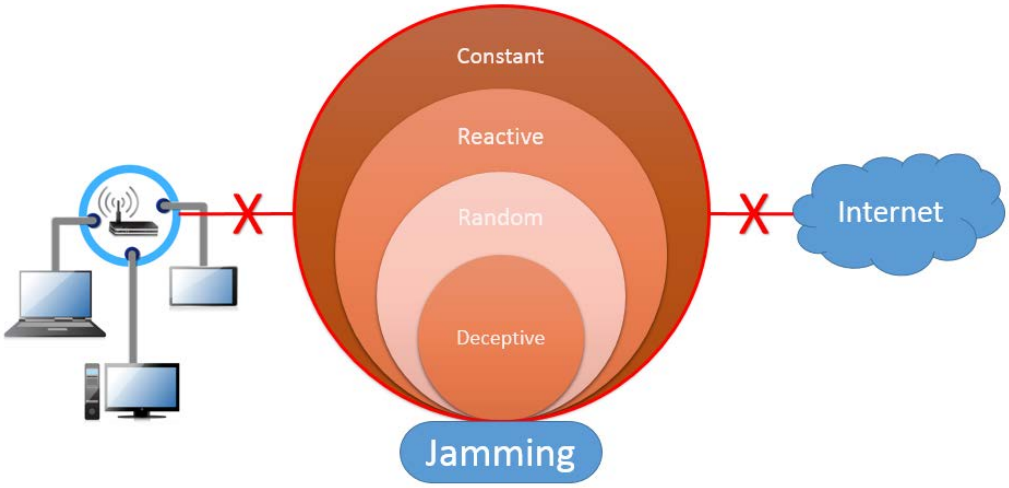


Figure 2.4: Network Influenced by Jamming Attack

In deceptive jamming, illegitimate packets are transmitted so that the channel appears busy. Since deceptive jamming is protocol aware, carrier sensing time for legitimate nodes is increased indefinitely. Reactive jammers initiate when data transmission is sensed on the channel. Once sensed, a jamming waveform is transmitted to cause corruption of data. Due

to protocal awareness, deceptive and reactive jammers can be considered as smart jammers.

2.7.2 Spoofing Attack Vectors

Spoofing attacks impersonates another device or user on a network to gain access to network hosts, inject malware, bypass access controls, and steal data. In the smart grid, spoofing attacks enables a malicious user to manipulate readings in order to manipulate pricing. Furthermore, a successful spoofer would have access to sensitve data such as social security numbers, account information, etc. Injecting malware ultimately reduces system performanc and can cause major delays. Most spoofing attacks fall into three primary categories: IP Spoofing, ARP Spoofing, and DNS Server Spoofing.

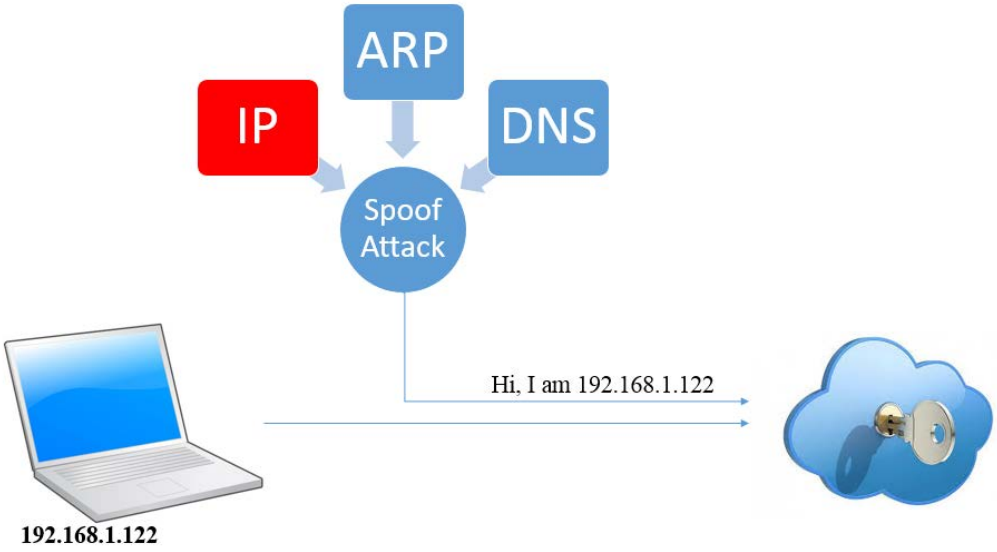


Figure 2.5: Network Influenced By Spoofing Attack

In IP spoofing, a malicious user emulates a desired IP address and transmits data as a legitmate user or device. IP spoofing is most commonly used to flood packets into the spoofed address. Likewise, a spoofed address can be utilized to transmit illegitimate packets to other recipients on the network. In addition, IP spoofing can be used to bypess IP based authentication protocols to gain "trust" on a network.

Address Resolution Protocol (ARP) spoofing enables a malicious user to send spoofed ARP messages across a network to link their respective MAC address with the IP address of a legitimate user. As a result, messages that are intended for the legitimate user are routed to the malicious user. ARP spoofing facilitates attacks such as session hijacking and man in the middle attacks.

Domain Name System (DNS) spoofing the DNS server is modified in order to reroute a specific domian name to a malicious IP address. DNS is typically used to resolve URLs and email addresses into their corresponding IP address. The malicious IP associated with the domain name reroutes a legitimate user to a server that hosts files infected with malware and thus is able to spread viruses.

2.7.3 Moving Target Attacks

Moving target attack (MTA) has gained way as the new era of attacks vectors where static attack detection approaches are rendered useless. MTA involves the modification of source and signature to bypass network security protocols. Therefore, anticipating a MTA becomes nearly impossible.

MTAs techniques are comprised of eight primary categories: Polymorphism, Metamorphism, Obfuscation, Self-encryption, Anti-VM, Anti-debugging, Encrypted and Targeted Exploits, and Behavior changes. Polymorphism uses multiple encryption keys to generate different instances of the same malware. Metamorphism functions in a similar manner by changing in-memory code with every execution. Obfuscation is the creation of code that is incomprehensible to human understanding which allows the evasion of manual code inspection. Self encryption changes malware signature to camouflauge malicous code and data. Anti-VM deactivates when in a virtual environment and commences malcious activity once released to real systems. Anti-debugging initiates malicous activity once runtime inspection and debugging tools are not detected. Encrypted and targeted exploits

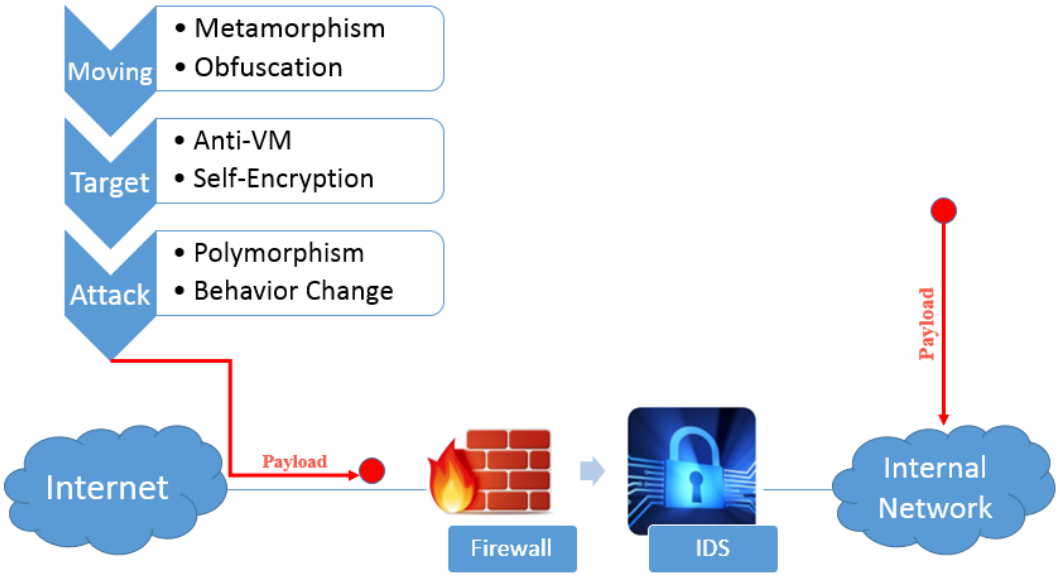


Figure 2.6: Network Under Influence of Moving Target Attack

changes host server, encryption keys, file names, etc. at every delivery. Behavior change triggers an attack during user interaction such as when a user navigates to a website and begins to scroll through the page.

2.8 Intrusion Detection in Smart Grid Systems

Networking devices in the smart grid possess limited computation capabilities thus being very susceptible to attack vectors. For this reason, utility companies need to understand the risks of deployment and the requirements for intrusion detection before they choose the monitoring architecture in which to invest [19]. In essence, attack detection algorithms are required to provide robust discovery of cyber attacks with minimal compuational complexity.

Intrusion detection systems (IDS) represent one of the first lines of defense in smart grid systems as they monitor the network and system for malicious activity and policy violations. When a compromisation is detected, the IDS alerts the system administrator or utility company of the event. The administrator then enacts measures to mitigate the

anomolous behavior.

The two primary types of IDS are network based (NIDS) and host based (HIDS). NIDS analyzes incoming network traffic whereaas HIDS monitors important system operation files. Detection approaches are either signature based, protocal analysis stateful-based, or anomaly based. Signature based recognizes bad patterns based on historical data. Anomaly approaches detect deviations from network traffic to determine abnormalities. Protocal analysis stateful-based is a variation of anomaly based dtection. In this approach, the IDS manufacturer provides behavior patterns collected in different networks and configures the system to detect these patterns [20, 21].

CHAPTER 3

SMART GRID INTRUSION DETECTION FOR JAMMING ATTACKS

3.1 Jamming Detection in WLAN

3.1.1 Overview

The most common type of DoS technique utilized by a malicious user is jamming. Jamming attacks are able to interfere with network operations instantaneously by emitting high powered radio signals of the same frequency. In the presence of a jamming attack, latency dramatically increases which could hinder a smart device's reception of important messages. An additional consequence is resource exhaustion because of a device's response to a compromised network channel. In the physical layer, jamming degrades channel availability. In the medium access control (MAC) layer, jamming further causes performance issues by invoking retransmission and other protocols in network devices. If transmission failure (e.g., ACK timeout) is reported by the MAC layer, the application will retransmit the same message until it succeeds [22]. Since communication in the smart grid behaves in a periodic manner, stability is of utmost importance.

3.1.2 Related Works

Several approaches have been proposed for jamming detection. In [23], Jamming Attack Detection based on Estimation (JADE) was introduced to achieve reliability for jamming detection in smart grid systems by using a gambling based model to determine probability of a jammed packet. The jamming probability is compared to a critical jamming probability determined for a network. An approach based on received signal strength indicator (RSSI) was implemented in [24] to detect jamming between smart phones and access points for Wi-Fi establishment. Their analysis insinuates that in the absence of jamming, high signal

strength usually corresponds to a low packet delivery rate. Low RSSI and high packet loss usually results from large propagation loss. In the presence of a jammer, both signal strength and packet loss rate are high. The algorithm employed similarly measures packet loss rate (PLR) and RSSI. This approach presents a great means for detecting jamming attacks but when applied to smart grids, time necessary to detect jamming attacks may not be sufficient.

The IDS presented in [25] detects jamming attacks by discovering the correlation between packet delivery ratio, signal strength variation, and pulse width of the received signal. The packet deliver ratio (PDR) of a node is obtained to be compared with a predefined threshold. If PDR is lower than the threshold, then the signal strength variation is compared with the signal strength variation in the normal network. PDR is then checked for consistency with signal strength variation. Next, the obtained pulse width is compared with predefined values for various jamming techniques. Though this approach is well defined, latency issues may arise due to limited computational resources of smart devices.

