**A DISSERTATION**

**ON**

**DATA INTEGRITY ASSAULT DETECTION IN A SMART GRID NETWORK**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD

OF

**MASTER OF TECHNOLOGY**

IN

**ELECTRICAL ENGINEERING**

(**POWER SYSTEM ENGINEERING)**

UNDER

**ASSAM SCIENCE AND TECHNOLOGY UNIVERSITY**

(SESSION: 2021-23)



SUBMITTED BY

**SONALI THAKUR**

PG -EE-09

ASTU Roll No 210620064017

ASTU Reg. No-244607117

UNDER THE GUIDANCE OF

**Dr. BIMAL CHANDRA DEKA**

Professor

**Mrs KUMARI NUTAN SINGH**

Assistant Professor

**DEPARTMENT OF ELECTRICAL ENGINEERING**

**ASSAM ENGINEERING COLLEGE, JALUKBARI, GUWAHATI-781013**

Department of Electrical Engineering

Assam Engineering College

Guwahati-781013

**DECLARATION BY THE CANDIDATE**

I hereby declare that the work presented in the dissertation entitled “DATA **INTEGRITY ASSAULT DETECTION IN SMART GRID NETWORK**” is an authentic record of my own work submitted in partial fulfillment of requirements for award of the degree of **MASTER OF TECHNOLOGY,** with specialization in **POWER SYSTEM ENGINEERING**, under supervision of **DR BIMAL CHANDRA DEKA** and **MRS KUMARI NUTAN SINGH,** Department of Electrical Engineering, Assam Engineering College, Assam-781013. The matter embodied in this work has not been submitted to any other institute for award of any degree or diploma. I have followed the guidelines provided by the Department of Electrical Engineering, Assam Engineering College as per Assam Science and Technological University regulation, in preparing the project report. I have also conformed to the norms and guidelines given in the ethical code of conduct of Assam Engineering College. Whenever I have used materials (example: data, theoretical analysis, figures, tables and text) from other sources, I have given due acknowledgement to them by citing them in the text of this project report and giving their details in the bibliography.

Sonali Thakur

Roll No. 210620064017

M. Tech (EE)

Department of Electrical Engineering

Assam Engineering College

Guwahati-781013

**CERTIFICATE FOR APPROVAL**

This is to certify that the project work entitled **“DATA INTEGRITY ASSAULT DETECTION IN A SMART GRID ENVIRONMENT**’’, undertaken by Ms. Sonali Thakur (Roll No 210620064017), a M. Tech 4th Semester student of Department of Electrical Engineering, Assam Engineering College,Guwahati-13, has been carried out by her during the academic year 2022-2023 under our supervision and guidance.

....................................… ..........................................

Dr .Bimal Chandra Deka Mrs Kumari Nutan Singh

Department of Electrical Engineering Department of Electrical Engineering

Assam Engineering College Assam Engineering College

Date: Date:

Department of Electrical Engineering

Assam Engineering College

Guwahati-781013

**CERTIFICATE FOR APPROVAL**

This is to certify that Sonali Thakur (Roll No. 210620064017), a M Tech 4th Semester student of the Department of Electrical Engineering, Assam Engineering College, has submitted her dissertation on “**DATA INTEGRITY ASSAULT DETECTION IN A SMART GRID NETWORK**’’ in partial fulfilment for the award of degree of Master of Technology in Electrical Engineering with specialization in Power System Engineering under Assam Science and Technology University.

Date

Place External Examiner

Department of Electrical Engineering

Assam Engineering College

Guwahati-781013

**CERTIFICATE OF ACCEPTANCE**

This is to certify that Ms. Sonali Thakur (Roll No 210620064017) a M. Tech 4th semester student of the Department of Electrical Engineering, Assam Engineering College, has submitted her dissertation on “**DATA INTEGRITY ASSAULT DETECTION IN A SMART GRID NETWORK**’’, in partial fulfilment for the award of the degree of “**Master of Technology**” in Electrical Engineering with specialization in **Power System Engineering** under **Assam Science and Technology University**.

