Distributed intrusion detection/prevention system design and implementation for secure SCADA communication in smart grid

by

Sathya Narayana Mohan

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Computer Engineering

Program of Study Committee:
Manimaran Govindarasu, Major Professor
Doug Jacobson
Venkataramana Ajjarapu

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

DEDICATION

I dedicate this research work to my family. I feel grateful to my loving parents, ***Mohan*** *and* ***Gnana sundari****, whose support and encouragement gave me the strength to progress through this journey. I would like to express my dedication to my close friends who have guided and supported me throughout the process.*

Finally, with all reverence, I also dedicate this work to all the hard working and respected Teachers.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis.

First, I would like to thank my Professor and Advisor, **Manimaran Govindarasu** for his guidance, patience and support throughout this research and the writing of this thesis. His insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education.

I would like to thank my committee members, Professor **Doug Jacobson** and Professor **Venkataramana Ajjarapu** for taking time to respond to my emails and agreeing to be a part of my POS committee.

I would like to thank my research associate Dr. **Ravikumar Gelli** for providing valuable feedback during the research meetings and his guidance throughout the initial stages of my graduate career.

ABSTRACT

Cybersecurity, one of the expanding research area, has tremendous importance towards critical infrastructures. Organizations like power, oil, and gas use SCADA communication to manage and control their outstations across wide area. Some of the standard SCADA protocols used are DNP3, Modbus, IEC 61850 to control, share, and exchange real-time information. The communication involves both cyber-physical system processes and requires high availability and integrity of the data. DNP3, a TCP based protocol, is widely used in these infrastructures. With the involvement of the cyber, the systems are susceptible to network-based intrusions and cyber attacks. Since the communication is between the control center and its vast network of outstations, it becomes a challenge to monitor and control the network activity of the whole system. It creates a demand in the visualization of different network areas and a need to monitor their network activity from a single console. This work presents a framework to bring the distributed setup of the intrusion detection system and provide an optimal solution to detect network intrusions and abnormal behavior. The main focus of the work is to provide a single dashboard view to monitor the network activities of different outstations.

Further, the design and implementation of the distributed setup are explained in various architectures. Different types of IDS rules based on packet payload, packet flow, and time threshold are generated to show how an attack surface of the system can be reduced and detect different types of cyber attacks. Then, IDS testing and evaluation is performed with a set of rules in different sequences. The detection time is measured for different IDS rules and the results are plotted. All the experiments are conducted in Power Cyber Lab, ISU using two-area and 39-Bus power model and presented in CPS and Grid-Ex based training. After successful testing and evaluation, the knowledge and implementation are transferred to field deployment. In the last section, the conclusion of the work is summarized and a possible extension of future work is discussed.

CHAPTER 1. INTRODUCTION

With the advent of modern technologies, internet communication has become seamlessly efficient and reliable, but simultaneously the cyber attacks are getting more advanced and sophisticated. In this backdrop, the network security of an organization turn as a primary goal to protect. In the infrastructures like power and energy, which has both physical and cyber layer are integrated with the legacy systems. These industrial control systems exchange control and operational data between the control and production networks frequently. Modern society depends hugely on electricity and any significant impact on the critical infrastructure may disturb daily life. The demand for electric power is increasing every day particularly in urban districts [1]. A stealthy coordinated cyber attack like 2015 Ukraine blackout can result in power outages affecting around 225,000 customers [2]. Malwares like Stuxnet infected the Industrial controls systems made by Siemens creating a disturbance of the processes controlling the centrifuges of Iran's nuclear reactor [3]. Availability and integrity of the data becomes one of the primary concerns in industrial communication. It creates a demand to achieve highly secure and reliable communication. Therefore, the need for efficient management of the power and reliable supply to the consumer is needed. Due to increasing load demands, the power distribution management needs to improve efficient and reliable power supply.

Smart grid which is slowly evolving, profoundly relies on ethernet-based communication protocols like DNP3. The power grid can be divided into three stages, like generation, transmission, and distribution. As shown in the Figure 1.1, the power grid is expanding the need to provide cost-effective, and reliable power becomes challenging. At this juncture, the smart grid evolves as a modernized solution. The central aspect of the smart grid is to create decentralized power distribution where the consumer plays a small intrinsic role towards power contribution locally [4]. To make cost-effective and energy-efficient, consumers today are installing solar energy panels

which helps not only in lowering the utility bills as well as the load demand is decreased. In some cases, solar farms are built, which helps in a small-scale amount of power production and feeding back to the grid as commerce. With the use of the latest smart energy meters it is possible to use the power resources more efficiently. Smart energy meters provide real-time data of consumer load consumption to the regional control center, which helps in balancing the power loads of different areas and reducing power outages [5]. Smart energy meters integrated with solar panels offer detailed feedback of the energy to the grid through incentives from the utilities. The network data exchanged through SCADA communication has high importance to availability, integrity, and confidentiality. In this order, availability is extremely critical as all the process that is being managed and control is in real-time. In traditional security, confidentiality and integrity of the message is given more importance. However, in critical communications availability of the data packets are given the primary importance [6]. Then follows the integrity; where the data exchanged between the server and the clients should be maintained highly integrated. Any slight compromise in the integrity of the data received may lead to disastrous events. Confidentiality has the lowest priority compared to other two. The data communication in most of the SCADA environment is not encrypted, since encryption creates an additional overhead time that hinders the availability of critical communication.

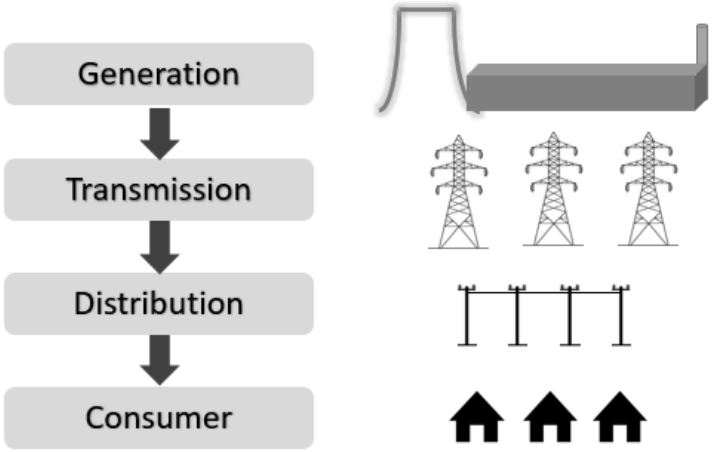

Figure 1.1 Power Transmission

The current state-of-the-art in intrusion detection and prevention system is presented in the upcoming section. The main focus of this work is to provide a distributed IDS/IPS to decrease the attack surface of the system by using various types of IDS rules. It describes the depth of the distributed intrusion and prevention system architecture, the need for secure SCADA communication and explains in detail about the network nodes and operation. Also in the evaluation section, how the order of IDS rules makes the detection time faster is shown. Then, several types of IDS rules are generated based on the given scope of cyberattacks. Chapter 3 gives an overview of SCADA protocols like DNP3 and Cyberattack classification and describes the IDS rule generation based on various types to deter malicious cyber activity.

1.1 Motivation for Secure SCADA Communication

Modern organizations use internet in different ways; their first-order security comes as firewall, IDS, IPS, security incident and event management (SIEM) and data analytics. The use of an advanced firewall which has an application layer filtering can evade most of the cyber attacks. On the other hand, cyber attacks are developing more stealthy. Intrusion detection and prevention systems like rule-based, analysis-based, and behavior-based help in detecting real-time attackers. There is also much demand in monitoring the network activity to analyze the network behavior of the system which has anomalous behaviors. Also, the bandwidth and network usage statistics are continuously monitored to find out any abnormal system behavior. When an organization has multiple areas, the operations face a challenging task to manage its network of clients. The top concern remains to protect their official data across multiple areas from a single standpoint, which leads to a distributed intrusion detection and prevention system. The security standardization has a different perspective view for information technology (IT), and operational technology (OT) environments as the risks are different in each other domains. The entire system process is dependent on both Cyber and Physical process, any discrepancies in time between the Sever and the client can create disastrous effects. Since OT systems use sensors to regulate critical processes and are related to legacy systems, the priority for availability is highest rather than integrity and confidentiality.

1.1.1 Background

Many operational networks are vulnerable to cyber-attack vectors due to a lack of in-depth security. The information security priority for the operational technology and Information technology is shown at Figure 1.2. It is not a practical approach to apply IT security measures to the Operational environment since availability is more concerned for Industrial control systems. The network of SCADA has a master server to connect to its client at outstations to perform real-time operations without any compromise in security, and are connected to provide mechanical stability. In this communication that uses a protocol like DNP3, Modbus, IEC 81560: the attacker can easily exploit the protocol layer fields like function codes, data objects, headers, and payload by observing stealthily. Distributed Network Protocol (DNP3) a widely used SCADA protocol has function codes present in DNP3 application layer that are exploitable by an adversary. There are 28 types of attacks related to DNP3 protocol exploiting the function codes and other data fields in the data-link and DNP3 application layer [7]. These protocol fields can create more vulnerabilities which can lead to attacks like man-in-the-middle, denial-of-service and data integrity attacks. There emerges a variety of threats conducted by various threat actors or activities supported by illegal groups against critical sector operated by SCADA. At this point, it is highly essential to provide robust technologies to prevent malicious activity and secure SCADA communications. Stealthy cyber-attacks like IP spoofing can manipulate the whole network header of the packet and can bypass the existing security measures. Firewall and Intrusion detection system lacks the scope of detecting IP-based spoofing attacks. These incidents and attacks happening at the DNP3 protocol is expected to increase because of the advanced development and besides the use of legacy SCADA systems. The use of legacy infrastructure paves ample space for the attackers to exploit using modern tools and technologies.

1.1.2 Limitations due to Legacy Infrastructure

The traditional architecture used to build critical systems are difficult to upgrade to the current technologies. Since many limitations are needed to overcome if the infrastructure is going to trans-

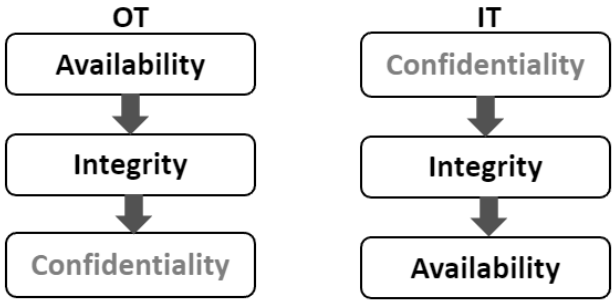

Figure 1.2 Security triad for OT versus IT

form, it requires a massive effort and difference in changing the technology used in the legacy and the current technologies. The operational technology consists of hardware and software processes that are dependent on each other. Any difference between these processes can deliver unreliable outputs. This connection creates a gap between existing and modern technology. Most of the systems, like process control domains (PCD), Management and distributed systems, programmable logic controllers(PLC), SCADA are dependent on each other. Standard DNP3 protocol lacks encryption and authentication, which gives attackers to perform sniffing and reconnaissance. The vulnerability of standard DNP3 protocol creates different types of cyberattacks like Man-in-the-middle and Data integrity attacks [8]. Many SCADA operators still use standard DNP3 version, although the latest DNP3-SA (secure authentication) is still available. DNP3-SA provides an authentication mechanism used to ensure that messages are exchanged between an authenticated Master and Client. The traditional operational networks are designed for specific security properties to comply with the availability, integrity, and confidentiality. The use of DNP3 in SCADA communications has security issues, and most existing DNP3 communication lacks authentication, data encryption, and access control. While enhanced versions of DNP3 like DNP3 secure authentication, DNP3sec have been reaching the industry slowly but still the majority of the SCADA is working with the legacy systems and protocol.

1.2 State-of-the-art of Cyber-Security in SCADA

1.2.1 Related work

DNP3, a protocol designed for distributed communications has some limitations. In order to identify the appropriate requirements for generating various IDS rule sets, the DNP3 packet structure and the protocol has to be studied. The work in [9] shows a list of attacks that are exploited in the SCADA environment. It shows that there are many data fields in the DNP3 layer like data-link, transport, and application which can be exploited. Also, in [10] shows the vulnerability of the protocol stack and how it can impact the secure communication. Using this information, we used to test the developed IDS rules at the Power cyber lab, Iowa. Therefore, we explored the types of attacks by exploiting the function code and examined the system analysis. The attack data sets from [11] show a wide variety to explore the exploitable features of the protocol. The work in [12] proposes a new type of IDS, where it dynamically modifies the firewall rules according to the IDS. It introduces a dynamic IDS, which unites with the firewall to prevent various DNP3 attacks. The work in [13] displays an intrusion detection framework based on analysis-driven IDS. It shows how the analysis of the traffic in SCADA can be used in the intrusion framework. In [14] proposes a game theory approach in finding the vulnerability of the DNP3 and showcase a model to define it. Also, it describes the DNP3 attack, which can be detected using advanced techniques used in game theory. The latest development in the DNP3 protocol is DNP3 secure authentication [15], which includes end-to-end cryptographic authentication. It provides more authentication in the application layer of the server and client, which makes it very hard to impersonate. Also, in [16] provides a framework for communication with confidentiality, integrity, and authenticity. The smart grid is a network of the different terminal device; the work in [17] uses this framework to propose a dendritic cell algorithm. The work focuses on the packet header that is exchanged between the different nodes to detect the anomalous traffic.