In [26], an IDS was proposed to identify the occurrence of malicious behavior and to notify the system operator about one specific type of electromagnetic interference in RFID systems. The system demonstrated jamming detection prior to having significant impact on communication between RFID reader and tag. The IDS consists of three modules: radio frequency signal processing (RFSP), attack detection, and event registration. RFSP module's first stage detects operating frequency of the signal and detects the received signal strength in the operating frequency respectively. The second stage detects existence of the preamble to be sent, extracts contents of the response from the tag and generates identifier, and performs cyclic redundancy check calculation based on response from the received tag. Parameters handled by the two stages are forwarded to the attack detection module. From there, the attack detection module performs comparisons of parameters provided by RFSP with predefined values in order to determine if an attack has occurred. The events are then

recorded to the event registration module. Similar to the previously mentioned IDS, this approach may suffer latency issues if applied to the smart grid.

3.1.3 Theoretical Analysis

In the smart grid, home area networks (HANs), neighborhood area networks (NANs), and wide area networks (WANs) hosts an advanced metering infrastructure (AMI) comprised of smart nodes with wireless capabilities. Based on the transmit power and distance from one node to another, the received signal strength indicator can be modeled as:

$$RSSI = P_T - 10\lambda log\left(\frac{d_1}{d_0}\right) + X_\sigma \tag{3.1}$$

where P_T is the transmit power, $PL\,(d_0)$ is the path loss for reference distance d_0, λ is the path loss exponent, d is the distances between the transmitter and receiver. Gaussian noise is present in all wireless mediums; therefore X_σ is the respective variable representing white Gaussian noise with zero mean and σ^2 variance. In the presence of a jamming attack, RSSI of incoming traffic is modeled as follows:

$$RSSI_T = (P_G + P_J) - 10\lambda log\left(\frac{d_G d_J}{d_0^2}\right) + X\,(\sigma_G, \sigma_J) \tag{3.2}$$

where $J\,(\mu, \sigma)$ is the jammer received signal with a mean power, μ_j, and variance, σ_j^2. Effectiveness of a jamming attack is directly proportional to distance from victim node. Due to spatial correlation, location of an attacker in reference to a victim node determines jammer effectiveness. As the signals interfere, associated Gaussian distribution shifts from expected mean and variance of genuine RSSI. Likewise, variance changes in relation to genuine traffic. Therefore, monitoring mean RSSI values in combination with variance proves to be adequate for jamming detection. Figure 3.1 illustrates the shift in RSSI Gaussian distribution as a result of jamming.

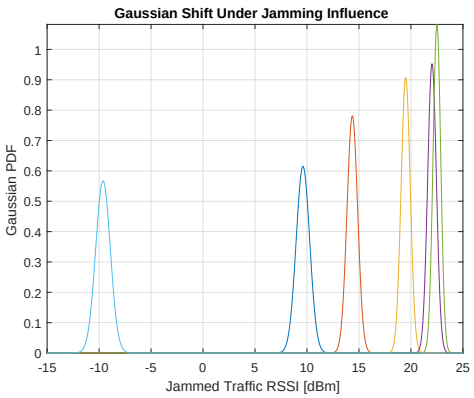


Figure 3.1: Gaussian shift under influence of jamming attack

The jamming causes the Gaussian distribution for genuine traffic with mean -10 dBm, to experience a drastic change in the mean and variance as illustrated in Figure 3.1. In addition to the observed variations in the RSSI Gaussian distribution model, a jamming attack usually increases the packet loss rate (PLR) significantly. PLR represents the ratio of the number of packets lost to the number of packets sent. Eq. 3.3 represents PLR for network traffic.

$$PLR = \frac{number\ of\ packets\ lost}{number\ of\ packets\ sent} \tag{3.3}$$

Whenever the jammer effectiveness increases, the PLR will exponentially decrease. Figure 3.2 displays the exponential increase in PLR as a result of jamming power gain. Undoubtedly, PLR serves as a good pre-indicator for detecting jamming attacks.

3.1.4 WLAN Jamming Detection Algorithm

Smart devices within the smart grid contain limited computational resources. Appropriately, IDS characteristics such as computational complexity and detection time determine the overall system viability. The proposed algorithm is comprised of one training phase and two sequential detection phases that analyze RSSI and PLR for attack detection. The 2-stage

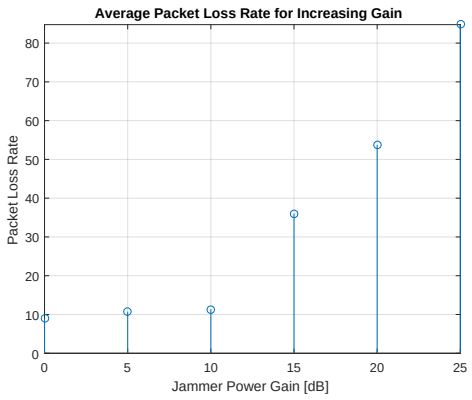


Figure 3.2: Packet loss rate under influence of jamming attack

detection phases are implemented sequentially to limit the computational complexity of the IDS algorithm. In addition, the algorithm performs analysis for every n RSSI samples attained by incoming network traffic where n is the number of samples to be analyzed at any instant of time as set by a utility company or vendor. The parameter n introduces a rolling window thus allowing real time attack detection. In case of a detection of an attack by the first detection phase the second detection phase is triggered and PLR samples are analyzed to confirm the attack.

3.1.5 Training Phase: Acquire Expected Mean and Variance

The training phase is only conducted once to establish baseline measurements. It's objective is to obtain RSSI and PLR measurements from nodes in the vicinity that will be communicating on regular basis. From the data, RSSI mean and variance are extracted to signify expected values of incoming traffic. The obtained expected RSSI statistics are then organized in pairs as follows:

$$E = (Expected\ Mean, Expected\ Variance) \tag{3.4}$$

where E represents a coordinate pair of expected RSSI mean and variance. The

coordinate pair serves as the basis for jamming detection in Detection Phase 1. In addition, the obtained normal PLR is recorded to estimate a PLR threshold PLR_{thresh} which will be used in the Detection Phase 2.

3.1.6 Detection Phase 1 - Euclidean Distance Detection

As stated, every n samples are analyzed in a given instant. The first detection phase measures the mean and variance for n samples to create a coordinate pair, O_i:

$$O_i = (Observed\ Mean, Observed\ Variance) \tag{3.5}$$

The i in O_i denotes the i^{th} instance of n observed samples acquired by the receiving node. The coordinate setup enables euclidean distance to be employed for comparison of data. As incoming data, sectored into O_i, is compared to expected RSSI mean and variance in E, the euclidean distance will be calculated as

$$d(O, E) = \sqrt{\sum (O_i - E)^2} \tag{3.6}$$

As previously mentioned, a jamming attack dramatically affects mean and variance components of incoming signal. Respectively, as mean and variance are augmented, an increase in distance, $d(O, E)$, is exhibited. An increase in $d(O, E)$ is directly proportional to a decline in network performance. Distance values provided will be compared against threshold d_{thresh} as follows:

$$d(O, E) < d_{thresh} \tag{3.7}$$

The network is considered stable if $d(O, E)$ remains below d_{thresh}. Ideally, $d(O, E)$ for a stable system is zero but due to multipath fading effects and noise, there will be variations. Accordingly, the threshold is set as the minimum distance allowable that will

enable satisfactory network performance. In the event that d_{thresh} is exceeded, Detection Phase 1 is presumed to be in a compromised state.

3.1.7 Detection Phase 2 - Packet Loss Rate Detection

To confirm the compromised state of the first detection phase, PLR detection is employed. Since jamming attacks dismantle channel availability, PLR is destined to surge. Similar to the first detection phase, PLR is observed for packets across n and compared to a normal operation PLR threshold as follows:

$$Observed\ PLR < PLR_{thresh} \tag{3.8}$$

PLR_{thresh} is the minimum allowed value that will satisfy network requirements. If $Observed\ PLR$ exceeds PLR_{thresh}, detection phase 1 is confirmed and the system is determined to be under influence of a jamming attack. Likewise, this phase is considered stable if $Observed PLR$ remains below PLR_{thresh}. On the other hand, detection phase 1 being in a compromised state while detection phase 2 signifies normal operation potentially indicates an anomaly in network traffic. An anomaly could mean that a different attack vector, such as false data injection, is being used to compromise the system.

3.1.8 WLAN IDS REsults

To validate the effectiveness of the proposed IDS, the system was tested in a wireless smart grid environment comprised of 3 Universal Software Radio Peripheral (USRP) LabVIEW devices. Two of the devices represented smart nodes communicating with each another were placed 30 ft apart. The third device, denoting the jamming device to cause DoS attacks, was placed 15 ft away from the receiver. Experimental parameters are listed in Table 3.1.

Table 3.1: Experimantal Environment Parameters

	Transmitter Node	Receiver Node	Jamming Node
Distance	0 ft	30 ft	5-25 ft
Frequency	2.4 GHz	2.4 GHz	2.4 GHz
Modulation	QPSK	QPSK	None
Packets	23000	23000	None

In this setup, our system analyzes 50 RSSI samples and associated packets at any instant of time. The objective of the transmitting node is to send 23,000 packets to the receiving node. During transmission, the jamming device is introduced, at several power gain values, in an attempt to degrade channel availability.

Communication between nodes is permitted following the generation of E. Figure 3.3 displays RSSI training data acquired by the smart node in the training phase. Expected mean and variance were obtained to determine the coordinates in E. Expected mean was calculated to be -5.89 dBm with variance 0.198. Thus E equals $(-5.89 dBm, 0.198)$.

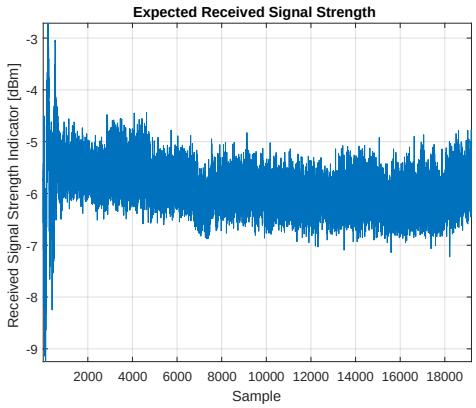


Figure 3.3: Expected RSSI for Network Traffic

As a result of jamming, the Euclidean Distance Detection phase transitions the system

from a stable to compromised state. Figure 3.4 displays the first detection phase for a jammer with 15 dB power gain. As demonstrated, when network traffic is genuine, the euclidean distance remains below the threshold of 5. Once the threshold is exceeded, the distance of the i^{th} sector of data increased instantaneously.

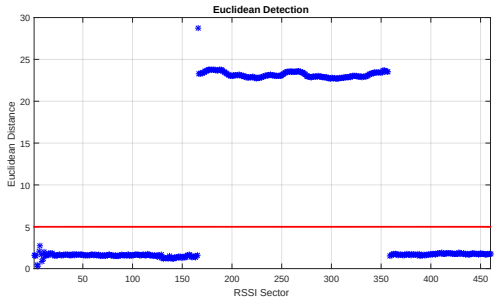


Figure 3.4: Euclidean Distance Detection

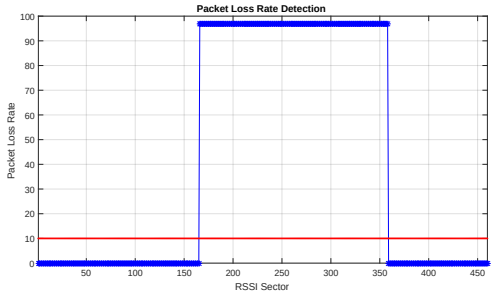


Figure 3.5: Packet Loss Rate Detection

The change of state results in detection phase 1 calling upon the PLR detection phase to confirm. Figure 3.5 highlights the PLR measured for each sector. Similar to the euclidean distance, PLR below 10% represents a stable system. In this test, a PLR of 97% was observed. PLR from detection phase 2 confirms phase 1 to conclude that the system is under a jamming attack. In both phases, sectors 166 through 358 were affected. These sectors indicate that packet numbers 8,300 through 17,900 were jammed thus proving a forensic advantage using the proposed IDS. Outcomes for the IDS tested under jamming conditions at different jammer power gain values are listed in Table 3.2.