**Dr. Aroop Bardalai**

Professor and Head of the Department

Date: Electrical Engineering Department

Assam Engineering College

Jalukbari, Guwahati-781013

Place:

**ACKNOWLEDGEMENT**

The ability to help, the patience to exercise diligence and to provide support is a quality admonished by very few. Any job in the world, however, trivial or tough cannot be accomplished without the assistance of others. I am extremely grateful and will remain highly indebted to my guide Dr. Bimal Chandra Deka and co-guide Mrs. Kumari Nutan Singh for being a source of inspiration and for their constant support. I also thank for their insightful comments and suggestions which continually helped to understand the insight of the project. I also extend my deep gratitude to Dr. Aroop Bardalai, Head of Department of Electrical Engineering Department and Dr. Bani Kanta Talukdar ,Assam Engineering College for providing the opportunity and all possible facilities towards this work. I sincerely thank all the other faculty members of the department.

With Regards

Sonali Thakur

**ABSTRACT**

A Smart Gridis an electricity network based on digital technology that is used to supply electricity to consumers via Two-Way Digital communication. Smart grid applications include managing demand, integrating smart meters and generated energy, storing and managing renewable resources, and utilizing systems that continuously provide and use data from an energy network. Smart grids are vulnerable to various types of cyber-attacks due to their high requirements of communication networks to convey, sense and control data to improve its efficiency in all matters such as energy generation, transmission, and delivery. The majority of surveys classify attacks based on the security requirements, confidentiality, integrity, and availability. Cyber-attacks in a smart grid could lead to energy theft, customer privacy breach etc. Machine learning, a subset of artificial intelligence uses [algorithms](https://builtin.com/learn/tech-dictionary/algorithm) born of previous datasets and statistical analysis to make assumptions about a computer’s behavior. The computer can then adjust its actions, even performing functions it was not programmed to do. These abilities have made machine learning a crucial cyber security asset. In the work presented, the main aim is to provide a cyber-security solution for a smart grid network. Machine learning based approach is used for the same. The initial requirement is to have a smart grid for the work. Therefore, the primary objective is to replicate a smart grid using software. For this purpose ‘Matlab’ software is used. At first a power system modeling is done to represent a traditional grid. IEEE 5 bus is modeled to understand and represent a basic grid. After that a higher order bus consisting of 14 buses is developed which is upgraded to a smart grid. To upgrade the traditional 14 bus system to a smart grid system, the fact that a smart grid integrates renewable sources is utilized. Renewable energy sources are purposefully added to the 14 bus system to have a Smart grid environment. Solar PV system is added to bus 12, Wind energy system is added to Bus 8 and Battery Energy storage is added Bus 1 of the 14 bus system. Hence, a smart grid is replicated. The next task is to have data generated from the smart grid to carry out the proposed work. Electrical parameters such as voltage, current, power etc. are recorded over time from software run Smart Grid to have large set of data. The project aims to provide a cyber-security solution so the next requirement is to have a problem to begin with, which is a scenario of cyber-attack. A scenario of attack on integrity of data is created by making subtle changes in the data set of one of the electrical parameters obtained from one of the buses of the smart 14 Bus system. Thus a set of corrupt data is obtained and thereby cyber-attack is created. The goal is to provide an intelligent system that detects this manipulation and identify the points of anomalous data. So the final step is the detection of the cyber-attack. Machine learning approach is used for the detection. An unsupervised learning algorithm Isolation Forest Algorithm is used. In the work presented, three observations are manipulated and isolation forest algorithm is applied on the corrupt data. Based on anomaly score obtained, the anomalous points are clearly indicated. Pictorial representation of the anomaly score is also presented for clear understanding of the detection. Therefore, the project work is an approach to provide a cyber-security solution for a Smart grid network and it covers the entire process from initially developing a traditional grid, converting it to a Smart grid, creating a cyber-attack scenario and ultimately providing an intelligent solution for the attack detection using a machine learning algorithm.