1.3 Scope of Research

1.3.1 Reduction of Attack Surface

Attack surface is the entry point for the cyber-attacks. An Adversary‘s main aim is to exploit these attack surface and create a cyber-attack. By reducing the attack surfaces, we can able to reduce vulnerability risk of the network. Since the organization contains many technology resources, from the point of security, we need to minimize the attack surfaces of the cyber resource we have. When these Attack surface are exploited using any technology or tools is called as Attack Vector. Attack surface is of three significant types, namely Software attack surface, Hardware attack surface, and Human attack surface. When a network is exposed, the vulnerability is exposed, then an attack surface of an organization increases. Finding an attack surface is quite challenging since attack surfaces are changing every day of its dynamic nature. There is a various number of vectors that lead to types of attack surfaces like open ports, weak firewall rules, web pages, weak Access list rules. The upcoming chapters show how attack surfaces are identified and decreased using different techniques.

However, before entering into the Classification of Attack surface, let us describe what an attack surface is?. An Attack Surface of a network can be defined as all possible endpoints which are having access directly or indirectly to the Internet without any security checkpoint. An Attack Surface is a medium for attackers to exploit using different types of attack vectors. Any Cyber resources that correspond with the Internet, directly or indirectly tend to be a form of Attack surface. An Attack surface can be viewed from two different perspectives as an attacker and defender. From the point of an attacker, it is more interested in the exposure of the systems and applications used inside the private network, whereas from the point of the defender, it is more concerned with the protection of the systems and applications. Hence, for a secure private network, learning an attack surface needs both attacker and defender point of view. In this work, we consider the software process running obscure, and the open ports which are listening tends to be the attack surface of the system. By

having this information, robust IDS and IPS rules are designed to decrease the attack surface of the system and for the network.

1.3.2 Distributed Intrusion Detection and Prevention System

Many technologies have been employed to combat unwanted cyber activities. An IDS in a network serves as a device that continuously monitors the network traffic events. IDS also analyze the system's behavior and protects critical systems against the malicious activity of the attackers. The primary focus of this work is to provide distributed intrusion detection and prevention system in a network to filter and log the inbound traffic and outbound traffic. This advanced functionality provides threat signature detection, virus signature detection, protocol, and port filtering, and includes network and application layer level protection. Unlike traditional firewalls, IDS with IPS has greater visibility and control in real-time.

There are different types of IDS like anomaly-based IDS and the signature-based IDS. This work focuses on signature-based IDS, which uses a set of signatures to identify any possible malicious activities. The structure of the IDS can be of two types centralized or distributed. Centralized IDS acts as a standalone system with no interaction with its clients. In the distributed setup, the network consists of multiple IDS clients connecting to the Master IDS. Distributed network monitoring allows an overall view of the network activity from the Master console. It provides the network administrator the flexibility to improve network monitoring and control various networks and take preventive measures in real-time. With this setup, the network admin can tremendously reduce the attack surface and increase the visibility and control of multiple network areas. However, the firewalls at the gateway of SCADA communication are whitelisted according to IP addresses, applications, websites, users, processes, devices to limit unwanted access. Also, the firewalls blacklist the rest, where any suspicious activity found on any application, user, IP addresses, websites can be blocked immediately. Next-Generation Firewall has advanced features like whitelisting and blacklisting in the level of physical-MAC addresses, network, and application layer. In this topology, all the Endpoints are considered as Client, and the Master controls all clients. The Master is

located in the Control center has an accessible server or gateway within the network, which installs a client software having security policies, IDS rules, and other features on each of the endpoints.

The Highlight of this work is to provide Network administrator complete control over the network of different areas in a single management dashboard. It mainly aims to secure every endpoint by applying policies to block unwanted access attempts and other risky activity at these points of entry. It can maintain greater control of the network access points and more effectively block threats. It also provides extended features such as monitoring and blocking risky or malicious activities

In chapters two and three, we discuss briefly on how SCADA traffic is studied. An IDS rule generation based on traffic behavior and the functionality of the relays is introduced. The network communication involving critical commands have a particular time sequence, which can be called as the behavior of the system is analyzed. Similarly, the protocol used in this work is DNP3, which has around thirty-five different functions are studied in our experiments. With all the above considerations, an algorithm is devised to show the generation of IDS rules. After the rule generation, the system is made to have the initial operations. Then it is conducted with various cyberattacks to check the performance of the IDS rule generation algorithm. It is validated by observing whether the IDS ruleset can detect numerous cyberattacks within the given scope. After this deployment and testing, the ruleset order is scrutinized to check the most suitable sequence to get the minimum detection time. Two types of rule order are used to test the minimum detection time, are then compared with the graphical terms, and the results are given in chapter 4. Finally, a valid rule sequence order with a rule generation algorithm within a given scope is proposed and evaluated with enough results.

1.4 Organization of Research work

The Figure 1.3 shows the roadmap of this work from July 2018 till July 2019. Most of the work in the initial phase was laid on the understanding and construction of the thoughts that carried us here. Attack surface Host Analyzer, a tool developed by Washington state university to

show how attack surfaces are connected with hidden processes and listening ports. After learning the different types of attack surfaces, there was an opportunity to find a possible way to reduce it. In October 2018, we demonstrated the standalone IDS testing and showed several IDS rules and reduction of the attack surface. After initial testing, there was a demand in distributed setup since the substations are located across different areas. In December the work focussed on the distributed configuration and developing advanced DNP3 rules based on traffic patterns. The testing of the advanced rules was done at the power cyber lab, Iowa. Phase I was started in the month of March; where Master and Client-1 is deployed at the control center and substation-1 network. Both systems were tested and are up in the production network. In May 2019, this work played a primary part in CPS training for the power utilities. Also, the work was presented in conferences like Electric Power Research Center (EPRC), Graduate and Professional Student Research Conference (GPSRC), Florida Reliability Coordinating Council (FRCC). More details regarding conferences and outreach are discussed in Chapter 5. Later, the evaluation of the IDS is performed in May 2019. Different rule order sequence was conducted to find the least detection time, and the results are shown in Chapter 4. In July 2019, Client-2 is deployed with a robust rule sequence and advanced DNP3 rules. Finally, the work is documented and paper regarding the rule generation and rule order sequence is published with the title "Distributed Intrusion Detection System using Semantic-based Rules for SCADA in Smart Grid". The results have a graphical picture for the rule order sequence to achieve the least detection time are presented.

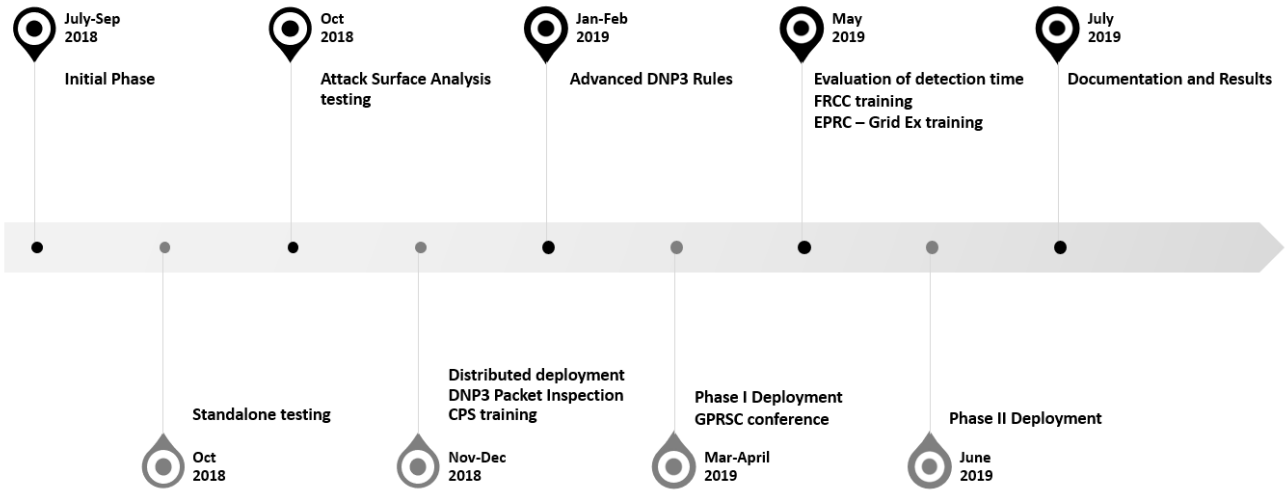

Figure 1.3 Organization of Research work

CHAPTER 2. DISTRIBUTED IDS DESIGN FOR SCADA COMMUNICATION

2.1 Introduction

The evolution of the power grid has slowly transformed into a smart grid for two decades. The industrial control system operation can be summarised as data communication between the sensors, PLCs, and other devices in the field on a real-time basis. ICS plays primary importance in controlling field hardware equipment and systems. In industrial automation, controllers and advanced sensors need the synchronization. SCADA Smart grid relies extensively on the poll and response communication of critical controls and data objects. In the existing background of cyber threats, attacks related to these infrastructures are increasing due to the use of legacy infrastructure and the vulnerabilities of the traditional architectures and protocols. Communication protocols used in the Industrial control system (ICS) such as DNP3, Modbus, IEC 61850 have existing vulnerabilities. The communication that happens between the different networks in the grid demands a distributed Intrusion detection system. Gaining an understanding of network activities are required in real-time to protect from the cyber attacks. The requirement for a detection system that communicates to a central node from all the networks is much expected. This chapter discovers the elements of a distributed IDS with different architecture [18]. Also, the Intrusion prevention system is built using In-line architecture. At the end of this chapter, the attack surface of a system is mapped using a tool called Attack surface Host Analyser, [19] a tool developed by Washington State University. Then using developed IDS rules, we demonstrate how to reduce the attack surface utilizing tool information.

2.1.1 Primary Features

The use of distributed intrusion detection systems gives the flexibility for the cyber network operator to monitor and control the network of its outstations. It eases the operator through providing a single console view of its entire grid network; in this way, the operator can potentially leverage this feature in the smart grid. Distributed IDS architecture, includes one Master node and many sensors. All forward nodes are connected to the master node. The Client has a sensor installed that operates in promiscuous mode to collect the network information and forwards to the master. In Standalone operation, the master node has its database server. Whereas in Distributed deployment has two architecture. With or without the storage node deployment. In distributed storage node deployment, the master maintains a storage node to forward the network data information for future queries. Storage of these data analytics are useful for analyzing the system states, and behaviors of the process to identify any abnormal behaviors. Advanced machine learning uses data analytics to train a proper model for IDS. Another feature of using this distributed IDS is the customized rule sets. The IDS rules are designed based on the traffic pattern, network packets, packet flow, packet content and the packet threshold time; this is one of the most significant features that can be tailored for different types of environment. This potential feature gives enormous flexibility for the Cybersecurity expert to analyze and monitor the network data from a single standpoint. Also, the operator can leverage the flexibility of writing different types of IDS rules that suits their network environment.

2.1.2 Network Nodes and Operation

In the distribution setup, we consider different types of nodes according to the respective functions. The nodes are Linux distribution endpoints. The nodes are categorized into Master Node, Sensor Node, Storage Node, Forward node, Heavy node.

Master Node: In standalone deployment, the master node functions the same as the sensor node. In a distributed deployment, we have a Master node and several sensor nodes. All Sensor nodes like Forward node, Storage node, and Heavy nodes are connected to the Master node. The

Master node controls the operation of its distributed network. All the Sensor nodes have a secure shell (SSH) connection to the Master node for communication. Any IDS rule update, network security updates, host intrusion updates done in the Master, is then easily pushed to the Sensor node. It provides flexibility for the network administrator to push the updates for the different clients from the single node. The feature of monitoring and controlling the updates of the IDS makes the distributed setup more comfortable to handle. The Sensor node in return provides the network activities which monitor through promiscuous mode and sends back to the Master, which then transfers the information to the storage node.

Sensor node: The sensor node can be of Forward node, Storage or Heavy node. The sensor node is Linux distribution Ubuntu-16.04 operating system. All sensor nodes are connected to the Master node and are deployed only after installing the Master node. The network configuration of the sensor node contains two or more interfaces. The primary interface is for the management IP address of the Client machine. The management interface is used to connect to the Master via SSH connection. The secondary interface is used for the sniffing of the Client network. There can be one or more sniffing interface depending upon the requirement of the network which is used to monitor the network traffic.