Table 3.2: IDS Results for Jamming at Different Distances

Jammer Gain	Observed Mean & Variance	Euclidean Distance	Packet Loss Rate	System Decision
0 dB	(12.71 dBm, 0.162)	19.34	11.4%	Jammed
10 dB	(17.27 dBm, 0.134)	17.81	48.7%	Jammed
15 dB	(18.97 dBm, 0.0781)	24.85	97.1%	Jammed
20 dB	(19.17 dBm, 0.069)	25.01	99%	Jammed
25 dB	(20.18 dBm, 0.080)	26.54	100%	Jammed

3.1.9 Conclusion: Jamming Detection in Wireless LAN

In this section, a 2-stage IDS algorithm was proposed for detecting jamming attacks on wireless communication enabled smart grid devices. The transmission of data from one network node to another, both within range of a jammer, was considered. The two primary parameters used to facilitate the detection of jamming attacks were received signal strength indicator (RSSI) and packet loss rate (PLR). These two parameters were used in two separate detection phases to reduce the computational complexity in the wireless devices. Experimental results proved that the proposed system is capable of real-time detection of jamming attacks. The real-time detection advantage in combination with our algorithm's low computational complexity gives way to robust attack detection throughout the smart grid. Furthermore, anomaly detection is possible in the event that Euclidean Distance and PLR detection phases produce conflicting results. The IDS is deployable in smart grid networks that utilize smart devices of differing antenna architectures. Integration of our IDS into smart grid networks will improve overall stability.

3.2 Synchronization Signal Jamming Detection in LTE

3.2.1 Overview

The smart grid features a relatively open communication network as it covers over large geographical areas. Therefore, ensuring invulnerability of every device within the network becomes a very challenging task. When wireless network is used this challenge becomes even more complex as wireless signal can be overheard and jammed by the attackers/jammers. There are various forms of jamming attacks. Jamming attacks target radio signals by continuously or randomly emitting random signals into the channel with an aim of jamming the channel or deteriorate the signal quality at the receiver. MAC layer jamming prevents a node from determining channel availability.

3.2.2 Threat Model

Channel jamming could be executed to block Synchronization Signal which is known as Synchronization Signal Jamming (SSJ) [27]. Note that whenever a node wants to connect to another, a series of synchronization steps are required before transmission begins. A Primary Synchronization (PS) signal is transmitted from one node to the other for initial establishment of communications. Following the PS signal is a Secondary Synchronization (SS) signal which consecutively makes way for reception of a Master Information Block (MIB).

Due to the critical nature of the smart grid, with respect to time-critical messages, launching any kind of jamming attack proves catastrophic. Cost wise, a utility company could potentially find themselves victims to the tariff of dispatching technicians to resolve issues. Also, a network rendered to this type attack may leave many homes subject to blackout and more.

3.2.3 Theoretical Analysis

The underlying technology of LTE is Orthogonal Frequency Division Multiplexing (OFDM). In OFDM, the entire channel is divided into many narrow bandwidth sub-channels, which maintain high data rate transmission and at the same time increase the symbol duration to combat inter-symbol interference (ISI) [27]. To achieve OFDMA, the subcarriers are dynamically divided amongst mobile devices to enable access. In general, the input data stream is divided amongst the subcarriers. The subcarriers are spaced 15 KHz apart to achieve orthogonality as shown in Figure 4.

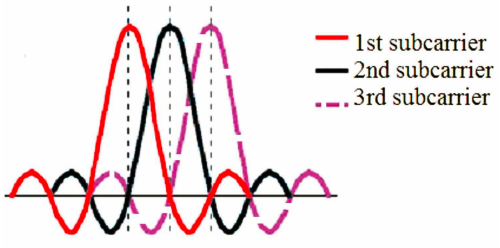


Figure 3.6: Orthogonality of Subcarriers in OFDM Systems

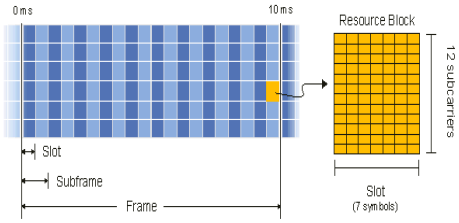


Figure 3.7: LTE Physical Layer Frame [2]

Each subcarrier in OFDM based LTE comprises of 7 symbols in which a scheme such as QPSK, 16 QAM or 64 QAM modulates the bits. After mapping and modulating the data onto the subcarriers, the Inverse Fast Fourier Transform (IFFT) is performed on data associated with each subcarrier. IFFT transforms the data into a time domain signal as

$$x(t) = \sum_{k=0}^{N-1} x_k e^{\frac{j2\pi kt}{NT_s}} \tag{3.9}$$

The IFFT signal, $x(t)$, is a function of the summation of the data symbols x_k , number of subcarriers N, and the symbol time T_s . To mitigate ISI, a cyclic prefix is appended in each subcarrier stream. Cyclic prefix is obtained by placing a copy of the end of a symbol at the beginning of the symbol thus increasing symbol length. The increase in symbol length counteracts multipath fading effects. Finally the parallel streams of data are converted to serial for transmission. In LTE, a Resource Block constitutes 12 subcarriers with 7 symbols each. Moreover, a resource block constitutes a 0.5 millisecond slot within the standard 10 ms time frame. Two resource blocks are associated with the 1 millisecond subframes. The number of resource blocks for communications is determined by bandwidth. Bandwidths utilized in LTE are 1.25 MHz with 6 resource blocks, 2.5 MHz with 12 resource blocks, 5 MHz with 25 resource blocks, 10 MHz with 50 resource blocks, 15 MHz with 75 resource blocks, and 20 MHz with 100 resources blocks [28].

In LTE, primary synchronization signal (PSS), secondary synchronization signal (SSS) and master information block (MIB) synchronization signal are designed to be detected by all types of UE. They are transmitted twice per 10 ms radio frame. Importantly, synchronization signals in LTE always occupy the central 62 subcarriers of the channel making the cell search procedure the same regardless of bandwidth [29]. The primary synchronization signal (PSS) enables a UE to access a cell ID as the initial step of communication establishment. Note that the PSS and the SSS are transmitted periodically on the last and second to last OFDM symbols of slot 0 of the first and the sixth sub-frames within a radio frame in frequency division duplex mode. PSS is constructed from a Zadoff-Chu sequence, which are complex valued sequences that have constant amplitude as

$$x_q(k) = e^{-j(\frac{\pi qk(k+1))}{N})} \tag{3.10}$$

where N is the length of the sequence and q is the Zadoff-Chu sequence root index. The secondary synchronization signal (SSS) provides timing information, FDD or TDD configuration, and cyclic prefix length. Lastly, the master information block (MIB) contains the downlink bandwidth of the cell, configuration, and system frame number.

In SSJ, the typical waveform deployed for jamming is noise. Though signals are subject to noise in an AWGN channel, the addition of a second source of noise proves to be effective in altering primary synchronization (PS) delivery. A received signal $x(t)$ can be represented by the OFDM signal generated at the transmitter as [30]:

$$x(t) = \sum_{k=0}^{N-1} x_k e^{j2\pi kFt} \tag{3.11}$$

where x_k is the symbol transmitted on the kth subcarrier, F is the kth subcarrier frequency given by:

$$F = k\left(\frac{B}{N}\right) \tag{3.12}$$

where B is the given bandwidth and N is the total number of subcarriers. Thus the jamming waveform is modeled as

$$n(t) = He^{j2\pi kF't} \tag{3.13}$$

with H denoting the jamming tone and F' constituting a given subcarrier frequency for a PS symbol during the fifth sub-frame.

3.2.4 LTE Jamming Detection Algorithm

Implementing an Intrusion Detection System (IDS) to detect whether a communication channel is jammed plays a major role in mitigating such attacks. With an IDS, the synchronization process is to be monitored to ensure completion. Additionally, periodic PS signals

could be transmitted for a desired number of packets to detect jamming during transmission. For the scope of this study, the primary focus is detection for the initial PS signal for each message needing to be transmitted.

The framework for the IDS [31] includes sensing a PS signal for establishment of a communication channel and comparing the received signal power to an expected power level. To do so, the IDS detects a PS signal when a channel is presumed idle. In this case, the PS signal is transmitted with an expected power which can be determined by using a statistical method or historic data as

$$|P_{received} - P_{expected}| < \alpha \tag{3.14}$$

Once the signal is received, the difference in signal strengths is compared to a predetermined threshold, α. The predetermined value is set by user to a quantity that best fits the respective application. If the threshold is exceeded, then the system is assumed to be compromised thus sending an alert to the utility company for corrective action. Furthermore, the instantaneous data stream a can be compared against the expected data value b in smart grid communication to detect an attack. One approach that could detect attacks is cosine similarity matching, that is,

$$Similarity = cos(\theta) = \frac{a \cdot b}{|a|\,|b|} \tag{3.15}$$

where $0 <= cos(\theta) <= 1$. Note that when observed value a is equal to expected value b, similarity score should be 0. If not, the similarity value would not be 0. This could help detect cyber-attacks that insert false data [12] or wrong value because of signal jamming.

3.2.5 LTE IDS Results

Performance evaluation is carried out using numerical results obtained from simulations. Firstly, the plotted power difference for PSS signal vs. the time as shown in Figure 3.8. To

see LTE synchronization signal jamming, the detection systems received PS signal strengths over a period of 200 milliseconds thus translating to 40 observed PS signals. Comparing power of the received signal with the expected to an applied threshold of $\alpha = 2.5$, a total of 13 PSS was presumed "jammed" as represented by Figure 3.8.

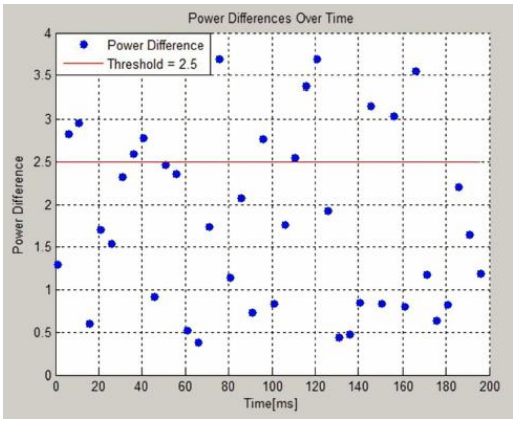


Figure 3.8: Detected Jammed PSS After Applied Threshold of 2.5

After calculating the difference in signal strengths and plotting them in relation to time, the results for every PS signal are obtained. Figure 3.8 infers, once the 2.5 threshold is applied at times t = 5ms, t = 35ms, t=100ms, t = 120ms, and t = 180ms, to name a few, the selected channel for transmission was deemed "jammed". For the mentioned times, utility companies are notified and corrective action, such as establishing an alternative or preferable link by switching to another frequency in place of the jammed frequency, are viable options.

Figure 3.9 demonstrates the expected OFDM signal in relation to the signal altered by jamming effect. Ideally, the difference between expected signal and observed signal was expected to be zero. Alternatively both signals should overlap with each other when there is jamming attack. Similarly Figure 3.9 illustrates the correlation between jamming signal and targeted PS signal. Ideally these two signal should have overlapped if there was no jamming attack. Difference indicates that there is jamming attack in the system. When an

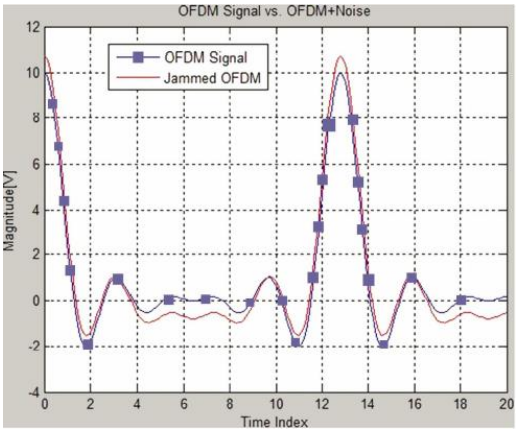


Figure 3.9: Expected OFDM Signal in Relation to Jammed OFDM Signal

attack is detected, the detector reports to the administer at monitoring center or takes the corrective action to avoid attack as quickly as possible.

3.2.6 Conclusion:Synchronization Signal Jamming

In the tiered WANs, NANs, and HANs infrastructure in cyber-physical smart energy grid system, wireless technologies are candidate solutions to establish communication between utility company and consumer. Connected systems bring networking opportunity to different entities along with several vulnerabilities. Since wireless mediums are very susceptible to several cyber-attacks including channel jamming. This section of the chapter has presented an intrusion detection system to detect/mitigate jamming attacks in LTE based smart grid communication system. Once a channel is declared jammed, system can send an alert signal and corrective actions could be taken by the administrators or systems dynamically. Performance is evaluated using numerical results obtained from experiments.