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| **Chapter** | **Title Page No.** |
|  | CANDIDATE’S DECLARATION 2 |
|  | CERTIFICATE FOR APPROVAL 3-4 |
|  | CERTIFICATE OF ACCEPTANCE 5 |
|  | ACKNOWLEDGEMENT 6 |
|  | ABSTRACT 7 |
|  | TABLE OF CONTENT 8 |
|  | LIST OF TABLES 9 |
|  | LIST OF ABBREVIATIONS 10 |
|  | LIST OF FIGURES 11 |
| **Chapter 1** | **Introduction**  12-14 |
| 1.1 | Introduction 12 |
| 1.2 | Related work 12 |
| 1.3 | Motivation 13 |
| 1.4 | Research problem 14 |
| **Chapter 2** | **Smart grid**  15-18 |
| 2.1 | A Smart Grid 15 |
| 2.2 | Smart Grid Architecture 15 |
| 2.3 | Advantages of a Smart Grid 16 |
| 2.4 | Smart Grid Pilot projects in India 17 |
| 2.5 | Smart Grid for the proposed work 18 |
| **Chapter 3** | **Cyber security** 19-22 |
| 3.1 | Cyber security 19 |
| 3.2 | Cyber security issues in a Smart Grid 20 |
| 3.3 | Artificial Intelligence techniques used for cyber-attack detection 21 |
| 3.4 | Machine learning for anomaly detection 22 |
| **Chapter 4** | **Methodology**  23-24 |
| 4.1 | Proposed Methodology 24 |
| **Chapter 5** | **Case study** 25-29 |
| 5.1 | Case Analysis 25-29 |
| **Chapter 6** | **Results and discussion** 30-33 |
| **Chapter 7** | **Conclusion**  34 |
|  | **Appendix** 35-39 |
|  | **References** 40-41 |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| Table No | Title | Page No |
| **1** | Global serious smart grid attacks and damages. | 13 |
| **2** | Comparison between existing grid and Smart Grid | 16 |

**LIST OF ABBREVIATIONS**

1 SG Smart grid

2 EPRI Electric Power Research institute

3 AI Artificial Intelligence

4 DDoS Distributed Denial Of Service

5 DoS Denial of Service

6 CDIA Covert Data Integrity Assault

7 IT Information Technology

8 MitM Man- In-The-Middle

9 ICT Information and Communication Technology

10 TRA Threat Risk Assessment

11 IPDS Integrated Power Development Scheme

12 CAPTCHA Completely Automated Public Turing Test to

Tell Computers and Humans Apart

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| Figure No | Title | PageNo |
| 2.1 | Smart grid | 15 |
| 2.4 | Smart grid pilot projects | 18 |
| 3.2 | Cyber security in a smart grid | 20 |
| 3.4 | Anomaly detection using machine learning | 22 |
| 4.1 | Isolation forest tree | 23 |
| 5.1 | Single line diagram of IEEE 5 bus system | 25 |
| 5.2 | IEEE 5 bus system model | 26 |
| 5.3 | Single line diagram of IEEE 14 bus system | 26 |
| 5.4 | IEEE 14 Bus system model | 27 |
| 5.5 | Grid with renewable energy | 27 |
| 5.6 | IEEE 14 Bus - Smart grid model | 28 |
|  | A) Battery energy storage system model | 28 |
|  | B )Solar (pv) system model | 29 |
|  | C) Wind energy model | 29 |
| 6.1 | Simulink results of current of bus 9 | 30 |
| 6.2 | Simulink results of voltage of bus 13 | 30 |
| 6.3 | Simulink results of of active power bus 8 | 31 |
| 6.4 | Simulink results of reactive power of bus 7 | 31 |
| 6.5 | Normal data sample | 32 |
| 6.6 | Attacked sample | 32 |
| 6.7 | Anomaly score | 33 |
| 6.8 | Graphical representation | 33 |

**CHAPTER 1**

**INTRODUCTION**

* 1. **Introduction**

The likelihood of cyber-attacks and other mishaps rising as more devices are interconnected is one of the forthcoming difficulties that the current smart grid will confront, according to a literature from the Electric Power Research Institute (EPRI). The size and complexity of the smart grid make cyber security an even greater challenge. To address cyber dangers and provide architectural and analytical countermeasures to stop such attacks, a number of researchers and standard bodies have created complete frameworks [1].