Forward node: In this node, the sensor assimilates and forwards all the network-related logs and information to the Master. Forward node contains network packet capture detection, rule detection engine, and analysis detection engine. The forward node updates the IDS rules and security updates and compares its network traffic using the network packet capture. After the network packet capture, it then checks the rule signature using the rule detection engine to check any pattern is matching. Once any of the IDS rule signatures match the network packet, immediately the detection engine sends an alert to the Master. The alert is then able to visualize at the dashboard of the Master node.

Storage Node: Storage node is an additional node connected to the Master, where all the network traffic logs are forwarded from the sensor to the Master, the Master then forwards the logs to the storage node. After the sensor node captures all network log and information, storing

this information requires huge memory, which is a challenge for the Master node, and doing this in the Master may reduce the computing time in intrusion alerts and lowers the performance. Hence storage node is employed. This node operates as a database server when the Master node queries about any information; it takes from the storage node. The storage node stores all the logs and are used as data analytics for advanced statistics and analyze the different network behavior [20]. The master node uses the storage node for the data that can be queried through the use of cross-cluster search.

2.2 Deployment Architecture

2.2.1 Standalone Deployment

The standalone deployment comprises of a single node, which combines the functions of the Master Node, forward node, and storage node. This type of implementation is used if the network area is limited. It is used locally to manage and monitor the network for testing. This type of deployment is used for testing in the labs and for evaluation. This implementation can be seen in Figure 2.1. In this type of deployment, the Master is employed as both the sensor and the manager. In this node, a single interface can operate as both the management and the sniffing. This type of deployment is useful for organizations having limited network areas.

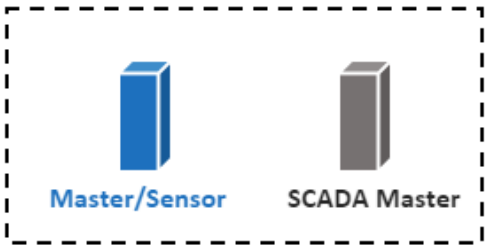

Figure 2.1 Standalone deployment

2.2.2 Distributed Deployment

The Distributed deployment comprises of one Master node, one or more forward nodes, one or more storage nodes. This architecture is widely deployed in the industry as production deployment compared to the standalone deployment, as it provides more scalability, performance and also handles the heavy network traffic and log management. This implementation is shown in Figure 2.2. It is highly recommended to use the distributed deployment for the production network, and the use of the storage node provides more extended options to learn about the data analytics and elastic search.

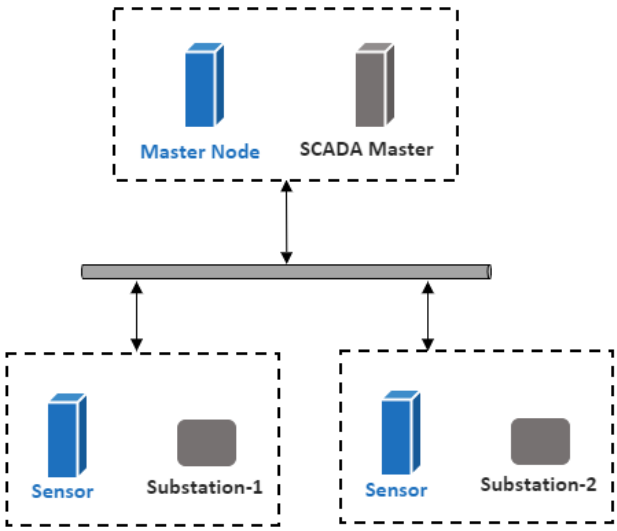

Figure 2.2 Distributed deployment

Another type of implementation called Heavily distributed deployment is used when there is a Storage node in the network. The Master node uses an additional node as storage to store a vast amount of the network logs. The heavily distributed deployment consists of a Master node, one or more heavy nodes. Heavy nodes are the forward node, which has both the functionality of the sensor and elastic search. This implementation is seen in the Figure 2.3.

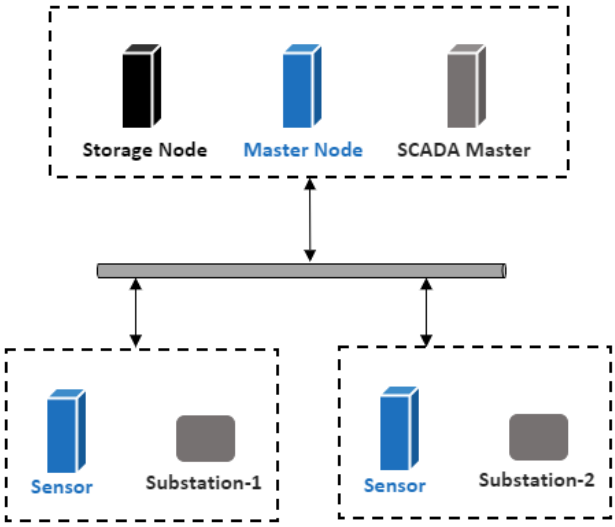

Figure 2.3 Heavily Distributed deployment

2.2.3 In-line IPS Deployment

The main aim of the IDS is to detect the malicious network activity and alert the admin. The IDS uses the signatures to match the network packets for finding an alert. When this IDS is placed at the gateway of the network or the in-line to the device, it can now operate as a prevention system. The intrusion prevention system is able to detect and defend against various types of cyberattacks like Denial of service. This type of implementation is one of the advantages for the administrator as it provides the functionality of the detection and prevention of malicious activities. The Figure 2.4 shows the IPS deployment architecture.

2.3 Design algorithm

In this section, we discuss IDS rule generation, as shown in the Figure 2.5. The primary step in IDS rule generation is to study the network traffic of the environment. In this case, we consider the DNP3 traffic to analyze different commands that are exchanged between the control center and substations. The communication of DNP3 commands follows a particular pattern that depends on the utilities network and conditions. The algorithm starts from regular SCADA traffic to pattern-

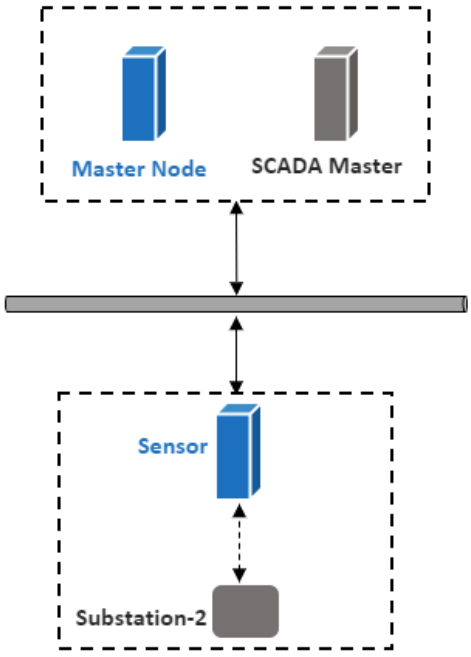

Figure 2.4 In-line IPS Deployment

1 and pattern-2. The traffic pattern can be based on the protocol, services, critical commands, packet payload, content, time threshold. In this consideration, a list of traffic patterns is defined. After having the traffic pattern, the IDS rules are tailored to detect a particular incident, which is discussed in chapter 3. In the Figure 2.5, it shows that when a new pattern is detected, a new IDS rule is generated else, it is resolved to the Normal operation and the Figure 2.6 shows properties to consider for designing robust IDS. After the new IDS rule is generated, the rule is sent to the repository, and finally, it is updated with the regular network operation. This brings a robust adaptive approach to detect different types of cyber incidents in the network. In this work, we define different sets of IDS rule based on various traffic patterns, which are discussed in Chapter 3.

2.3.1 Flow chart for rule generation

The general SCADA traffic from the server to substation to the endpoint is shown in the Figure 2.7. In this picture, we can see that each substation is having many Relays, and each relay is controlling many endpoints like generator, power transformer, circuit breaker. The traffic

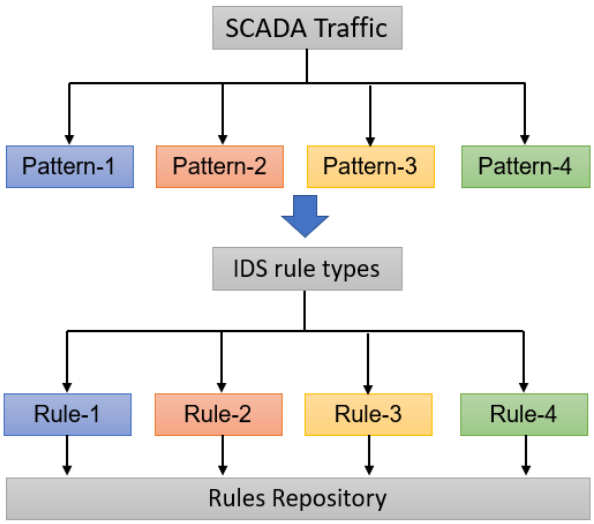

Figure 2.5 IDS Rule generation algorithm

pattern for each end device is different. Table 2.1 and Table 2.2 shows the traffic route from the SCADA server to the end devices. Now, each relay has different endpoints and are controlled by different function codes. The network traffic is different for each end-device for different relays and substations. This creates different traffic patterns based on the station IDs, relay IDs, and device addresses. The Figure 2.8 shows IDS rule design flow, and Figure 2.9 shows the functions to consider for creating robust IDS rules for various outstations.

Table 2.1 Network traffic pattern in SCADA for Substation-1 Relay 1

Pattern	**Network traffic**
Pattern-1	SCADA server - SS1 - Relay1 - GEN
Pattern-2	SCADA server - SS1 - Relay1 - TX
Pattern-3	SCADA server - SS1 - Relay1 - CB

Table 2.2 Network traffic pattern in SCADA for Substation-1 Relay 2

Pattern	**Network traffic**
Pattern-4	SCADA server - SS1 - Relay2 - GEN
Pattern-5	SCADA server - SS1 - Relay2 - TX
Pattern-6	SCADA server - SS1 - Relay2 - CB

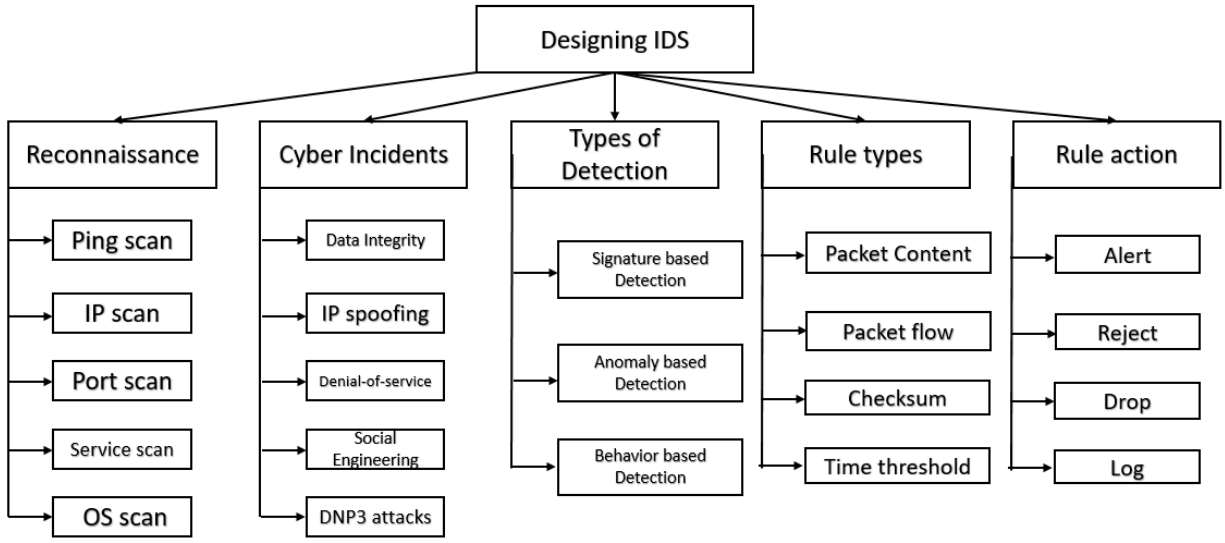

Figure 2.6 Proposed Multistage Properties for the Robust IDS Design

2.4 Attack Surface Analysis

In this section, we are using a tool called Attack surface Host Analyzer, a tool developed by Washington state university to check the attack surface such as open ports, hidden processes running in the machine. The idea is to show the attack surface of the system by showing the open ports and then by using the help of the IDS and in-line IPS, we try to reduce the attack surface of the system. Then using the same tool, we rerun the analysis to show the reduction of the attack surface. The Figure 2.10 shows the attack surface of the system before the IDS and in-line IPS deployment. As we see in this figure, many hidden processes are still listening to ports that are open. These are the established connections before and now are in listening state. These open listening ports creates a path for the attacker to learn about the system and the processes running in the respective ports. Then after deploying the In-line IPS with the customized rules. Figure 2.11 shows that most of the open listening ports are blocked and the hidden processes are no more in connection other than the expected address. The IDS rule for designing the system are discussed in the next chapter.