CHAPTER 4

SPOOFING DETECTION IN SMART GRID

4.1 Overview

In the smart grid, Home Area Networks (HAN) are comprised of home appliances (or smart devices) of consumers that report their load demand and electricity usage patterns in real-time, to control and monitor the real-time power consumption [32]. HANs host an array of wireless communications infrastructures ranging from Wi-Fi (802.11) to Zigbee (802.15.4). Consequently, the incorporation of a wireless infrastructure into the traditional power grid causes an inheritance of associated vulnerabilities.

The primary vulnerability in HANs is spoofing attack. In a spoofing attack, an adversary masquerades as one or more legitimate nodes, and by forging their identities, injects malicious traffic to affect normal operation of the network [33]. In relation to the smart grid, spoofing attacks will enable a malicious user to inject false power readings to manipulate pricing, enable domain access for pivoting through a network, cause Denial of Service (DoS) effecting network performance, etc.

4.2 Related Works

Recent related studies for spoofing detection in smart grids and wireless sensor networks are detailed in [33, 34, 35, 36, 37, 38, 39, 40]. In [33, 35, 36, 37], Discrete Haar Wavelet Transform (DHWT), summation of detailed coefficients (SDC) of received signal strength indicator (RSSI) streams, and ratio of out of bounds frames were used to detect spoofing attacks in wireless sensor networks and HANs. The study in [34] detected identity-based attacks by means of signal print verification. A given node broadcasts RSSI values to all nodes within the vicinity. From there, a decision is made by referencing the broadcast RSSI. This approach could generate many false positives due to multipath fading and does not

mention adaptability for different antenna types. In [38], the empirical path loss model is utilized for attack detection. Each node broadcasts packets to its neighbor and stores RSSI values for given distances. Then, the nodes evaluate an approximate path loss, fading, and standard deviation coefficients. Finally, the approximated coefficients are applied to the empirical model to determine a threshold. This approach heavily relies on a large number of devices to obtain high detection rate. Therefore, a significant drop in detection rate is seen when fewer devices are used.

4.3 Threat Model

HANs encompass an assortment of smart appliances and devices that are capable of transmitting power readings to a smart meter. In the case of HAN, a smart meter is presumed to be the target of spoofing attacks. Furthermore, smart meters serve as the HAN gateway for transmission of power readings to Neighborhood Area Network (NAN) gateways. Zigbee (802.15.4) and Wi-Fi (802.11) are the predominate communication infrastructures used in HANs. The scenario at hand is based on an 802.11 infrastructure.

We assume that an attacker has impersonated a smart device in an attempt to infiltrate and maneuver around the network. The attacker could possibly gain administrative privileges, manipulate power readings, or deliberately cause DoS attacks.

4.4 Theoretical Analysis

The proposed algorithm takes into account spatial correlation of RSSI [41]. The dynamic nature of the environment causes multipath fading. Accordingly, parameters such as distance have a predominate effect on RSSI. RSSI can be modeled as:

$$RSSI = P_o - 10\lambda log\left(\frac{d_1}{d_0}\right) + X_\sigma \tag{4.1}$$

Where P_o is the transmitted power by node i, λ is the path loss exponent, d_1 is the distance between transmitter and receiver with d_0 as the reference distance. X_σ is the Gaussian distribution $N(0, \sigma)$. In a spoofing attack, an attacker is presumed to have knowledge of the frequency and modulation scheme of the smart device of interest.

At the receiver end, RSSI will be the addition of the spoofed and genuine signal represented as follows:

$$RSSI_T = (P_S + P_G) - 10\lambda log\left(\frac{d_S d_G}{d_0^2}\right) + X\left(\sigma_S + \sigma_G\right) \tag{4.2}$$

where $RSSI_T$ denotes the total received signal strength. P_S and P_G are the addition of the spoofed signal and genuine signal. The distances d_S and d_G are the distances of the spoofing and genuine node. d_0 is the reference distance. $X(\sigma_S + \sigma_G)$ represents the Gaussian distribution affected by both signals.

The sectoring nature of the proposed algorithm takes RSSI streams and divides them into 2^n sectors with a mean and variance associated with each as shown in training phase 1. Accordingly, when two signals interfere at the receiver, a significant increase in mean and variance per sector is noticed. The increase in mean and variance correlate to a decrease in the cosine similarity.

4.5 Spoofing Detection Algorithm

The proposed spoofing detection algorithm features sector analysis of RSSI samples obtained by a smart meter from neighboring devices. Sectoring of RSSI data has shown to be effective in terms of accuracy. Additionally, deterring potential attackers increase due to the daunting task of having to mimic spatial correlation properties for each RSSI sector. This algorithm is highlighted in Algorithm 1.

The spoofing detection system is comprised of two training phases and an operational

Algorithm 1 Proposed Spoofing Detection Algorithm.

Precondition: Training Phase 1: For a smart meter, $\{k_1, k_2, ...k_{max}\}$ is RSSI training data from respective smart device i.

function SECTORING($k_1, k_2, ..., k_{max}$)
  for $k_x \leftarrow x = 1$ to max **do** ▷ Training data sets
    $k_j \leftarrow reshape(k_x, 2^n$ ▷ Converts k to sectors
    $j \leftarrow j + 1$ ▷ Increments j with k_x
  end for
  for $k_j \leftarrow j = 1$ to max **do**
    $k_{j\mu} \leftarrow mean(k_j)$ ▷ Mean for each sector
    $k_{jvar} \leftarrow var(k_j)$ ▷ Variance for each sector
    $j\mu \leftarrow j\mu + 1$
    $jvar \leftarrow jvar + 1$
  end for
  for $k_j\mu \leftarrow j\mu = 1$ to max **do**
    $G \leftarrow G + k_{j\mu}$ ▷ Add sectoral mean
  end for
  for $k_jvar \leftarrow jvar = 1$ to max **do**
    $V \leftarrow V + k_{jvar}$ ▷ Add sectoral variance
  end for
  $E_{ij} \leftarrow (G/max, V/max)$ ▷ μ,var for device i

Precondition: Training Phase 2: Compare average mean and variance to genuine traffic. Sectoral mean and variance extracted to create G_{ij}. Cosine similarity to compare E_{ij} to G_{ij}

  $G_T \leftarrow (E_{ij} * G_{ij})/(mag(E_{ij}) * mag(G_{ij}))$

Precondition: Detection Phase: In deployment network traffic sectored to generate O_{ij}. Cosine Similarity to compare O_{ij} to E_{ij}.

  $O_k \leftarrow (E_{ij} * O_{ij})/(mag(E_{ij}) * mag(O_{ij}))$
  $Detect \leftarrow |O_k - G_T|$
  if $Detect > \alpha$ **then**
    System ← Compromised
  else
    System ← Not compromised
  end if
end function

phase. If an attack is detected, the smart meter sends an alarm to the utility company or enacts predefined preventative measures. Training and operational algorithms are detailed as follows.

4.5.1 Training Phase 1: RSSI Training Data

As aforementioned, HANs consist of smart devices capable of transmitting power readings to the smart meter. Smart meters are also capable of communicating with one another. Thus, the initial training phase consists of obtaining RSSI mean and variance values for each node in proximity. As a result of sectoring RSSI streams, mean and variance values are stored for each sector, j, respectively.

For the i^{th} node, k RSSI training datasets are collected. Each dataset is sectored by a factor of 2^n samples with n denoting sensitivity measure. Therefore, the number of sectors, ρ, for each RSSI stream is depicted as:

$$\rho = \frac{m}{2^n} \tag{4.3}$$

where m indicates the total number of samples for the k^{th} RSSI training dataset. Note that, m must be factorable by 2^n. Accordingly, mean and variance values for the j^{th} sector of the k^{th} RSSI dataset for given node i can be represented as a coordinate pair, $(\mu_{ijk}, \sigma^2_{ijk})$.

After obtaining mean and variance pairs, an average is then obtained for sector pairs of identical sequence (i.e. the first sector pair of the first dataset will be averaged with the first sector pair of the second dataset). An example matrix for node i where each pair corresponds to a given sector is demonstrated as:

$$\begin{matrix} (\mu_{i11}, \sigma^2_{i11}) & (\mu_{i21}, \sigma^2_{i21}) & \left(\mu_{i31}, \sigma^2_{i31}\right) \\ (\mu_{i12}, \sigma^2_{i12}) & (\mu_{i22}, \sigma^2_{i22}) & \left(\mu_{i32}, \sigma^2_{i32}\right) \\ \left(\mu_{i13}, \sigma^2_{i13}\right) & \left(\mu_{i23}, \sigma^2_{i23}\right) & \left(\mu_{i33}, \sigma^2_{i33}\right) \end{matrix} \tag{4.4}$$

Thus, newly obtained averages for each sector is represented by E_{ij}^{g} as follows:

$$E_{ij}^{g} = \left(\mu_{ij}, \sigma_{ij}^{2}\right)_{g} = \left(\frac{\sum_{k} \mu_{ijk}}{k}, \frac{\sum_{k} \sigma_{ijk}}{k}\right) \tag{4.5}$$

where E_{ij}^{g} signifies average expected mean and variance reference for each sector j for genuine node i.

4.5.2 Training Phase 2: Cosine Similarity Data

The secondary training phase tests the received traffic from node i against the authenticated reference E_{ij}^{g}. As a result of multipath fading, mean and variance for received traffic, $G_{ij} = \left(\mu_{ij}, \sigma_{ij}^{2}\right)$, will deviate from the reference E_{ij}^{g}. For this reason, cosine similarity is employed to obtain expected cosine similarity values for each sector. Cosine similarity is a measure of similarity between two nonzero vectors of an inner product space that measures the cosine of the angle between them. Output values range from -1 to 1 with 1 representing two vectors with the same orientation, -1 representing opposite orientation, and 0 representing orthogonal orientation.

In order to enact cosine similarity, E_{ij}^{g} and G_{ij} will need to be transformed into vectors. To do so, origin (0,0) is used as a reference point to translate E_{ij}^{g} and G_{ij} into $\vec{E_{ij}^{g}}$ and $\vec{G_{ij}}$. Now cosine similarity can be computed as:

$$Similarity = cos\left(\Theta\right) = \frac{\vec{E_{ij}^{g}} \cdot \vec{G_{ij}}}{\left\|\vec{E_{ij}^{g}}\right\| \left\|\vec{G_{ij}}\right\|} \tag{4.6}$$

The number of cosine similarity values is inherently equal to ρ. Furthermore, genuine traffic generates cos-sim readings closer to 1. The values are then stored in an array, G_T, to serve as the point of reference in the operational phase.

$$G_T = \left\{CS_1, CS_2, CS_3, \cdots, CS_j\right\} \tag{4.7}$$

4.5.3 Detection Phase: Correlate Cosine Similarity

If the difference exceeds α, the system is presumed compromised by a spoofing attack.

4.5.4 Threshold Selection and False Positive Rate

Threshold and false positive rate are directly correlated to sensitivity factor. As n decreases, the sensitivity increases. Likewise Increasing n decreases the sensitivity. As noticed, n is indirectly proportional to sensitivity.

The presence of multipath fading causes a distribution of cosine similarity values that deviate from 1. Therefore, for determining the threshold a value at the lower end of the distribution tail is chosen to reduce the false positive rate. Each measure of sensitivity will have a minimum cosine similarity value for genuine traffic. By using the minimum point as the frame of reference for threshold selection, α can be denoted by:

$$\alpha = 1 - min\,(Similarity) \tag{4.8}$$

4.6 Experimental Results

Performance of the proposed system was tested in Georgia Southern University's Optical Network and Smart Grid Application (ONSmart) lab. The wireless network test-bed contained 3 National Instrument Universal Software Radio Peripheral 2921 devices operating under QPSK modulation. Two of the devices were placed 30 feet apart representing a smart device communicating with a smart meter. The third device, acting as the spoofing node, was placed in several locations to test effectiveness. Figure 4.1 shows the experimental setup used.