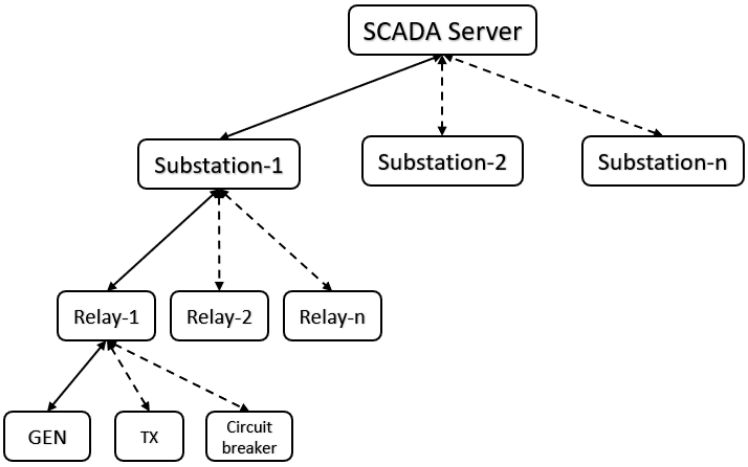

Figure 2.7 Traffic pattern in SCADA

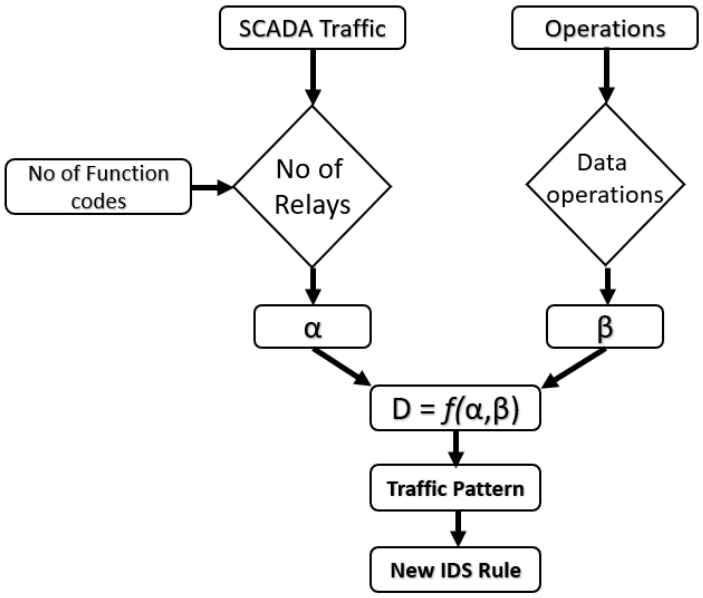

Figure 2.8 Design flow of IDS

$$D = func\ (\alpha, \beta)$$
$$\alpha = (No\ of\ Relays) * (No\ of\ functions\ per\ relay)$$
$$\beta = Behavior\ of\ SCADA\ traffic\ and\ Data\ operations$$

Figure 2.9 Design function of IDS

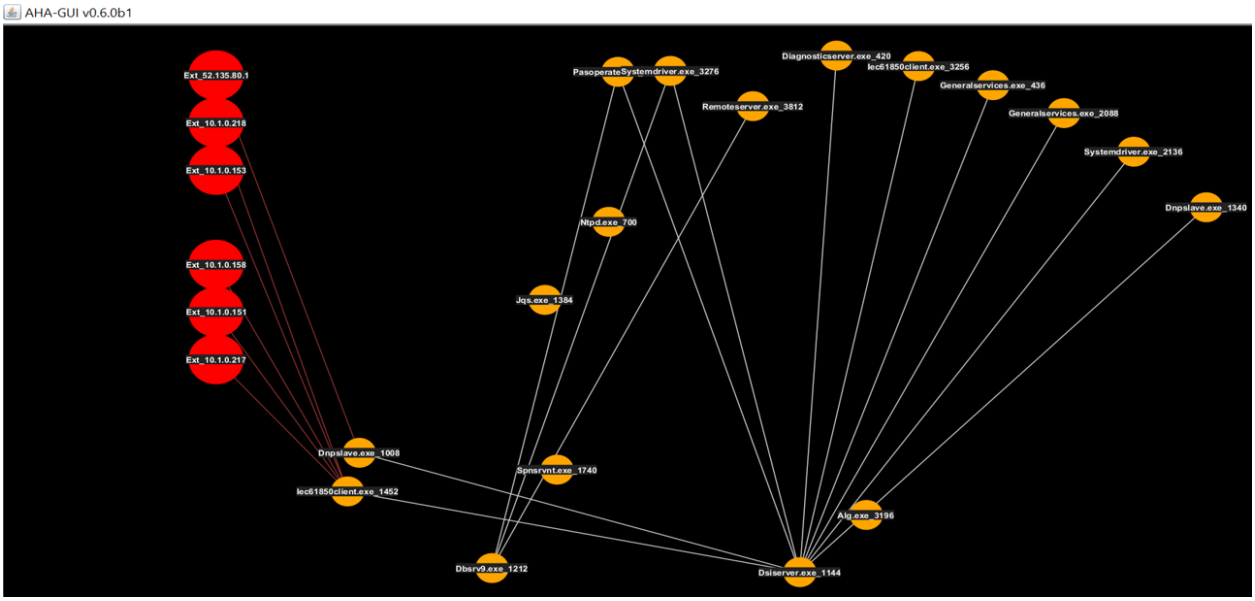

Figure 2.10 Before IDS/IPS Deployment

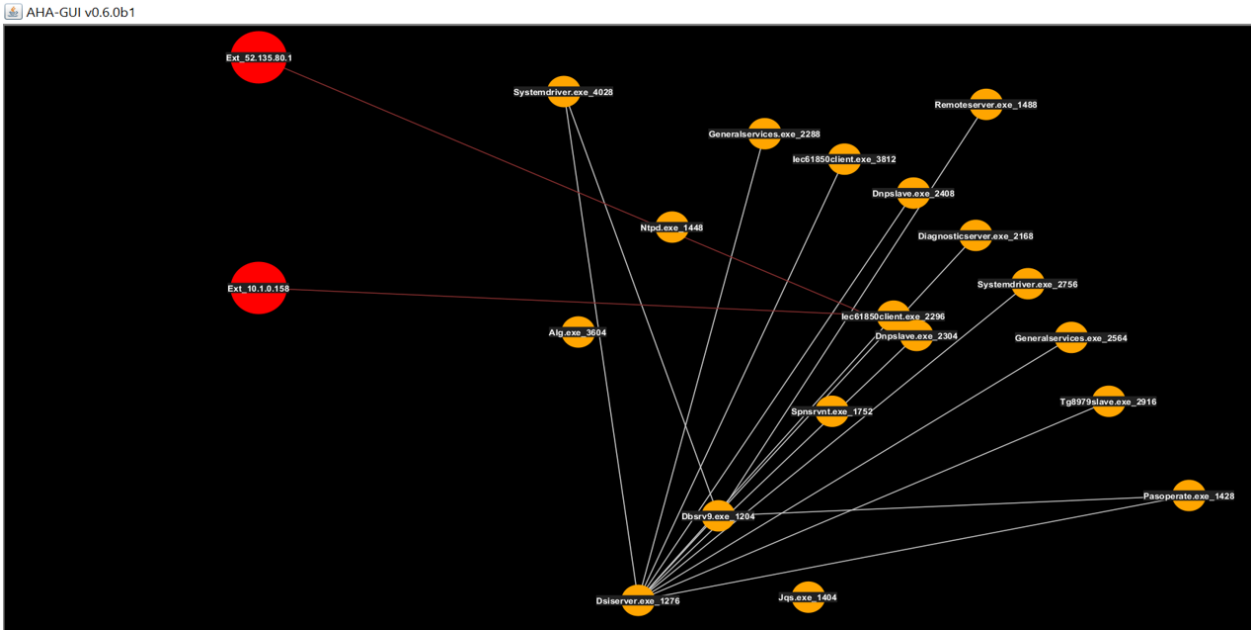

Figure 2.11 After In-line IPS

CHAPTER 3. CYBER-ATTACK CLASSIFICATION AND IDS RULE GENERATION

The primary step of any cyber attack is reconnaissance. Before launching any attack, the attacker first tries to learn, analyze, and understand the network, then plot the vulnerable target and direct the attack. During the period of reconnaissance, the process of discovering hosts, IP address, port scanning, identifying the OS and services used in a network is initiated. It depends on the level of exposure due to insecure security device configurations and firewall access rules, which leads to the exposure of vulnerability. It exposes the organization's network and hardware surface, which is exploitable by the attackers. Once the attacker gets the ground information of the target, then they proceed to identify the vulnerability of that system and find the ways to exploit it. Finding these vulnerabilities inside the network is a crucial step in determining a cyber attack. The cyber attacks are basically categorized into four types reconnaissance, with spoofing, without spoofing and flooding. All the cyber attacks fall into one of the four categories mentioned above. After scanning the entire network, the vulnerability report for each system in the network is collected. The vulnerabilities in the report are measured in terms of CVE (Common Vulnerabilities and Exposure) [21], which is a database for known information security vulnerabilities and exposures. The CVE database is a catalog of vulnerabilities with an identification number for each known vulnerability or exposure. The CVE is operated by MITRE corporation, to standardize the security vulnerability of software and hardware. These CVEs are sorted based on a measurement called CVSS (Common Vulnerability Scoring System), which is a numerical score provided by NIST [22], National Vulnerability Database, National cybersecurity division. CVSS are represented in High, Medium, and Low and are ranked based on the priority of the vulnerability. After we obtain the vulnerability report of all the assets in the network based on the CVSS ranking, a detailed analysis is done. From this report information, the attack surface of the network is classified

based on software attack surfaces, hardware attack surfaces, and physical/human attack surfaces. In the software level, any application program interfaces, storage files that are listening to the external network or the hidden process are vulnerable to the system. Similarly, in the hardware level, unused ports, and hardware device exposure, remote connection creates an entry for attack vectors. Employees unaware of cyber-related resources, social engineering attacks, and the latest technologies used for security are considered to be human attack surface. The below table 3.1 lists all the types of attack surfaces.

Table 3.1 Classification of Attack surface

Software attack surface	**Hardware attack surface**	**Human attack surface**
Unused API	Unused ports	Unaware of security policies
Storage files	Network interfaces	Unaware of social Engineering attacks
Databases	Operating systems	Human machine interface
unused directories	remote access	weak passwords
Social networking messages	Plug-in devices	P2P connection
Executable files like .exe	Hardware patching	

After examining the attack surface, the classification is done on the level of priorities for different networks. By now, we know that attack surfaces are narrowed down to three different areas. When any hardware related feature is exposed to the outside network, we try to obscure the level of system exposure to the outside network; this reduces the hardware attack surface. Any software and applications that are used in the critical network should be thoroughly tested and verified the genuinely before the deployment. When the software running code that has an exploitable vulnerability should be notified immediately. Similarly, software attack surface can be reduced when any vulnerable software feature in the level of coding and testing is identified. Human attack surface can be measured in terms of awareness of the employees in the organization. Hence, increasing the awareness of security and strict policies can bring down the human attack surface.

3.1 Analysis of Cyber-attacks in SCADA environment

From the work [23] it shows that 50% of the incidents are accidental, 30% are due to malware caused by social engineering techniques. 11% is the potential attackers, which can be a group or an organization having immoral intentions. 9% counts on the internal attackers, which can be related to likely or unaware of the act. An internal attacker can be caused due to human attack surfaces. Most of the incident is related to the cyber intrusions like ping scan, port scanning, OS detection, process and service scanning. The attacker's primary step is to perform the reconnaissance to gather all the information regarding the target system. The attacker then enumerates all the learned vulnerabilities of the target to execute the next action. Once the adversary learns the IP address, open ports, services running, it eases to launch an attack vector. The attacks based on IP address are denial-of-service, IP flooding, data integrity attacks, DNP3 flooding. Once the IP address of the system is exposed, it is bound to receive any network packet. Sometimes the whole IP layer of the network packet is crafted stealthily to appear as the outbound packet of the trusted network. This type of attack is termed as IP-based spoofing, which has a spoofed IP layer to bypass the firewall and IDS. IP spoofing is used to deliver a malicious payload compromising the security features of the host.

Table 3.2 Common types of Cyber attacks in SCADA

Type of Incident	**IDS Alerts**
Reconnaissance	ICMP Ping alert, TCP Scanning alert
Denial of service	DOS alert, Timing factor alert
Spoofing	Timing factor and connection not established
Replay attack	Sequence mismatch and timing factor alert
Data Integrity	Incorrect checksum alert

3.1.1 Cyber-attack Classification

The Figure 3.1 shows the scope of cyberattacks considered in this study. The incidents are categorized based on four types, as shown in the figure. Reconnaissance includes all kinds of system scanning like IP addresses, ports, services, operating systems. Denial-of-service is one of

the frequent attacks that occur for any network. It includes all types of packet flooding such as Ping-to-death, IP flooding, DNP3 flooding, Distributed Denial of service (DDOS), brute force attack. Other attacks are classified into two types like with or without spoofing. Attacks due to without spoofing like data integrity, replay attack, Man-in-the-middle, social engineering attack uses a random IP header of the packet to deliver the malicious payload to the target. However, the firewall and IPS can prevent these attacks. In spoofing based attacks, the attacker spoofs the whole IP header and tries to deliver the malicious payload to the target. Spoofing based attacks are difficult to prevent since the IP header looks the same as the IP header of the actual sender. Spoofing attacks can happen on the DNP3 protocol, where the attacker uses a spoofed IP header and add the DNP3 packet header on top. However, based on traffic analysis, we can detect spoofing based attacks. In the following section, there are IDS rules which can identify these spoofing attacks are discussed.

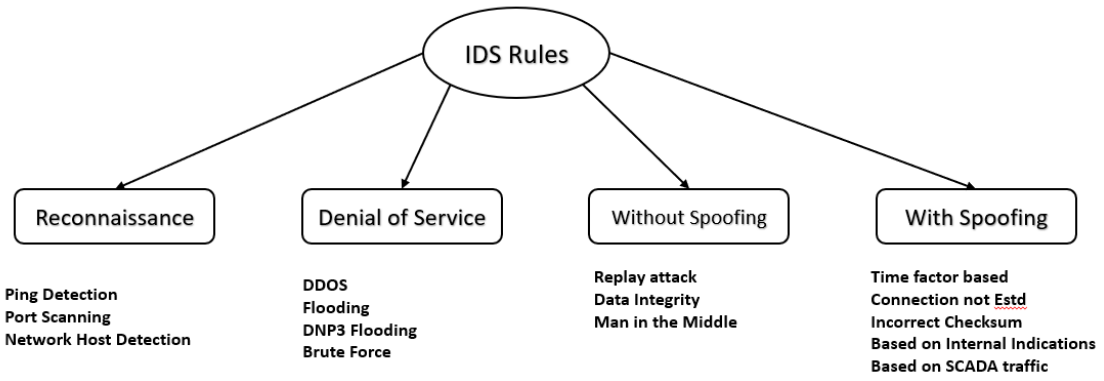

Figure 3.1 Cyber-attack Classification

3.1.2 Existing Vulnerabilities and threats

DNP3 an open standard communication protocol, where the features are available on the internet. Since all details are available, it becomes easy for an adversary to understand the existing vulnerabilities and the way to exploit it. As DNP3 communication is involved in a distributed architecture, the remote devices which are located in different areas will have remote network ac-

cess. This remote network access point creates an attack surface and thus can be countable on the vulnerability. There are many categories of threats and risks related to SCADA communication using DNP3 protocol [8]. The DNP3 protocol has three layers; the application layer contains the critical function code fields. There exist many threats and attacks when these fields are exploited. In the next section, we study the DNP3 protocol stack and discuss different types of threats and cyber-attacks related to it.

3.2 Network Protocols used in SCADA

SCADA Server needs a reliable communication protocol to control physical operations carried out at remote outstations. The system uses an application that involves the Programmable logic controller (PLC). These remote physical devices are monitored and controlled using critical commands. The basic SCADA architecture contains multiple communicating devices, as shown in Figure 2.7. This diagram shows that the relays are controlled by the Remote Terminal Unit (RTU), which sends the control signal. In some cases, we have Master Terminal Unit (MTU) which controls many RTUs and the following relays. The commonly used SCADA protocols are DNP3, Modbus, IEC 61850, Profibus. In this work, we are studying the DNP3 protocol, which is widely used in North America.

3.2.1 DNP3 Protocol stack

DNP3 a flexible, reliable, and open standard communication protocol that supports distributed communication to multiple devices at different outstation levels. It also supports multiple data types such as critical data, time-synchronized data, time-stamped data, data with priorities, data with responses, data with acknowledgments. It handles two-way communication, which can be a request-response or unsolicited response communication. It provides unicast, multicast, anycast, and broadcast communication, which gives the protocol to handle the distributed architecture seamlessly. The DNP3 protocol is a TCP based connection-oriented protocol. It has three different layers, as shown in the Figure 3.2. The main fields in the DNP3 link layer are the outstation source

and destination address. The control field includes function codes related to link status, request, reset, confirmation, ACK, NACK of user data; this builds synchronization in the communication. The transport control layer creates the sequence number for the packets in the communication. It is responsible for dividing the application layer messages and assigning the sequence number to provide reliable and uninterrupted transmission. Finally, in the DNP3 application layer, there are many critical fields such as application control, function code, internal indications, DNP3 objects. Among all these fields, the most commonly exploited field is the DNP3 application layer function codes. In this chapter, we are focusing on the exploitation and incidents related to DNP3 application layer function codes. Then IDS rules are designed to prevent these types of incidents, and their performance is evaluated.

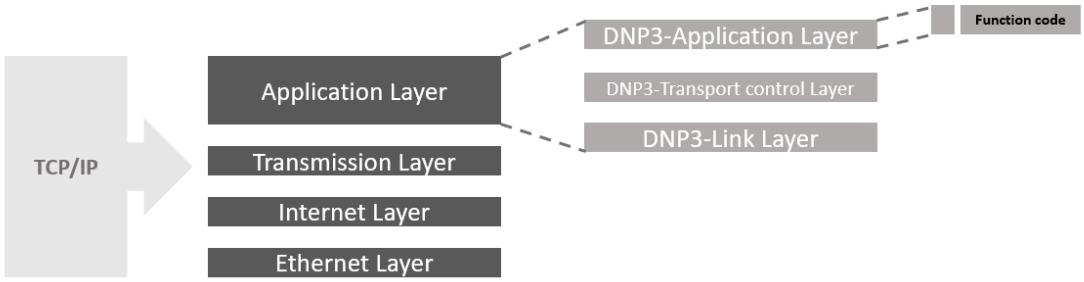

Figure 3.2 Schematic of DNP3 packet

In DNP3 communication, as shown in the Figure 3.3, the master and the outstation exchanges two types of the messages; request and response. The DNP3 Master sends a request message to its respective outstation to enquire about the status of the outstation. Similarly, the outstation replies a response message with the status information to the Master. The Figure 3.4 shows the list of request message the DNP3 Master sends to the outstation and the Figure 3.5 shows the two types of response message replied by the outstation to the DNP3 Master.

Figure 3.3 Communication between DNP3 Master and Outstation

3.2.2 Cyber attacks related to DNP3

DNP3 Data Integrity attack: As we discussed, the DNP3 has three layers datalink, transport and application layer. Most of the critical commands are set in the application layer. These critical commands directly operate the remote system when the packet is delivered. The attacker tries to exploit these commands to create instability. There are many types of data integrity attacks related to these three layers. In this work, we study the following DNP3 data integrity attacks. All the attacks mentioned here have experimented with real-time testing in SCADA traffic, and a suitable IDS rule is designed to prevent it. The cyber attack and defense work is conducted, tested, and the rule detection performance is analyzed in Power cyber lab, Iowa state university.

DNP3 Select and Operate attack: Most of the existing DNP3 communication is not encrypted. Since encryption takes additional overhead time, which limits availability. In DNP3, Select and Operate are two Functions. During system operation in SCADA, Select command is always initiated before Operate. The attacker sniffs the payload of both the command and delivers it to the target. This attack can lead to disastrous situations like a blackout, power outage. It is conducted in the lab and an IDS rule is made to prevent it.

DNP3 Direct operate attack: As compared to the previous attack, this is more severe. We have a function code 5; Direct operate in DNP3. This function does not involve with the Select function of the DNP3, as it directly operates the system. Now, it becomes simple for an attacker to modify one payload instead of two. It can affect the power system and leads to a blackout. Similar IDS rule with content and time threshold is given to check whether the traffic pattern is normal.

Code	Function	Code	Function
00	Confirm	10	Initialize application
01	Read	11	Start application
02	Write	12	Stop application
03	Select	13	Save configuration
04	Operate	14	Enable unsolicited
05	Dir operate	15	Disable unsolicited
06	Dir operate-No resp	16	Assign class
07	Freeze	17	Delay measurement
08	Freeze-No resp	18	Record current time
09	Freeze clear	19	Open file
A	Freeze clear-No resp	1A	Close file
B	Freeze at time	1B	Delete file
C	Freeze at time-No resp	1C	Get file information
D	Cold restart	1D	Authenticate file
E	Warm restart	1E	Abort file
F	Initialize data		

Figure 3.4 DNP3 Request Function codes

Code	Function
81	Response
82	Unsolicited response

Figure 3.5 DNP3 Responses Function codes

DNP3 Stop application attack: The function code related to Stop application is 12; an attacker can use this information to exploit and perform unauthorized action. In Stop application attack, the attacker tries to shut down or stop a functioning application. The impact of this could cause instability in the power system or any control systems. An IDS rule with the content field is used to counter it.

DNP3 Cold restart: Cold restart is used to clear the communication sequence and restart the end device. This is generally used to check the self-test of the device. The function code related to Cold restart is D (in hexadecimal); the attacker attempts this attack to perform an unauthorized operation. The impact leads to the unavailability of the targeted device, and if the attacker is

persistent, it can lead to Denial-of-service. An IDS rule with a time threshold to check the number of packets is employed to prevent this.

DNP3 Data reset attack: It is a type of Data integrity attack. In this type of attack, the attacker modifies the application layer field in the DNP3 protocol stack. The function code field in the application has various control operations, which are exploitable. In this attack, the Function Code F (in hexadecimal) is used, which is to Initialize data. It causes the substation to reinitialize the data objects and creates a timing mismatch in the running process. This attack was tried in the lab, using a python script with the above function code. And an IDS rule using the content field is designed to prevent this attack.

DNP3 Disable Unsolicited Messages: Any updates from the outstation are given as a response to the SCADA server. There are two types of responses sent to the server; Normal response and Unsolicited response. In Normal response, the outstation replies to the server only when there is a request message. Whereas in Unsolicited response, the outstation is bound to send the response without any request from the server. In this type of attack, the attacker modifies the Function code 15, which is a request message from the server that implies the outstation to disable the Unsolicited response. It intercepts the communication between the server and the substation, which makes the server unaware of the system status. This attack can lead to a situation where the system can be separated from real-time communication. It is implemented in the lab with a python script, and an IDS rule is designed to prevent it.

3.3 Intrusion Detection System and Rules

3.3.1 IDS rule structure

The Figure 3.6 shows a sample IDS signature. The IDS rule consists of different fields like action, protocol, source and destination address, source and destination port, the direction of the packet [24]. It also contains the rule option field like content, packet flow, time threshold, offset, and depth of the packet.

```
action protocol src-addr & port direction dest-addr & port message sid
```

Figure 3.6 IDS Rule Structure

Ping detection is one of the common means to check if the host is alive. The Figure 3.7 shows a sample ping rule, where an alert is generated if any external host tries to ping the home network.

```
alert icmp !10.1.150.1 any -> 192.168.1.210 any (msg:"Ping from Unknown source"; sid:100001;)
```

Figure 3.7 Sample Ping IDS rule

3.3.2 IDS Rule Generation based on Network Packet Payload

The network payload in standard DNP3 communication is not encrypted. Since the communication is not encrypted, there is a possible chance of sniffing. Critical commands from the SCADA are operated to control the physical processes of the substation. Now, these commands if sniffed, can be exploited by the threat actors. SCADA control commands like select, operate, direct operate, response, unsolicited response, direct operate without response can be easily exploited. There is a list of function codes in the DNP3 application layer stack that can be exploited. An IDS rule based on the packet payload is generated to prevent this type of attack. Here we inspect every network packets coming inbound with the payload of the given critical function whether it is matching with the given destination and source IP address, port number, and the packet flow. This rule enables the detection if the attacker injects any corrupted packet having the critical function to execute the system. This rule detects and alerts the Master; the rule sample is shown in the Figure 3.8. In this rule, only the communication between the DNP3 server and Home network is allowed. The IP address of the DNP3 server and the Home network is predefined in the Master and sensor system. Then the sensor checks for the content field in the rule which matches “04”. This is the DNP3 operate function code that server controls the client. The rule options field has both the offset and

depth of the packet to identify which DNP3 function code is used. This rule prevents data integrity attacks that happen at the DNP3 layer.

```
alert tcp !$DNP3SERVER any -> $HOMENET any (content:"|04|"; offset:12; depth:1; msg:"DNP3-Operate from Unknown source"; sid:4;)
```