In this experimental setup, 10 training datasets were transmitted from smart device to the smart meter. Moreover, the sensitivity factor n was set to 3 thus mean and variance

Figure 4.1: Experimental setup

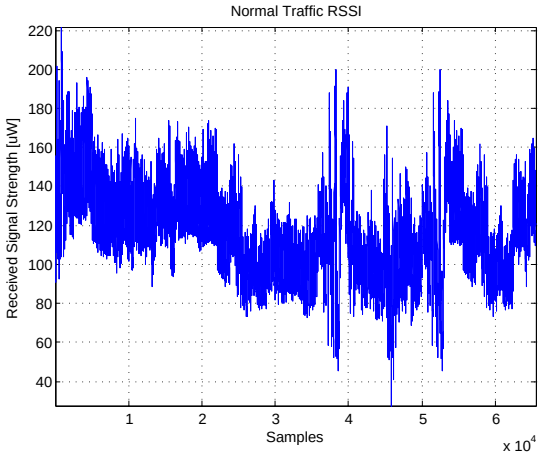
values were obtained for every 2^3 samples. Additionally, the operational phase threshold was set to 0.2. Table 4.1 summarizes the experimental parameters used in this experiment.

Table 4.1: Experimental Parameters Used

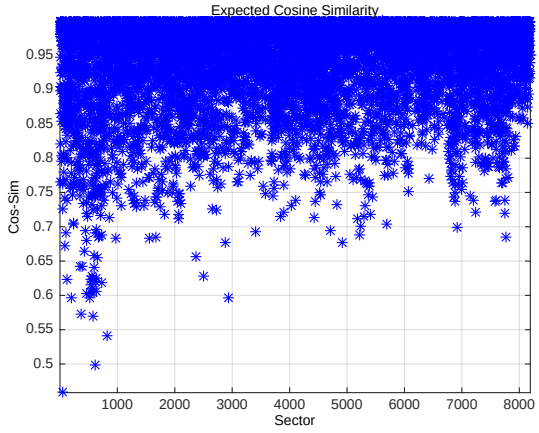
Parameters	Values
Distance between AMI & Device	30 ft
Distance between Spoofer & AMI	8, 12, 15, 30 ft
Frequency	2.4 GHz
Modulation	QPSK
Sensitivity (n)	3
RSSI Training Data Sets	10
Threshold (α)	0.2

The second phase of training obtained RSSI measurements and expected cosine similarity values as shown in Figure 4.2. Figure 4.2 (a) illustrates RSSI values from a smart device without spoofing. Figure 4.2 (b) exhibits expected cosine similarity measurements of normal traffic when compared to the average mean and variance calculated in training phase 1. As can be seen, due to multipath fading, the RSSI measurements deviate significantly.

Subsequent to the training phase is the operational phase. In this phase, the spoofing



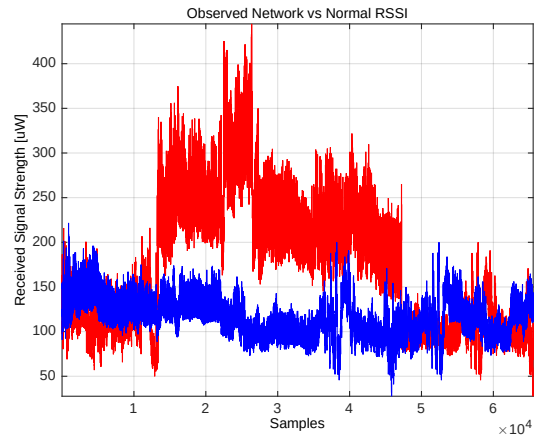
(a)



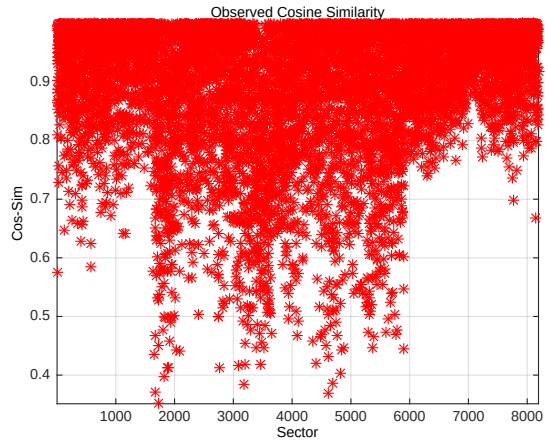
(b)

Figure 4.2: (a) Normal Traffic RSSI measurements (b) Expected Cosine Similarity values per sector

node was placed at distances of 8 feet, 12 feet, 15 feet, and 30 feet respectively from the smart meter node. Figure 4.3 (a) shows the spoofed RSSI measurements versus normal RSSI measurements for spoofing node placement of 15 feet. Figure 4.3 (b) displays the cosine similarity values obtained for the spoofed node. It could be noticed that the cosine similarity values are much lower than the expected values thus proving feasibility of detection in the operational phase.



(a)



(b)

Figure 4.3: (a) Observed vs. Normal RSSI measuremnts (b) Cosine Similarity of spoofing node at 15 ft from smart meter

Figure 4.4 shows detection results of compared values. This figure illustrates the effectiveness of the proposed detection algorithm under the influence of multipath fading. With the threshold set to 0.2, the algorithm has a false positive rate of 3.05% and a false negative rate of 9.09%. This distance represented the worst case scenario for this experiment. To highlight the effect of changing the distance between the spoofing node and the AMI on the performance of the proposed IDS method, the false positive rate, false negative rate, and accuracy at 8, 12, 15, and 30 ft are summarized in Table 4.2.

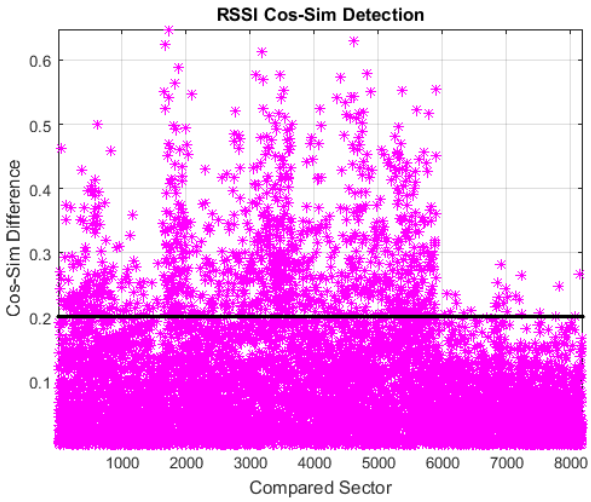


Figure 4.4: Cosine Similarity detection output

Table 4.2: Performance at different distances between AMI and spoofing node

Distance	False Positive Rate	False Negative Rate	Accuracy
8 ft	2.37%	6.12%	91.51%
12 ft	2.89%	7.79%	89.32%
15 ft	3.05%	9.09%	87.86%
30 ft	5.53%	3.2%	91.27%
Average	3.46%	6.55%	89.99%

Table 4.3 provides the average false positive rate, false negative rate, and accuracy for the IDS as the threshold varies across all tested distances. As expected, the false positive and false negative rates are inversely propositional with the optimal accuracy value at threshold level (α) of 0.2. At $\alpha = 0.2$, the maximum average detection accuracy across all tested distances is 89.99%.

Table 4.3: Performance at different threshold values

Threshold	False Positive Rate	False Negative Rate	Accuracy
0.1	13%	2.6%	84.4%
0.2	3.46%	6.55%	89.99%
0.3	0.9%	44.9%	54.2%
0.4	0.057%	78.5%	21.443%
0.5	0.011%	94%	5.989%

4.7 Conclusion

In this chapter, a cosine similarity based algorithm using RSSI characteristics was introduced for spoofing detection in smart grid 802.11 based Home Area Networks. The optimal threshold was found to be equal to 0.2 for the experimental setup used. At optimal threshold level, the proposed algorithm was able to detect spoofing with an accuracy of 87.86% at 15 ft distance between AMI and spoofing node (worse case scenario). The results demonstrate the effectiveness against multipath fading which is a major factor that affects false positive rates. Moreover, the nature of the algorithm enables adaptability for different antenna types associated with smart devices. Therefore, the algorithm provides high detection rate for HANs comprised of devices with different antenna structures. Also, the proposed algorithm is adaptable to any kind of communication infrastructure. For example, the same approach can be used for Zigbee based HAN. This algorithm is best utilized for AMI environments in which smart devices are stationary but could be used for mobile smart device. In essence, the change in distance between the smart meter and mobile smart device will cause variation in false positive and false negative rates thus varying the detection accuracy dynamically. Finally, the proposed algorithm can be easily applied to Neighborhood Area Networks (NAN) and Wide Area Network (WAN).

CHAPTER 5

MOVING TARGET DEFENSE INTRUSION IN SMART GRID

5.1 Overview

Moving Target Defense (MTD) has gained popularity in combating moving target attacks. MTD is the concept of controlling change across multiple system dimensions in order to increase uncertainty and apparent complexity for attackers, reduce the window of opportunity, and increase the cost of their probing and attack efforts [42]. Once exploit cost exceeds attack gain, an intruder will deem the operation unprofitable.

5.2 Related Works

Several researchers have examined moving target defense methods for the IPv6 address space. Moving Target IPv6 Defense (MT6D) was developed in [43, 44] that leveraged the immense address space of IPv6. Security is obtained by rotating addresses of both the sender and receiver. Also, the addresses are capable of rotating mid-session to prevent an attacker from discovering node identities. In the smart grid, nodes in AMI are limited in computational resources. This approach requires the host device to enact address rotation which may cause latency issues.

In [45], a Sliding Window and Full Transparent (SWIFT) scheme was proposed for IPv6 address mutation. The scheme focuses on address mutation with very high frequency and reduction in packet loss by means of real and virtual IPs generated by a rolling window. Here the algorithm faces issues of computational complexity. When combined with limited resources, this approach can give way to denial of service (DoS) vulnerabilities.

Moving target defense for smart grid defense was applied in [14] in which MT6D was adapted for smart grid communications. This approach dwells mostly on the relationship between client and server. In the case of smart grid, the relationship correlates to com-

munication strictly between smart meter and utility company not amongst smart meters. Moreover, MT6D encapsulation calls for increased storage capacity due to the amount of bytes added in communication.

5.3 Threat Model

Ordinarily, in the reconnaissance stage, an attacker will have a profile of the network using reconnaissance software for packet sniffing (e.g. Wireshark). After analyzing traffic patterns, there is a possibility that an attacker may attempt an array of network layer attacks for domain access. Recent development in attack tools and techniques has brought about Moving Target Attacks (MTA). MTA gives intruders the upper hand in network penetration by means of randomization in attack vectors to evade detection. MTA, to name a few, include polymorphism, metamorphism, obfuscation, and encrypted exploits. Polymorphism changes malware signature whereas metamorphism changes malware code on the fly. Obfuscation conceals code and logic. Encrypted exploits are able to bypass investigation by changing signatures and parameters.

This new era of attack vectors has brought great attention to anomaly detection algorithms. In anomaly detection, data patterns are analyzed to decipher whether the data is of genuine or malicious origin. Another advantage of anomaly detection is the ability to identify attack vectors independent of historic signatures.

5.4 MTDIDS Algorithm

The proposed Moving Target Defense Intrusion Detection System (MTDIDS) [46] features analysis of network entropy for attack detection. Integrating entropy into network operations such as IP, port, and packet selection creates a moving target effect. Furthermore, the inherited dynamic attack surface correlates to significant cost increase of reconnais-

sance tools utilized by intruders. Additionally, MTDIDS demonstrates anomaly detection therefore proving effective for detecting zero day attacks. Zero day attacks represent attack vectors in which there is no prior knowledge. Without prior knowledge, attack detection signatures in static intrusion detection systems cannot be generated.

MTDIDS is comprised of three training phases and two detection phases. Two of the training phases take place in the coordinator node. The last training phase and detection phases take place in nodes wishing to communicate, i.e. smart meters. The coordinator node manages routing for all smart meters in a specified area network. If an attack is detected, the smart meter sends an alarm to the utility company or enacts predefined preventative measures. Training and operational algorithms are detailed as follows.

5.4.1 Training Phase 1: Random Routing Table Generation

As aforementioned, smart grid advance metering infrastructure consists of smart devices (i.e. smart meters) communicating amongst one another for distributed control or cooperation. Moreover, the hierarchical topology (Home, Neighborhood, and Wide Area Networks) is indicative of area networks that contain a respective number of nodes. For that reason, the first training phase employs a coordinator to generate random session routing tables for packet transmission by nodes in a given area.