Figure 3.8 IDS rule based on packet payload

3.3.3 IDS Rule Generation based on Network Packet flow

After inspecting the packet payload, this rule is made to detect whether the inbound traffic is coming as expected. It provides an option of checking whether the packet flow is established or not established. Before any TCP session is established, there exists a three-way handshake between the client and the server. This rule checks if the packet flow between the two sides is established or not established depending on the source and destination IP address, port number, and the packet payload. Any new TCP connection starts with SYN, SYN-ACK, and ACK; this rule finds whether any of the three packets are getting exchanged. Below Figure 3.9 shows, the sample IDS rule based on packet flow. Here the rule option field has flow-not established, states that any TCP handshake other than the DNP3 server is alerted. This rule alerts all types of scanning from the IP address, port scanning, service scanning that happen at the IP layer.

```
alert tcp !$DNP3SERVER any -> $HOMENET any (flow: not_established; msg:"Unknown TCP connection"; priority:1; sid:3;)
```

Figure 3.9 IDS rule based on packet flow

3.3.4 IDS Rule Generation based on Time-threshold

In SCADA traffic, the critical commands follow a specific pattern concerning time. For example, DNP3 commands like select are not operated more than two times in 30 seconds. Similarly, there are many network data patterns. If the traffic violates the standards, it is expected to be an anomalous behavior. Similarly, this rule is written based on the time threshold of the DNP3 packet; it checks

the network traffic, whether it is exceeding the number of attempts of that particular packet. If there is any attempt at an excessive number of the DNP3 packet, an alert is triggered. This rule detects all types of flooding attacks that happen in the SCADA environment. Flooding attacks like TCP flooding, SYN flooding, MAC flooding, reverse TCP flooding, scanning, DNP3 flooding, denial-of-service, distributed denial-of-service are detected. Below Figure 3.10 shows, the sample IDS rule based on the time threshold. Here the rule option field has threshold type, which has the count of the packet in a given time. This rule checks packet count and alerts if any count exceeds in a given time.

```
alert tcp !$DNP3SERVER any -> $HOMENET any (msg:"Anomalous operation DNP3-Operate"; content:"|04|";
        offset:12; depth:1; threshold: type both, track by_src, count 5, seconds 10; sid:4;)
```

Figure 3.10 IDS rule based on time threshold

3.3.5 IDS Rule Generation based on Incorrect Checksum

The checksum of the network packet shows whether the packet has arrived without losing data. This rule is made to detect whether the checksum is matching at the receiving end. The alert is generated if there is any incorrect checksum, thus notifies that the data has been compromised. It provides an option of checking whether the packet flow is established or not established. The sample rule is shown in the Figure 3.11. The rule option field has gid (generator id); these are predefined identifiers in the rule detection engine to check the checksum of the network packet.

```
alert tcp !$DNP3SERVER any -> $HOMENET any (msg:"DNP3-Bad-CRC"; sid:1; gid:145; metadata: rule-type preproc;)
```

Figure 3.11 IDS rule based Incorrect Checksum

3.3.6 IDS Rule on TCP Application Layer

The IDS rule on the TCP Application layer directly checks the application-layer header. In the sample rule shown in the Figure 3.12 which has DNP3 in the protocol field. This rule directly

checks the DNP3 layer header. Thus the detection time is less compared to other rules that use TCP protocol, offset and depth as rule option fields.

```
alert dnp3 !$DNP3SERVER any -> $HOMENET any (msg:"Unknown Source trying DNP3"; priority:1; sid:3;)
```

Figure 3.12 IDS rule based on the TCP Application layer

CHAPTER 4. TESTBED IMPLEMENTATION AND EVALUATION

In this chapter, we discuss the implemented IDS/IPS with testing and evaluation of different rules. As discussed in the previous chapter of designing the IDS rules, we generated around 40 advanced IDS rules. The design and implementation are entirely carried out at the Power Cyber Lab, Iowa state university [25]. Different power model like a two-area model and 39-Bus power model is used for the testing. The rules that have different option fields like content, time threshold, and packet payload, packet flow are tested concerning the time. Testing of the rules with different rule order sequence, detection time, and false positive and false negative rates are performed. Different attacks are conducted to check the detection rate and compared different rule sequence to get the most reliable detection time.

4.1 Two-area Model

In this power system model, we deployed two sensor nodes at RTU-1 and RTU-2 network. The master node is installed at the control center. The Figure 4.1 shows two-area power model from a Siemens SCADA. The SCADA server and the two RTUs are in regular operation. The sensor is installed and has different sets of IDS rules based on DNP3 and other fields. Another virtual machine having Kali operating system is created for an attacker from a different network. All the attacks are initiated from the Kali machine are directed to RTU-1 and RTU-2. The results are then noted from the sensor to check the detection time and the alert generated. The attacks performed are within the scope of four types, as discussed in chapter 3. The detection time is measured for different types of rules in sequence, and the result is plotted in the graph. Later, the same rule set is arranged in a different sequence to get the best detection time, and the result is plotted in the graph. Finally, the two graphs are compared to see the effective rule sequence to achieve the least detection time.

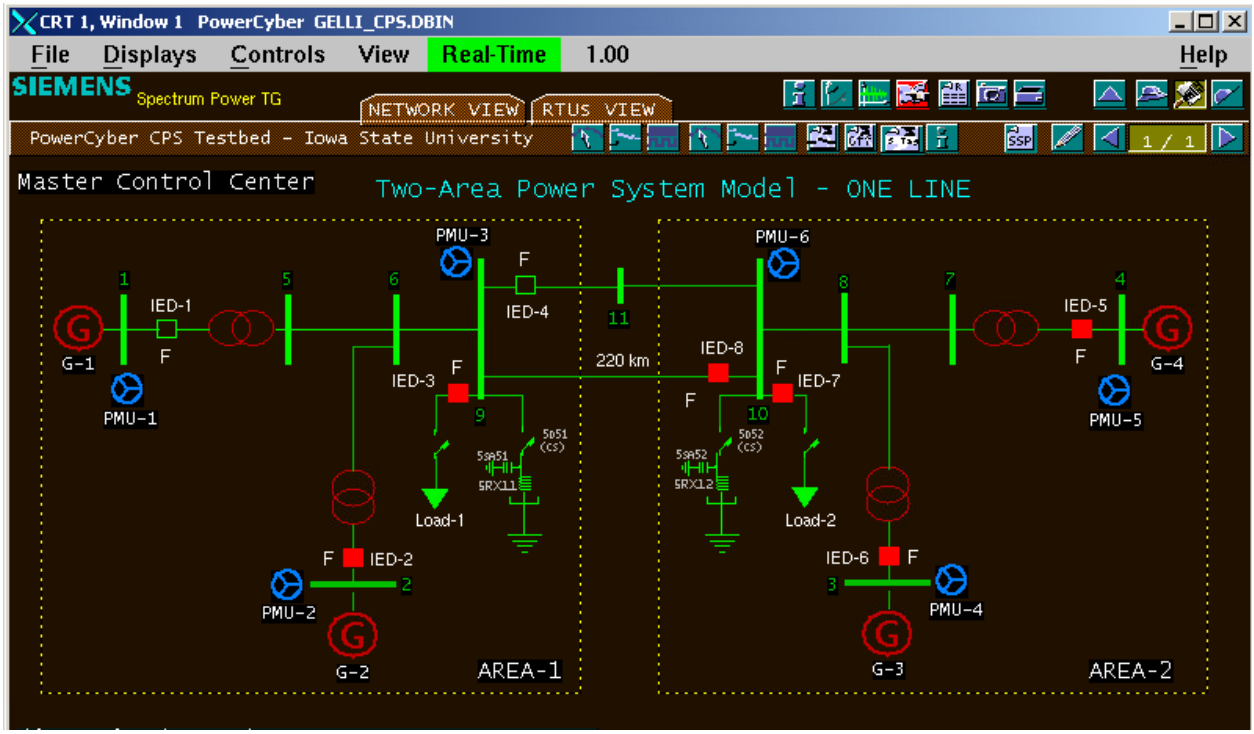

Figure 4.1 Two-area model

The Figure 4.2 shows how the detection time is calculated. We first measure the trip time from the attacker machine to the RTU, then the difference between the wireshark time at the sensor and the trip time gives the detection time. The IDS detection time is calculated for all the rules in this fashion.

The Figure 4.3 shows the first rule order sequence; in this sequence, the rules are arranged uniformly based on the action, protocol, content, and count. The detection time for each of the rule is noted and plotted in the Figure 4.4.

The Table 4.1 shows the detection time for Ping alert, Table 4.2 shows the detection time for flow not established alert, Table 4.3 shows the detection time for content rule, Table 4.4 shows detection time for time threshold rule.

Similarly, Figure 4.5 shows the second rule order sequence; in this sequence, the rules are arranged in reverse order. The order is based on the count, content, protocol, and action to check

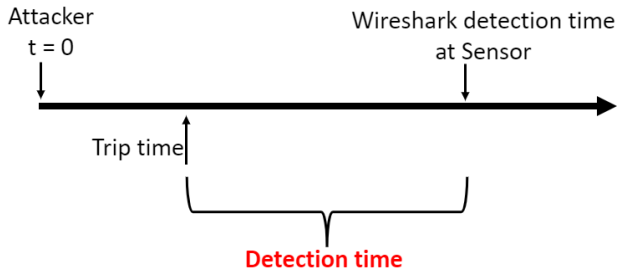

Figure 4.2 Detection Time Calculation

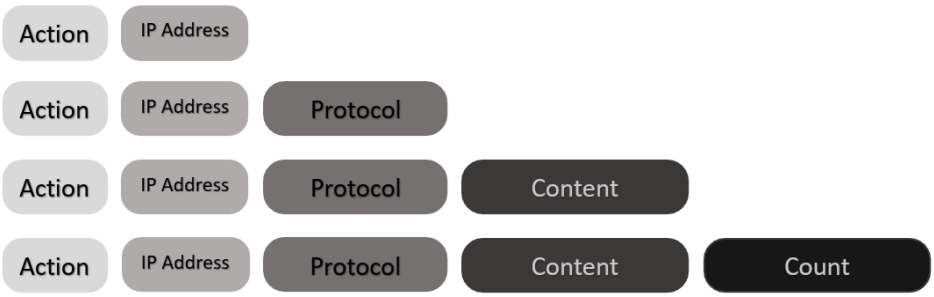

Figure 4.3 Rule order sequence 1

the detection time and performance. The detection time for each of the rule in Sequence-2 is recorded and plotted in the Figure 4.6.

The table 4.5 shows the detection time for Ping alert, table 4.6 shows the detection time for flow not established alert, table 4.7 shows the detection time for content rule, table 4.8 shows detection time for time threshold rule.

The Figure 4.7 shows a comparison between the two rule sequences. It shows that in rule sequence-2 the Ping, packet flow and content rule was detected with a small delay, as these rules were placed in the bottom of the rule repository. The detection time for the time threshold rule

Table 4.1 Ping Rule - Sequence 1

One-way trip time	Time in Wireshark	Detection time
0.85	265	265-0.85 = 264.15
0.74	218	218-0.74 = 217.26
0.69	191	191-0.69 = 190.31
0.52	215	215-0.52 = 214.48

Table 4.2 Flow not established Rule - Sequence 1

One-way trip time	Time in Wireshark	Detection time
0.85	422	422-0.85 = 421.15
0.74	434	434-0.74 = 433.26
0.69	427	427-0.69 = 426.31
0.52	329	329-0.52 = 328.48

Table 4.3 DNP3 Content rule - Sequence 1

One-way trip time	Time in Wireshark	Detection time
0.85	745	745-0.85 = 744.15
0.74	626	626-0.74 = 625.26
0.69	818	818-0.69 = 817.31
0.52	767	767-0.52 = 766.48

is lower than the sequence-1. It infers that the processing time is less for the rules placed on top order compared to others. Thus the critical IDS rules are placed on the top of the rule repository to get minimum detection time. The detection time also depends on the rule length, if the rule contains many option fields, the processing time increases.