Length of the routing table is determined by the packet analysis length parameter. The packet analysis length parameter is set by utility company and governs the increment in which packet trajectories will be mapped by the routing table for transmission and analyzed by the receiver. For example, if the packet analysis length was set to be 1024 then outgoing and incoming packets will be sent and analyzed 1024 at a time in accordance to the routing table. Thus MTDIDS features rolling window capabilities for real-time anomaly detection. Each packet in the table is assigned random IP and port for a given session. The address range, pool of addresses and ports employed for an area, is determined by IP acquisition

and allocation set by utility companies. Session constitutes the time period in which the generated table is considered valid. Once the time period has elapsed, the current session's routing table is invalidated and a new random routing table is generated for the new session. As the sessions transition, the number of IPs and ports used to transmit packets are chosen at random from the provided pool of IPs and ports. Once selected, the IPs and ports are distributed randomly across the packets set by the packet analysis length parameter.

Table 5.1 illustrates the general layout of a randomly generated routing table for a given session. In the first column, the predetermined number of packets to transmit and analyze at a time is set. The second column randomly selects an IPv6 address from the pool of addresses designated for the system. Finally, the third column randomly selects a port from the available pool of ports.

Table 5.1: Random IP and Port Assignment Per Packet

Packet Number	IPv6 Address	Port Assignment
1	Rand(IP)	Rand(Port)
2	Rand(IP)	Rand(Port)
3	Rand(IP)	Rand(Port)
.	Rand(IP)	Rand(Port)
.	Rand(IP)	Rand(Port)
Packet Analysis Length	Rand(IP)	Rand(Port)

5.4.2 Training Phase 2: Parity Packet Selection

In this system, parity packets represent packets from the routing table that are to be appended with security bits. The sequence of bits can be determined by vendor or utility company at time of deployment. Accordingly, parity rate is a randomly generated number that

constitutes the increment in which parity packets are selected from the randomly generated routing table. For example, if the parity rate is set to 6 then every 6th packet in the randomly generated table will serve as a parity packet. Parity rate is randomly generated at the time of routing table generation. Adding parity bits serve two main purposes which are, 1) adding a second security dimension through the randomization of parity rate across sessions, 2) in case of intruder obtaining the routing table with correct IP addresses and port numbers, the parity check will allow us to detect data injection or falsification as a second tier detection.

5.4.3 Trianing Phase 3: Planar Key Development

Consecutively, Training Phase 3's objective is to securely deliver this data to respective nodes in the area. Once received, the nodes will take the routing table and parity information to create expected planar signatures that will serve as a session's planar key for incoming traffic to be used in the detection phase. Planar keys are created for both regular and parity packet distributions with the total number of signature planes in a key being equal to the number of selected IPs. Signature planes are created using information for each packet in the routing table and mapping the points onto planes. Each element on a signature plane signifies a coordinate in the form of (Packet Number, IP Address, Port Number). Thus a planar key can be generated for a given session with the array of mapping points as detailed in Table 5.2.

In Table 5.2, N is the maximum value set by the packet analysis length parameter, while Rand(IP) and Rand(Port) are associated IP and port addresses determined in Training Phase 1. Planar key for parity packets is obtained in the same fashion. Parity mapping also follows the same nomenclature, (Parity Packet Number, Parity Rand(IP), Parity Rand(Port)). More details and visualizations of planar keys and comprised signature planes are provided in Section 5.5.

Table 5.2: Creation of Mapping Points

Packet	Planar Key Coordinate
1	(Packet Number 1,Rand(IP),Rand(Port))
2	(Packet Number 2,Rand(IP),Rand(Port))
3	(Packet Number 3,Rand(IP),Rand(Port))
.	.
.	.
N	(Packet Number N,Rand(IP),Rand(Port))

5.4.4 Detection Phase: Planar Signature Analysis

In the detection phase, incoming packets are incrementally analyzed according to the packet analysis length. The packets are thus mapped onto planes to create an observed planar distribution that inherently is identical in size to the planar key. Anomaly detection is possible when the observed planar distribution of incoming network traffic is compared against the planar key generated for the session. Therefore, due to controlled entropy characteristics, anomaly detection is governed by the relationship in Eq. 5.1.

$$|E\left(Packet, IP, Port\right) - O\left(Packet, IP, Port\right)| \stackrel{?}{=} 0 \tag{5.1}$$

where $E(Packet, IP, Port)$ denotes each packet's expected IP and Port determined by the randomly generated routing table in Training Phase 1. $O(Packet, IP, Port)$ represents incoming packets to be compared. Incoming packets collected by nodes are expected to match the planar key. Because of this, the difference should always be zero (threshold) thus representing system stability. Nonzero outputs from the detection equation result in the creation of difference points. The difference points intrinsically populate into difference planes to signify an anomaly. In the presence of genuine traffic, the difference should equal

zero for every packet which implies that a difference plane will not exist. Instead, one would notice a singularity located at the origin. As anomalies from malicious packets are detected, the singularity ceases to exist and difference planes begin to populate. This validates that difference in expected packet and observed packet is suitable for attack detection in this system. The same approach is applied to parity packets thus creating another layer of security an attacker would have to bypass. Anomaly detection in parity packets is governed by the relationship in Eq. 5.2.

$$|E\left(Parity, IP, Port\right) - O\left(Parity, IP, Port\right)| \stackrel{?}{=} 0 \tag{5.2}$$

$Parity$ indicates the packets appended with security bits as determined in Training Phase 2. The parity detection layer is advantageous because it presents two obstacles to an intruder. The first obstacle is having to determine the parity rate for the current session. The second obstacle requires an intruder to decipher security bits of the parity packets. Bypassing both obstacles, while simultaneously handling obstacles in the first detection layer, prior to session invalidation requires attack techniques of high complexity and cost. Costs associated with acquiring resources become substantial therefore diminishing profitability.

5.5 Results and Analysis

Evaluation of MTDIDS is carried out via smart grid AMI simulated in MATLAB in which two nodes, A and B, wish to communicate; a coordinator node generates randomized routing tables and parity rate for packet trajectory and planar key creation. A malicious node attempts to mimic network traffic to acquire domain access. As previously stated, reconnaissance tools could provide an attacker information about network patterns.

5.5.1 MTDIDS Session 1

In session 1, the coordinator node has selected five random IPv6 addresses from a pool of IPv6 addresses set by utility company for a respective area network. The session is set to be valid in 10 minute intervals. Hence, a new random routing table is generated every 10 minutes. Additionally, parity rate of 3 has been selected for session 1. Table 5.3 shows the IPv6 addresses chosen by the coordinator for this session.

Table 5.3: Session 1 IPv6 Address Selection

1	f3d3:1999:9616:40c9:5e39:bfc9:99b2:3ef7
2	842d:e0e6:93da:d10a:dab7:4ea2:d754:9943
3	d8a9:a45e:174f:a42a:52f7:2912:96c9:e095
4	5421:7418:8171:71c7:d1a3:aaa8:99db:6ee8
5	b02f:8d56:f541:b686:6185:8bac:2931:ab5e

For session 1, the packet analysis length has been set to 25,000. The five chosen IPs are then randomly distributed amongst the 25,000 packets. Furthermore, the utility company has chosen to utilize all 65536 ports associated with TCP/UDP for random selection. Similar to IP address selection, an assortment of ports are chosen in random fashion and distributed. The total number of ports chosen is directly proportional to the packet analysis length. Therefore, every 25,000 packets will follow the trajectory set by the routing table. Likewise, every 25,000 packets are compared against the planar key generated by MTDIDS for the current session. MTDIDS is able to create planar keys once routing table and parity rate are securely delivered to nodes A and B. Figures 5.1 and 5.2 exhibit planar keys created by MTDIDS for session 1 where each signature plane represents the signature for traffic across selected IPv6 addresses. Figure 5.1 illustrates planar key for all data packets. Figure 5.2 shows planar key for parity packets.

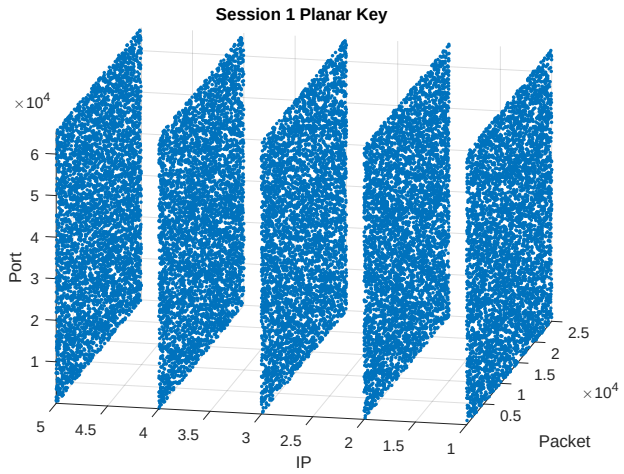


Figure 5.1: Session 1 Data Planar Key

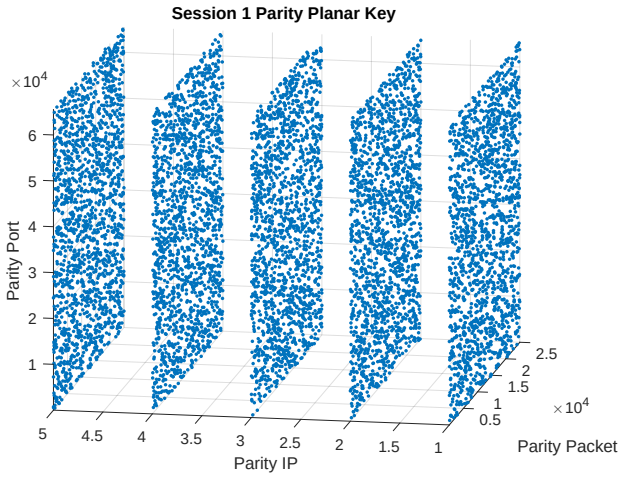


Figure 5.2: Session 1 Parity Planar Key

5.5.2 MTDIDS Anomaly Detection

Incoming packets from node A are compared against the planar key's signature planes created by MTDIDS in node B. Equally, incoming packets from node B are compared against the same planar key created in node A. As shown in Figure 5.3, when network traffic is genuine, only a singularity exists in the difference plane. The singularity represents a secure system with not intrusion.

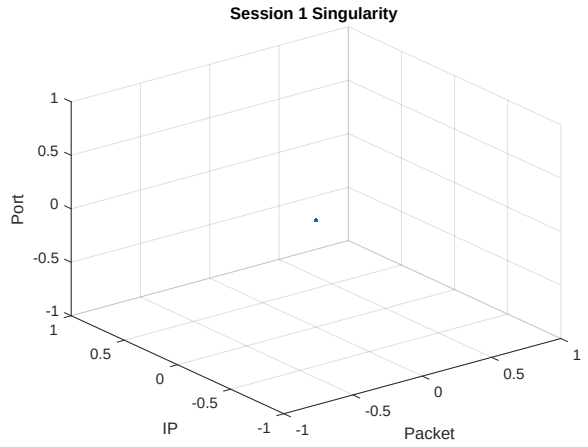


Figure 5.3: Session 1 Singularity

The goal of a malicious node is to inject false data into either A or B for impersonation purposes. In Figure 5.4, incoming traffic from the malicious node has been mapped and superimposed against session 1's planar key. Malicious data is denoted as red whereas blue represents expected.

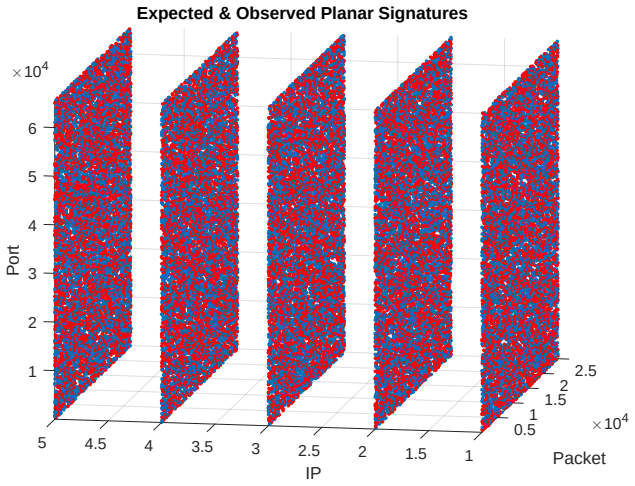


Figure 5.4: Expected and Observed Data Planar Signatures

As the second layer of defense, observed parity packets are compared to session 1's parity planar key. Figure 5.5 shows expected and malicious signature planes for parity

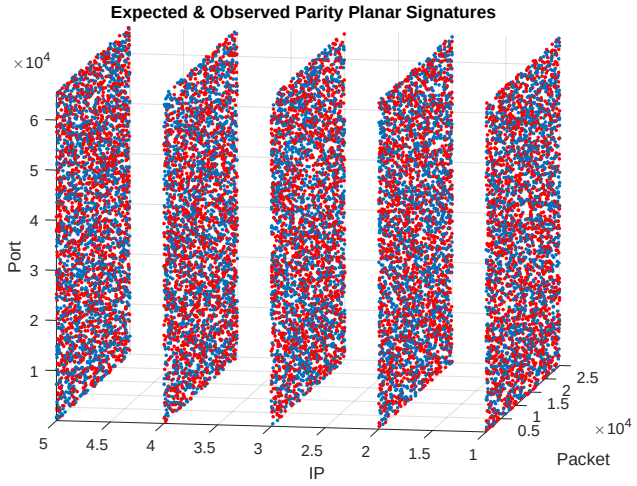


Figure 5.5: Expected and Observed Parity Planar Signatures

packets received. Since parity packets are appended with security bits, an attacker is tasked with determining packet trajectories and security bit sequence while evading detection from the obstacles in MTDIDS initial detection phase.

Figures 5.6 and 5.7 exemplify the generated difference planes for anomalies detected in the packet and parity packet detection phases. Each element that populates the difference plane provides IP, port, and packet in which the anomaly took place. For that reason, MTDIDS greatly compliments forensic efforts.

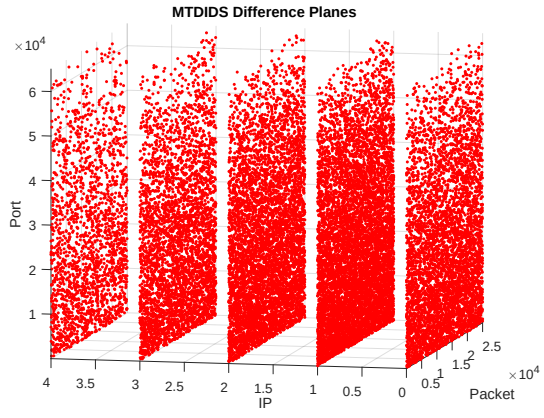


Figure 5.6: MTDIDS Data Difference Planes

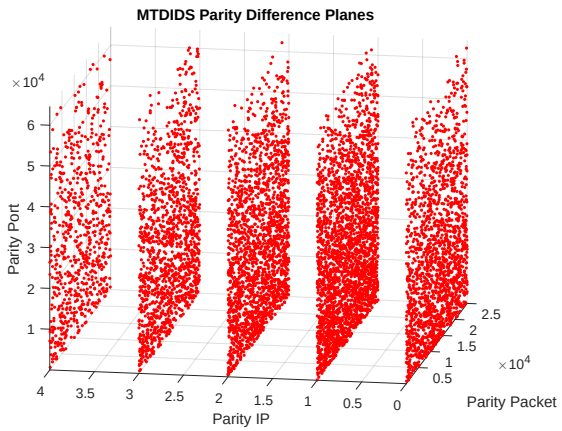


Figure 5.7: MTDIDS Parity Difference Planes

5.6 Dynamics of Multiple Sessions

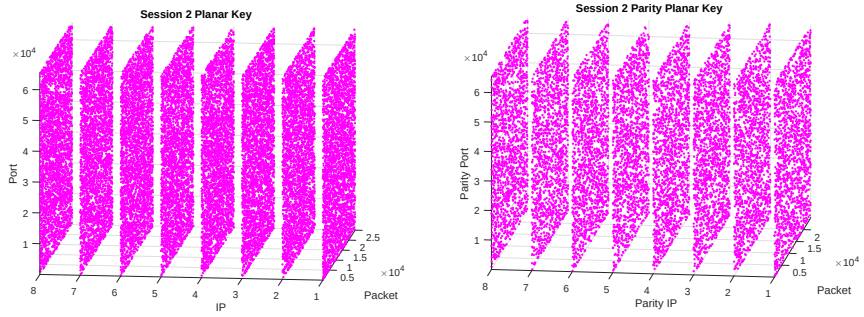
By rendering planar keys valid for a given session, the element of dynamic planar signatures is introduced as another layer of defense. For a single session, as seen in session 1, an attacker is tasked with mapping packets, IP, and ports while simultaneously evading detection. Varying the size of planar keys for each session dramatically decreases allotted time to exploit a system.

For sessions 2 and 3, the packet analysis length of $25,000$ and the ports are the same are identical to session 1. The number of IPs utilized and parity rate for each session has been altered due to the randomization of the routing table in the generation phase. For session 2, the number of IPs selected from the pool of IPS is 8 with parity rate 2. Selected IPs are shown in Table 5.4 with planar and parity keys displayed in Figure 5.8. In session 3, the number of IPs selected is 3 with parity rate 3. IPs and planar keys for session 3 are presented in Table 5.5 and Figure 5.9.

Figures 5.10-(a,b) exhibit all three sessions superimposed to demonstrate dynamic planar keys over the duration of 30 minutes. Blue represents session 1, magenta represents session 2, and green represents session 3. As noticed, the number of signature planes per

Table 5.4: Session 2 IPv6 Address Selection

1	26c2:d97d:b35a:71d8:c4de:373:ea17:2551
2	6645:7e43:b8e9:562d:ba14:2d1c:df9c:a031
3	eb38:963d:b771:2fc7:8f34:4f6a:2d05:7eba
4	f5e3:4e47:e091:473d:7f62:69d:a77e:4c4f
5	99d7:57e9:a42a:d3e3:88d0:682e:2a79:d296
6	d1cd:f1dd:108:28fa:b19a:d2d6:79a0:172b
7	f0a4:1d65:e52a:d4ea:7b01:f6df:5338:a0dc
8	e858:60ab:76c2:26de:8506:def:d07d:1081



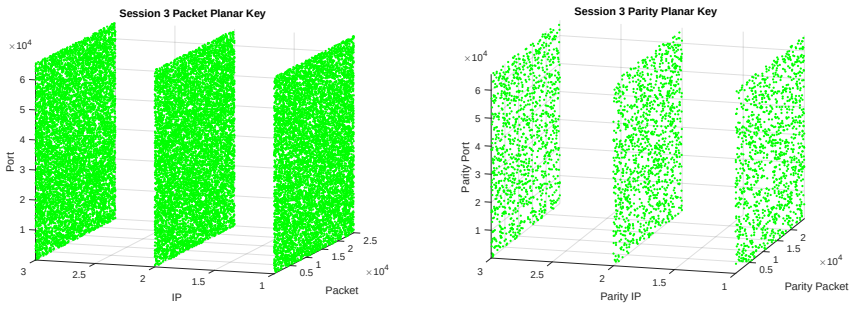
(a) (b)

Figure 5.8: Session 2 (a) Planar Key (b) Parity Planar Key

Table 5.5: Session 3 IPv6 Address Selection

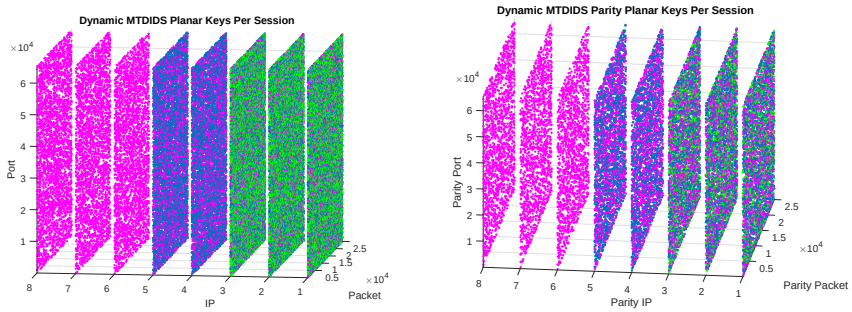
1	836e:ccdf:7017:da18:6c92:3c20:800a:6f88
2	be8c:36db:a799:9b2d:dd1c:a2d3:3d:3f82
3	aeb3:6658:98ae:4acf:b100:9f88:67d6:742c

planar key varied over the course of 30 minutes. Therefore, an attacker would have to figure 3 different trajectories for the same 25,000 packets.



(a) (b)

Figure 5.9: Session 3 (a) Planar Key (b) Parity Planar Key



(a) (b)

Figure 5.10: Dynamic Session (a) Planar Keys (b) Parity Planar Keys

5.7 Conclusion

In this chapter, a network based Moving Target Defense Intrusion Detection System for anomaly detection in smart grid AMI has been introduced. Remarkably, experimental data confirms MTDIDS capabilities of real-time anomaly detection. In addition, the dynamic IPv6 address space featured in MTDIDS gives way to a varying attack surface throughout the smart grid. As a result, the costs of exploits rise dramatically which deters attackers by virtue of exploits no longer being profitable. In the event an attacker is able to compromise a node, the amount of harm done will be minimal because of invalidation caused by changing session planar keys. MTDIDS is adaptable for wired and wireless infrastructures. In both

environments, packets are analyzed to create packet trajectories and planar signatures on receiving ends. Additionally, the proposed approach can be applied to systems that function outside of smart grid applications. The entropic nature of MTDIDS is ideal for securing patient data in health-care systems under HIPAA regulations (Health Insurance Portability Accountability Act). For Ad-Hoc networks such as VANET, Vehicular Ad-Hoc Network, cars in proximity can remain secure for autonomous operations. In wireless networks, MTDIDS is functional independent of antenna type, modulation scheme, etc. In all, the versatility of MTDIDS adds stability across many functional areas thus having significant impact.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

The Smart Grid has been introduced as the next generation smart energy grid which features improved reliability, manageability, and interconnectivity. Due to the integration of a two-way communication network, vulnerabilities associated are thus inherited by the smart grid. Furthermore, the vastness of the smart grid network creates a much larger attack surface that can lead to catastrophic events if exploited. Additionally, smart network devices possess limited computational resources which increases ease and effectiveness of attack vectors. For this reason, the primary requirement for smart grid security is an algorithm that is robust and within the computational confines of smart devices. To meet such a requirement, intrusion detection system algorithms have been presented in this work.

Chapter 2 provides an overivew of the smart grid area networks along with communication and address space protocols. Home, neighborhood, and wide area networks provide the hierarchial topology in which the commmunication infrastructure can be integrated. Moreover, the convergence of energy and communication systems usher in jamming and spoofing attacks that can lead to blackouts, access to sensitive information, or loss of control of the network. Historical data of said attack vectors enable the use of signature based intrusion detection systems to monitor network traffic and recognize attack signatures. In the case of new era attacks where the is no historical data, signature based approaches become useless. In this case, anomaly detection is necessary to detect abnormilities in network traffic.

Chapter 3 demonstrates intrusion detection algorithms for jamming attacks in WLAN and LTE networks. Theoretical analysis for both algorithms is provided. In WLAN jamming, received signal strength indicator and packet loss rate were the two parameters

used to detect jamming attacks. In the presence of a jamming attack, RSSI and packet loss rate increases simultaneously. The algorithm takes the data and compare to a threshold that represents normal network traffic. Exceeding the threshold signifies that the system is under a jamming attack. In LTE jamming, synchronization signals were explored and characteristics were utilized to detect jamming. In this type of attack, a jaming attack with great precision is necessary. The jammer must first detect when synchronization signals are transmitted and generate a jamming signal with the frequency of the observed subcarrier. In the event that the attack is effective, smart devices will not be able to establish communication with cellular towers. By means of signal strength and cosine similarity of primary synchronization signals, jamming attacks are detected in similar manner to WLAN. Moreover, both algorithms provide real time attack detection.