4.2 39-Bus Power system Model

In this power system model, we deployed one master and sensor nodes at RTU-1 and RTU-2 similarly as in the Two-area power model. The Figure 4.8 shows 39-Bus power model, here the model is having three areas, where each area has many relays. In this case, the sensor-1 deployed in Substation area-1 is monitoring many relays. As discussed in Chapter 2, the traffic pattern in SCADA for a substation is different if it has many relays. Similarly, in this implementation, the

Table 4.4 Time threshold rule - Sequence 1

One-way trip time	Time in Wireshark	Detection time
0.85	1411	1411-0.85 = 1410.15
0.74	2861	2861-0.74 = 1449.26
0.69	4129	4129-0.69 = 1267.31
0.52	5640	5640-0.52 = 1510.48

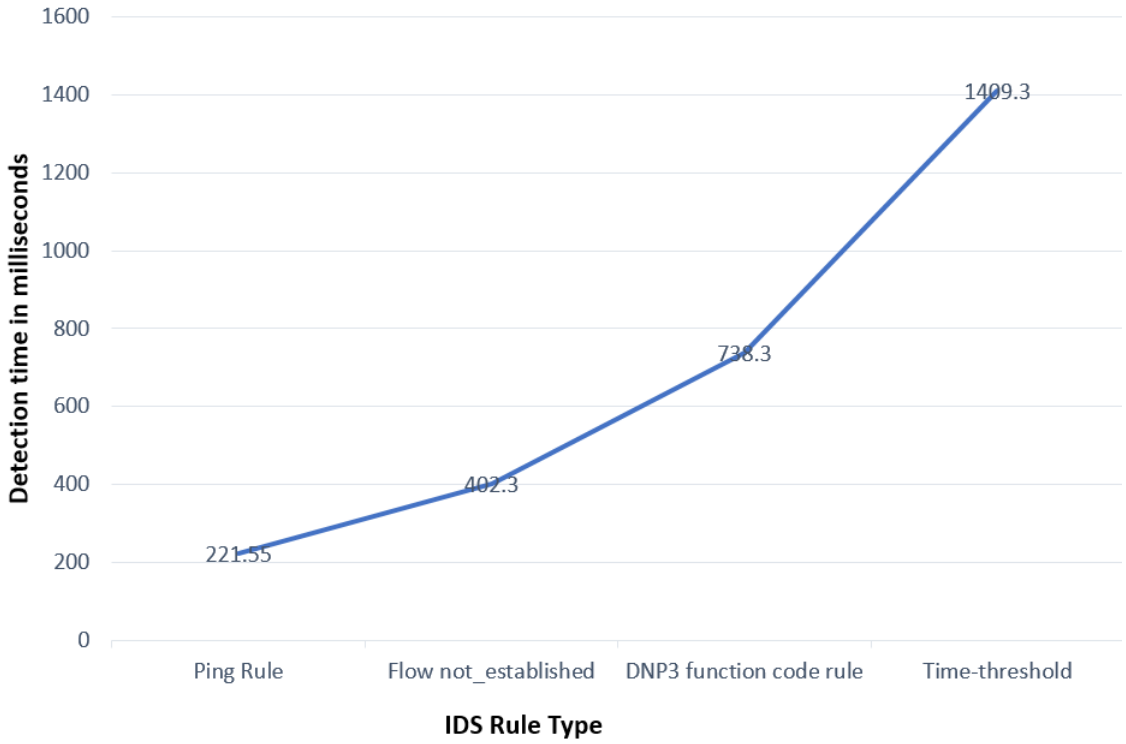

Figure 4.4 IDS Rule Sequence 1

SCADA server is now controlling a substation, which is having many relays. This implementation is different compared to the Two-area power model, although the same sensor is deployed. In this type, the SCADA traffic has many relays to a single substation; the packet rate and traffic pattern are now changed, as compared to the two-area power model. Hence the IDS rules are designed to match and prevent this type of network traffic. This implementation requires more traffic analysis and provides more IDS rules compared to the previous. The main advantage of having the distributed IDS is it offers the simplicity in designing the rules, for a complex network.

Table 4.5 Ping Rule - Sequence 2

One-way trip time	Time in Wireshark	Detection time
0.85	419	419-0.85 = 418.15
0.74	308	308-0.74 = 308.26
0.69	349	349-0.69 = 348.31
0.52	327	327-0.52 = 326.48

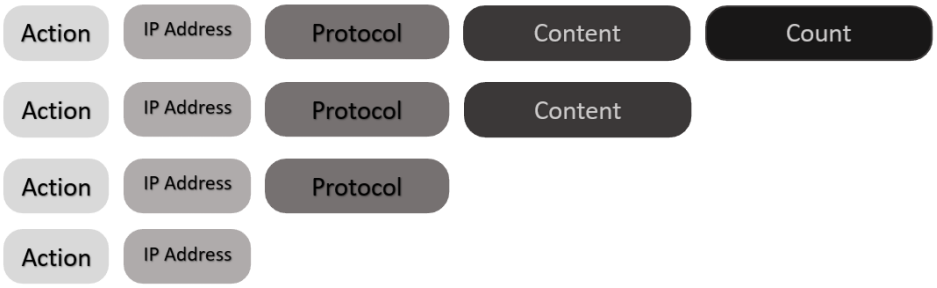

Figure 4.5 Rule order sequence 2

Table 4.6 Flow not established - Sequence 2

One-way trip time	Time in Wireshark	Detection time
0.85	560	560-0.85 = 559.15
0.74	474	474-0.74 = 473.26
0.69	554	554-0.69 = 553.31
0.52	478	478-0.52 = 477.48

Table 4.7 DNP3 Content Rule - Sequence 2

One-way trip time	Time in Wireshark	Detection time
0.85	890	890-0.85 = 889.15
0.74	701	701-0.74 = 700.26
0.69	760	760-0.69 = 759.31
0.52	805	805-0.52 = 804.48

Table 4.8 Time threshold Rule - Sequence 2

One-way trip time	Time in Wireshark	Detection time
0.85	1346	1346-0.85 = 1345.15
0.74	1396	1396-0.74 = 1395.26
0.69	1239	1239-0.69 = 1238.31
0.52	1302	1302-0.52 = 1301.48

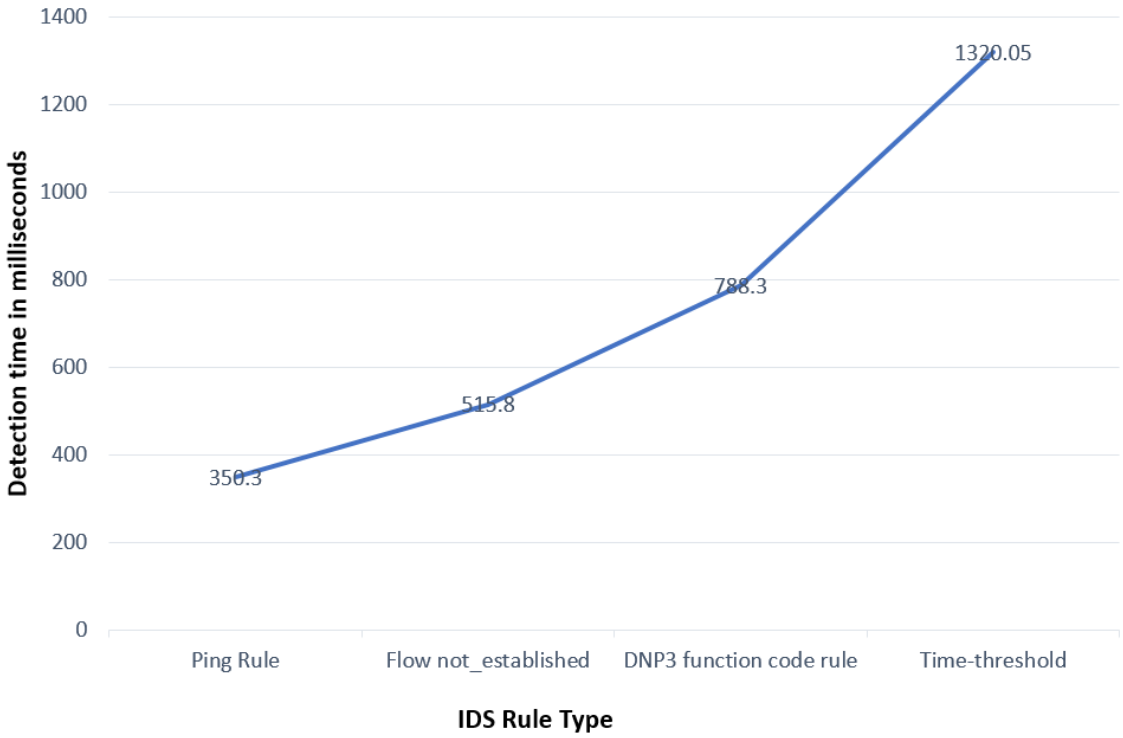

Figure 4.6 IDS Rule Sequence 2

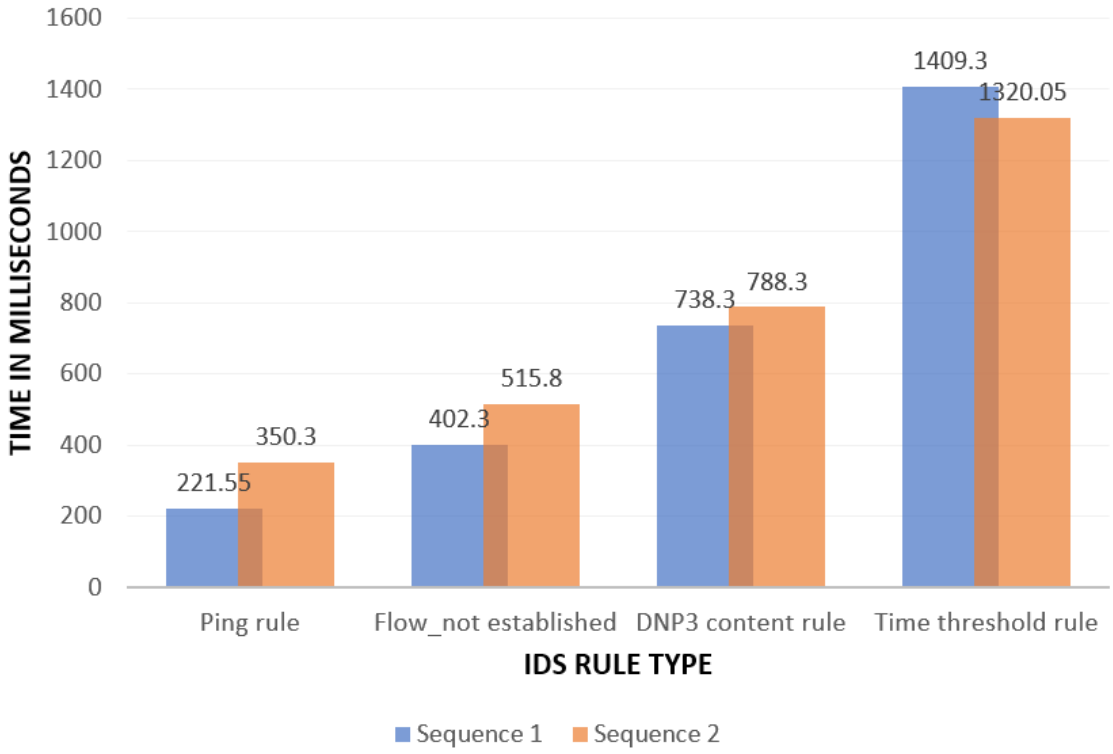

Figure 4.7 Sequence 1 vs Sequence 2

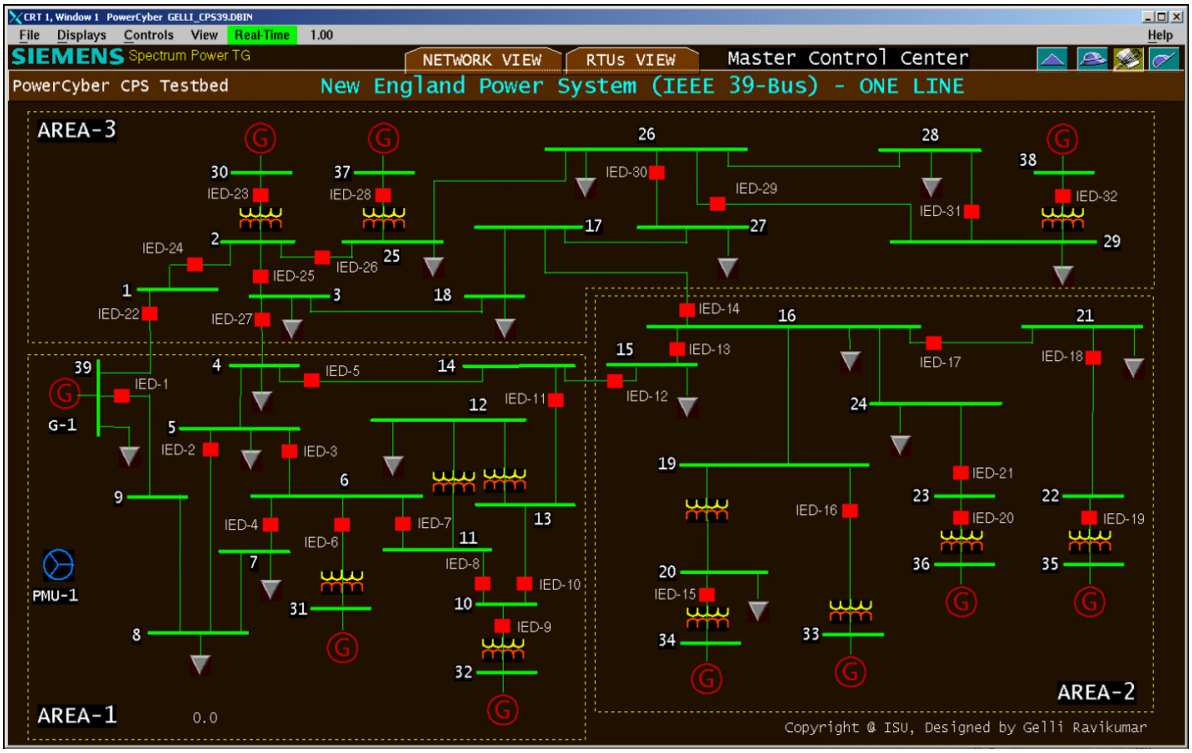

Figure 4.8 39-Bus Power system model

CHAPTER 5. CASE STUDY AND FIELD DEPLOYMENT

In this chapter, we discuss entirely on how distributed IDS was successfully transferred to local power utility in Iowa. For security reasons, the utility name is not disclosed. Before the start of the deployment, we showed a presentation of a standalone master. A single master node performing the operation of forward , storage and sensor node. Phase I and Phase II were started in January 2019 to bring out the advantage of having distributed IDS and to show the effectiveness of the IDS. After conducting repeated experiments and testing at power cyber lab for different IDS rules based on various scenarios, both phases were completed. The power utility had a primary control center, where all the data operations are conducted. The control center has a network of substation which are connected in ring fiber link. The deployment was assisted by a senior network engineer and network administrator of the organization. The master node was transferred through a VMDK virtual file. Later, master virtual machine was created from the VMDK. The client-1 and client-2 were installed and configured from the scratch and made the connection up to the master node.