Chapter 4 explores the usage of RSSI to detect spoofing attacks in 802.11 home area networks. In this algorithm, RSSI training data is utilized to create unique signatures for all smart devices within a consumer's home. The RSSI training data streams from each device enables the IDS to learn variations in RSSI of respective devices because of the effects of spatial correlation. Additionally, a sectoral cosine similarity is used as a second layer of defense to improve detection rate. In the presence of a spoofing attack, a malicious user attempts to impersonate a genuine user or device to gain access to the network. To be effective, a spoofer would need to imitate device behavior. On the network and data link layers, imitation of device characteristics is feasible. On the physical layer, imitating RSSI affected by spatial correlation becomes a very daunting task. Another obstacle for an attacker is to imitate thousands cosine similarity sectors.

Chapter 5 introduces Moving Target Defense Intrusion Detection System (MTDIDS). MTDIDS serves as an anomaly based IDS to detect irregularities in network traffic. A new era of attacks known as moving target attacks has provided malicious users the upperhand in the cyber world. These attacks are able to change characteristics during exploits which

enables them to easily bypass signature based intrusion detection systems. MTDIDS incorporates entropy into network operations to create an ever changing attack surface. Routing tables are generated in random fashion for a given session by a coordinator node. The routing tables are composed of multiple IP, ports, packet number, and parity value. Once received by respective nodes in the area, planar keys are generated to serve as the basis of the IDS portion of MTDIDS. To further increase adversity for a malcious user, IPv6 address space is utilized to minimize the ability of sniffing tools. In essence, a dynamic attack surface in turn decreases the chances of success for an attacker while increasing exploit cost. When exploit costs exceed profitability, the exploit is deemed ineffective.

In this work, four intrusion detection algorithms were proposed for jamming, spoofing, and anomolous attack detection. In chapters 3 and 4, the algorithms are signature based (static) in nature and are capable of detecting jamming and spoofing in WLAN and LTE. Additionally, these algorithms can be applied to Zigbee networks. MTDIDS in chapter 5 is an anomaly based approach that detects irregularities in network traffic while deterring attackers by signficantly reducing the profitability of exploits. MTDIDS serves as the next generation intrusion detection algorithm that uses characteristics of the new era of attacks against attackers themselves.

6.2 Future Works

This thesis provides extensive analysis of intrusion detection algorithms in smart grid WLAN and LTE communication networks. Furthermore, the algorithms applied Physical, MAC, and Network layer characteristics that enable layer 1, 2, and 3 in smart devices throughout the smart grid. Future works include further investigation of RSSI characteristics that are unique to attack vectors in order to generate physical layer attack vector signatures. Additionally, spatial correlation of RSSI is a critical factor in detection rate. Therefore, complete command of the nature of signals is necessary. In the spoofing detection algorithm,

sensitivity factor determines false positive and false negative rates. By studying the effect in different network setups, accuracy can be improved.

MTDIDS serves as the new era of IDS because of the ability to detect anomalies with a need for signature and dynamic attack surface that deems existing exploits unprofitable. Further study on MTDIDS includes testing against larger network setups to investigate latency, feasibility, potential bottlenecking, and computational exhaustion. With that said, MTDIDS is ideal for smart devices comprised of solid state technology. MTDIDS has been tested under standard SATA and SSD conditions. The outcome demonstrated that MTDIDS operating on a SSD reduces the latency by roughly four and a half times. By incorporating solid state technology, the algorithm becomes very robust. In essence, future work focuses on the abilities of solid state technology in conjunction with intrusion detecton algorithms to reduce computational complexity and amount of time needed to detect compromisations within the network.

REFERENCES

[1] H.J.Liao, C. R. Lin, Y. Lin, and K.Y.Tung, "Historic and projected u.s. electricity demand, 1950-2050," Rocky Mountain Institute, 2010, http://www.rmi.org/RFGraph-US_electricity_demand.

[2] I. Keysight Technologies, "Lte physical layer overview," Keysight Technologies, Inc., 2015, http://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/lte/content/lte_overview.htm.

[3] G. Lu, D. De, and W. Z. Song, "Smartgridlab: A laboratory-based smart grid testbed," in *2010 First IEEE International Conference on Smart Grid Communications*, Oct 2010, pp. 143–148.

[4] Y. Zhang, L. Wang, W. Sun, R. C. G. II, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 796–808, Dec 2011.

[5] R. Berthier and W. H. Sanders, "Specification-based intrusion detection for advanced metering infrastructures," in *2011 IEEE 17th Pacific Rim International Symposium on Dependable Computing*, Dec 2011, pp. 184–193.

[6] "Multi-vendor Penetration Testing in the Advanced Metering Infrastructure," ACSAC Applied Computing Security Assoc, 2010, http://www.patrickmcdaniel.org/pubs/acsac10b.pdf.

[7] Y. Sun, X. Guan, T. Liu, and Y. Liu, "A cyber-physical monitoring system for attack detection in smart grid," in *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2013, pp. 33–34.

[8] D. Evans, A. Nguyen-Tuong, and J. Knight, "Effectiveness of moving target defenses," *Effectiveness of Moving Target Defenses*, vol. 54, p. 29ï£¡48, Aug 2011. [Online]. Available: https://www.cs.virginia.edu/~evans/pubs/mt2011/

[9] M. Harvey, D. Long, and K. Reinhard, "Visualizing nistir 7628, guidelines for smart grid cyber security," in *2014 Power and Energy Conference at Illinois (PECI)*, Feb 2014, pp. 1–8.

[10] A. Aggarwal, S. Kunta, and P. K. Verma, "A proposed communications infrastructure for the smart grid," in *2010 Innovative Smart Grid Technologies (ISGT)*, Jan 2010, pp. 1–5.

[11] D. B. Rawat and C. Bajracharya, "Cyber security for smart grid systems: Status, challenges and perspectives," in *SoutheastCon 2015*, April 2015, pp. 1–6.

[12] D. Rawat and C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1652–1656, Oct 2015.

[13] M. J. Karam and F. A. Tobagi, "Analysis of the delay and jitter of voice traffic over the internet," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, vol. 2, 2001, pp. 824–833 vol.2.

[14] S. Groat, M. Dunlop, W. Urbanksi, R. Marchany, and J. Tront, "Using an ipv6 moving target defense to protect the smart grid," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, Jan 2012, pp. 1–7.

[15] "U.S. and World Population Clocks," U.S. Census - Website, 2011, http://www.census.gov.

[16] Y. Xu and C. Fischione, "Real-time scheduling in lte for smart grids," in *2012 5th International Symposium on Communications, Control and Signal Processing*, May 2012, pp. 1–6.

[17] F. Granelli, D. Domeniconi, N. L. S. D. Fonseca, and B. Tsetsgee, "On the usage of wifi and lte for the smart grid," in *2014 7th International Conference on Ubi-Media Computing and Workshops*, July 2014, pp. 1–5.

[18] B. Holfeld, S. Jaeckel, L. Thiele, T. Wirth, and K. Scheppelmann, "Smart grid communications: Lte outdoor field trials at 450 mhz," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[19] D. Grochocki, J. H. Huh, R. Berthier, R. Bobba, W. H. Sanders, A. A. Cï£¡rdenas, and J. G. Jetcheva, "Ami threats, intrusion detection requirements and deployment recommendations," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, Nov 2012, pp. 395–400.

[20] H.J.Liao, C. R. Lin, Y. Lin, and K.Y.Tung, "Intrusion detection system: A comprehensive review," Journal of Network and Computer Applications, 2012, http://dx.doi.org/10.1016/j.jnca.2012.09.004.

[21] P. Jokar, H. Nicanfar, and V. C. M. Leung, "Specification-based intrusion detection for home area networks in smart grids," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct 2011, pp. 208–213.

[22] Z. Lu, W. Wang, and C. Wang, "From jammer to gambler: Modeling and detection of jamming attacks against time-critical traffic," in *2011 Proceedings IEEE INFOCOM*, April 2011, pp. 1871–1879.

[23] ——, "Modeling, evaluation and detection of jamming attacks in time-critical wireless applications," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1746–1759, Aug 2014.

[24] G. Liu, J. Liu, Y. Li, L. Xiao, and Y. Tang, "Jamming detection of smartphones for wifi signals," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–3.

[25] S.Nadeem, A. Nazar, and Z. Muhammad, "Detection of jamming attacks in 802.11b wireless networks," EURASIP Journal on Wireless Communications and Networking, 2013, http://jwcn.eurasipjournals.springeropen.com/articles/10.1186/1687-1499-2013-208.

[26] L. Avanco, A. E. Guelfi, E. Pontes, A. A. A. Silva, S. T. Kofuji, and F. Zhou, "An effective intrusion detection approach for jamming attacks on rfid systems," in *2015 International EURASIP Workshop on RFID Technology (EURFID)*, Oct 2015, pp. 73–80.

[27] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "Ofdm and its wireless applications: A survey," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[28] D. Astely, E. Dahlman, A. Furuskï£¡r, Y. Jading, M. Lindstrï£¡m, and S. Parkvall, "Lte: the evolution of mobile broadband," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 44–51, April 2009.

[29] "LTE and the Evolution to 4G Wireless-Design and Measurement Challenges," John Wiley and Sons, Ltd., 2013, http://onlinelibrary.wiley.com/book/10.1002/9781118799475.

[30] "OFDM Based Relay Systems For Future Wireless Communications," River Publishers, 2012, http://http://www.riverpublishers.com/book_details.php?book_id=117.

[31] B.Chatfield and D.Rawat, "Detecting synchronization signal jamming attacks for cybersecurity in cyber-physical energy grid systems," IGI Global Inc., November 2016, http:
//www.igi-global.com/book/security-solutions-applied-cryptography-smart/166368.

[32] S. Clements and H. Kirkham, "Cyber-security considerations for the smart grid," in *IEEE PES General Meeting*, July 2010, pp. 1–5.

[33] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Spoofing detection in ieee 802.15.4 networks based on received signal strength," *Ad Hoc Netw.*, vol. 11, no. 8, pp. 2648–2660, Nov. 2013.

[34] S. Misra, A. Ghosh, A. P. S. P., and M. S. Obaidat, "Detection of identity-based attacks in wireless sensor networks using signalprints," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*, Dec 2010, pp. 35–41.

[35] P. Jokar, H. Nicanfar, and V. C. M. Leung, "Specification-based intrusion detection for home area networks in smart grids," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct 2011, pp. 208–213.

[36] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Spoofing prevention using received signal strength for zigbee-based home area networks," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct 2013, pp. 438–443.

[37] P. Jokar and V. Leung, "Intrusion detection and prevention for zigbee-based home area networks in smart grids," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016.

[38] A. Atassi, N. Sayegh, I. Elhajj, A. Chehab, and A. Kayssi, "Malicious node detection in wireless sensor networks," in *27th International Conference on Advanced Information Networking and Applications Workshops*, March 2013, pp. 456–461.

[39] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, "Detecting and localizing identity-based attacks in wireless and sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2418–2434, Jun 2010.

[40] R. Maivizhi and S. Matilda, "Distance based detection and localization of multiple spoofing attackers for wireless networks," in *International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, April 2014, pp. 63–67.

[41] B. Chatfield and R. J. Haddad, "Rssi-based spoofing detection in smart grid 802.11 home area networks," in *IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, April 2017, pp. 1–5.

[42] "Moving Target Defense," Department of Homeland Security- Website, 2011, https://www.dhs.gov/science-and-technology/csd-mtd.

[43] M. Dunlop, S. Groat, W. Urbanski, R. Marchany, and J. Tront, "Mt6d: A moving target ipv6 defense," in *2011 - MILCOM 2011 Military Communications Conference*, Nov 2011, pp. 1321–1326.

[44] O. Hardman, S. Groat, R. Marchany, and J. Tront, "Optimizing a network layer moving target defense for specific system architectures," in *Architectures for Networking and Communications Systems*, Oct 2013, pp. 117–118.

[45] S. Yan, X. Huang, M. Ma, P. Zhang, and Y. Ma, "A novel efficient address mutation scheme for ipv6 networks," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2016.

[46] B. Chatfield and R. J. Haddad, "Moving target defense intrusion detection system for ipv6 based advanced metering infrastructure," in *SoutheastCon 2017*, April 2017, pp. 1–7.