5.1 Network Architecture

The overall diagram of the network architecture is shown in Figure 5.1, which shows all the substations are inside the ring topology. Since the substation are having the critical control systems they are kept in the ring topology. The topology shows a primary control center and two substations. The primary control center is connected to the fiber ring which has an active link and backup link. The links are connected to Layer 3 switch, which manages in providing different VLANs for different areas of substations. All the IP addresses are assigned with respect to VLANs of that area substation. For example: if the RTU-1 IP address is 10.1.200.210 in VLAN 20 then the Sensor-1 IP address is given as 10.1.200.110 in VLAN 20. Similarly, it is done for Substation-2, once both

the sensors are up, they are connected to the Master, which is located in primary control center.

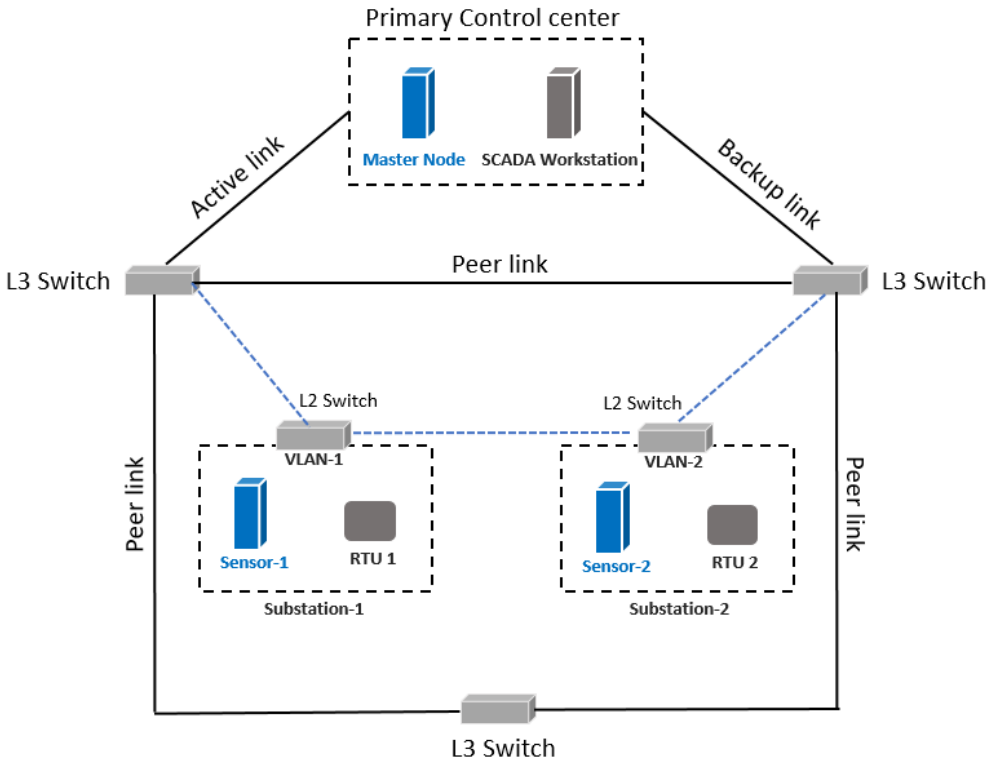

Figure 5.1 Network Architecture

5.2 Phase-I Deployment

5.2.1 Master node and Client-1

As shown in the Figure 5.1 the master node is started in Phase I. The master node is installed in the operations network. After the master is up, client-1 is then installed in substation-1 network. The deployment started with installing the packages and operating systems after analyzing SCADA traffic and generating various IDS rules at the power cyber lab. The rule set is transferred and used in phase I deployment. The rule set contains around 40 IDS rules using DNP3 packet content and time threshold. The rule set is updated in the master; then it is pushed to the client-1 at substation-

1 virtually. After the production is up, the time notifications of the IDS and features are recorded. Also, the network setup and deployment agenda are saved for the phase II deployment.

5.3 Phase-II Deployment

5.3.1 Client-2 at Substation-2

In phase II deployment, client-2 is installed in the substation-2 network. After, phase I, the network setup and experience made it comfortable to deploy the system. The previous IDS rules were modified with additional detection features. Then the rule set is updated at the master. After the master is up the rules are pushed to client-2. Similarly, in phase II, after the production is up, the time notifications of the IDS and status are noted. Finally, the distributed IDS setup commenced in Power cyber lab successfully achieved a field deployment. This work was presented in conferences and played an essential part in the CPS training for various occasions.

5.4 Outreach and Conferences

After successful deployment, this work was presented at the Electric and Power Research Center (EPRC) conference held in Ames, Iowa. The presentation was delivered by Dr. Ravikumar Gelli, and senior network engineer from the power utility joined to share their experience of the deployment. Also, we took this work in demonstrating the need for the distributed Intrusion detection setup for many CPS training. In May 2019, a team from Power cyber lab including Dr. Gelli, Sathya Mohan and Burhan Hyder from Iowa state university provided cyber-defense training for Florida reliability coordinating council (FRCC). This work was showcased in various CPS training and provided knowledge transfer to employees of the power utility. This work also played an integral part during the NERCs GridEx based training conducted during the EPRC conference. The work titled " SIEM based Distributed IDS for SCADA" which engrossed the audience at the 6th Graduate and Professional and Student Research Conference (GPSRC). Also, The work was awarded as the best research poster. The paper titled "Distributed Intrusion Detection System using Semantic-based Rules for SCADA in Smart Grid" is submitted for the conference.

CHAPTER 6. CONCLUSION AND FUTURE WORK

6.1 Summary

Securing data communications is one of the essential aspects to consider in critical infrastructures. While no network is immune to attacks, stable and efficient network security is necessary to protect industrial communications. Good network security helps in reducing the risk of cyber-attacks and decreases threats and vulnerability. There are several levels of network-level protection to prevent cyber attacks like man-in-the-middle, data integrity attacks, social engineering, denial-of-service attacks. Hence achieving secure communication is much needed in today's world. The security measures taken to provide security are encryption, authentication, authorization, access control, firewall, IDS/IPS, Security incident, and event management (SIEM). This work describes a typical approach in IDS and IPS by making it distributed to monitor the network activity of the different networks. The primary approach is to protect the distributed network against attacks on the DNP3 protocol. Several attack signatures were formulated and to prevent IDS rule were designed. The design of the rule generation algorithm is explained based on different SCADA traffic patterns. In this traffic pattern, various IDS rules are designed, and the flow chart is shown. The implementation was deployed using several network architectures like standalone, distributed, and In-line. The cyber attack classification is prepared within the scope of the lab, and the testing is performed for designed IDS rules. Then, the evaluation of the IDS rules is completed to check the best detection time. Experimental results are collected to show the best rule order sequence, where the processing time is less for the rules placed on the top of the repository. The standard defense strategy, as shown in the Figure 6.1 should be employed to make the network robust against malicious threats and incidents.

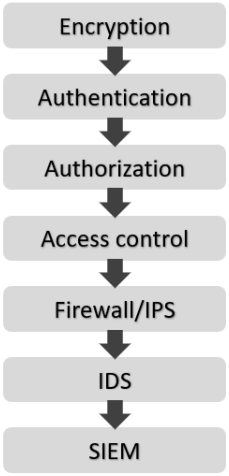

Figure 6.1 Defense Strategy

6.2 Future work

The possible extension of this work would be introducing data analytics. Data science is becoming a significant component in all cyber security research that will help in coming up with new techniques to develop advanced detection. Use of advanced machine learning algorithm using the network traffic data provides vast scope for future research. The network traffic has many features that can be extracted and used for both supervised and unsupervised learning techniques. Other techniques like Function code hashing in DNP3 communication can bring up the integrity of the critical commands. Similarly, the concept of moving target defense on the link-layer station address prevents attacks like Denial-of-service. The scope of this work can be extended by dynamically changing the IDS rules depending on different types of cases. Cyber security, a modern and futuristic research area, has a wide scope for different techniques. In the future, there is a plan to integrate this work for upcoming experiments and will provide a methodology for designing an IDS/IPS rules.

BIBLIOGRAPHY

[1] "US Energy Information Administration," 2017. [Online]. Available: https://www.eia.gov/state/?sid=NY#tabs-3

[2] "Cyber-Attack Against Ukrainian Critical Infrastructure," 2016. [Online]. Available: https://www.us-cert.gov/ics/alerts/IR-ALERT-H-16-056-01

[3] D. Kushner, "The real story of stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, March 2013.

[4] N. iek and H. Deli, "Demand response for smart grids with solar power," in *2014 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, May 2014, pp. 566–571.

[5] W. Luan, D. Sharp, and S. LaRoy, "Data traffic analysis of utility smart metering network," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–4.

[6] S. Prabhu and R. Venkat, "High availability for network management applications," in *The Second International Conference on Availability, Reliability and Security (ARES'07)*, April 2007, pp. 493–498.

[7] M. P. Samuel East, Jonathan Butts and S. Shenoi, "A TAXONOMY OF ATTACKS ON THE DNP3 PROTOCOL," 2016. [Online]. Available: http://dl.ifip.org/db/conf/ifip11-10/cip2009/EastBPS09.pdf

[8] Y. Xu, Y. Yang, T. Li, J. Ju, and Q. Wang, "Review on cyber vulnerabilities of communication protocols in industrial control systems," in *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Nov 2017, pp. 1–6.

[9] Z. Drias, A. Serhrouchni, and O. Vogel, "Taxonomy of attacks on industrial control protocols," in *2015 International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS)*, July 2015, pp. 1–6.

[10] I. Darwish, O. Igbe, O. Celebi, T. Saadawi, and J. Soryal, "Smart grid dnp3 vulnerability analysis and experimentation," in *2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing*, Nov 2015, pp. 141–147.

[11] "DNP3 PROTOCOL PCAP files," 2018. [Online]. Available: https://github.com/ITI/ICS-Security-Tools/tree/master/pcaps/dnp3

[12] L. Li and H. Peng, "A defense model study based on ids and firewall linkage," in *2010 International Conference of Information Science and Management Engineering*, vol. 2, Aug 2010, pp. 91–94.

[13] A. Sharma and M. Sharma, "Analysis and implementation of bro ids using signature script," in *2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI)*, Oct 2015, pp. 57–60.

[14] X. Zhang, C. Guo, and L. Wang, "Using game theory to reveal vulnerability for complex networks," in *2010 10th IEEE International Conference on Computer and Information Technology*, June 2010, pp. 978–984.

[15] G. Gilchrist, "Secure authentication for dnp3," in *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, July 2008, pp. 1–3.

[16] "DNPSec: Distributed Network Protocol Version 3 (DNP3) Security Framework," 2005. [Online]. Available: https://www.acsac.org/2005/techblitz/majdalawieh.pdf

[17] O. Igbe, I. Darwish, and T. Saadawi, "Deterministic dendritic cell algorithm application to smart grid cyber-attack detection," in *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*, June 2017, pp. 199–204.

[18] "Security Onion," 2017. [Online]. Available: https://github.com/Security-Onion-Solutions/security-onion

[19] "Attack surface Host Analyzer," 2018. [Online]. Available: https://aha-project.github.io/

[20] "Elastic search," 2018. [Online]. Available: https://www.elastic.co/

[21] "CVE," 2018. [Online]. Available: https://cve.mitre.org/

[22] "National Institute of Standards and Technology," 2018. [Online]. Available: https://www.nist.gov/

[23] Y. E. G. Jinan Fiaidhi, "SCADA Cyber Attacks and Security Vulnerabilities:Review." [Online]. Available: https://pdfs.semanticscholar.org/5ec6/597ac98af5032500d73cf5c3ba73c4c2a197.pdf

[24] "Snort- Network Intrusion and Prevention system," 2018. [Online]. Available: https://www.snort.org/

[25] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 847–855, June 2013.