Machine learning is of rising importance in cyber security. The primary objective of applying machine learning in cyber security is to make the process of assault detection more actionable, scalable and effective than traditional approaches, which require human intervention. To create intelligent security systems, it is essential to identify hidden trends and insights in network data and to build a corresponding data-driven machine learning model to stop these attacks. Machine learning can identify cyber security threats more effectively than other software-oriented methodologies, which lessens the workload on security analysts even though it cannot automate a full cyber security system [2, 3]. Fei Tony Liu and Zhi-Hua Zhou had created an algorithm, called Isolation Forest in 2008. It was put forth as a substitute for established anomaly detection techniques, which were costly computationally and prone to over-fitting. The technique is built on the hunch that anomalies can be more quickly distinguished from regular data points.

First concern is the data availability, then the data confidentiality and finally the data integrity. The first wave of cyber-attacks can be ascribed to Malware which are designed to cut access to the data or DDOS attacks which deny access to a website under attack. The second wave of cyber threats are characterized by “data theft”. Experts are now talking about a third wave known as Data Integrity Attacksthat neither prevent from accessing data like an ordinary malware does nor do they steal data like those ransom wares. Instead they compromises the integrity (i.e., accuracy and reliability) of the data.

**1.2 Related Work**

Information security is increasingly dependent on AI and machine learning, which can quickly analyze millions of data sets and find a variety of cyber threats, from malware threats to dubious activity that could lead to a phishing attempt. For the proposed work firstly research is done on the architecture of a smart grid [4]. The cyber security threats to a smart grid network is studied from various scholarly articles, papers and journals [1,2]. Data Integrity Assaultis chosen as the threat to be considered for the study. Use of artificial intelligence in cyber security is studied [5]. An extensive research on use of machine learning for cyber security is done and “Isolation Forest” algorithm has been chosen for the cyber-attack detection method for the project. Previous studies on use of Isolation Forest algorithm are studied. Power system modeling is done using *Matlab* software. Designing of the power system is done with careful consideration of our main aim of detection of cyber-attack in a smart grid environment. Related works that are studied are cited accordingly in reference. Previously designing of a smart grid and collecting the electrical parameters and then specifically providing an approach to contribute a method for data integrity assault detection is not made and hence this is done in this work based on study of all related works.

**1.3 Motivation**

The power industry is facing cyber-attacks that can impact the power grid with serious implications. The work is motivated by global serious smart grid attacks and damages [1]. Due to the high dependence of the smart grid on computer networks and other related technologies, cyber-attacks interfere with the normal operation of power systems. Once the power grid is attacked, it will cause immeasurable losses to normal production and life. The serious smart grid attacks and damages happen globally in recent times are shown in Table 1.1.

Summarizing the attacks against smart grids, it is found that the reasons for these losses mainly include two aspects, namely, the vulnerability of the smart grid and cyber-attacks launched by exploiting vulnerabilities. Therefore, it is necessary to find ways to protect the system from these attacks to prevent damages.

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Attack Targets and Place** | **Attack methods** | **Consequences** |
| June 2017 | Chernobyl nuclear power plant  Ukraine | Petya blackmail virus | Several national power facilities were infected, resulting in abnormal operations. |
| March 2019 | 1. S Power grid   United States | Denial of service (DoS) attack | The accident did not cause power failure, and the machine failure time was less than five min. |
| August 2019 | City Power company  South Africa  (Johanseburg) | Blackmail software | The attack prevented users from buying electricity, recharging, processing invoices and accessing the official website of city power. |
| June 2020 | Enel Group  Europe | Snake blackmail software. | The internal IT network was temporarily blocked, resulting in temporary interruption of customer service activities. |
| October 2022 | Tata Power  India | Not known | Impact on some IT systems |

Table 1.1 - Global serious smart grid attacks and damages.

**1.4 Research Problem**

This project work is on cyber-attack and detection in a smart grid network. The aim of the project is to provide a security solution for a data integrity attack in the communication network of a smart grid. Machine learning, a subset of artificial intelligence uses [algorithms](https://builtin.com/learn/tech-dictionary/algorithm) born of previous datasets and statistical analysis to make assumptions about a computer’s behavior. The computer can then adjust its actions, even performing functions it was not programmed to do. These abilities have made machine learning a crucial cyber security asset. In this work machine learning based approach is used for the cyber-attack and detection.

The main objective of the work is described below:

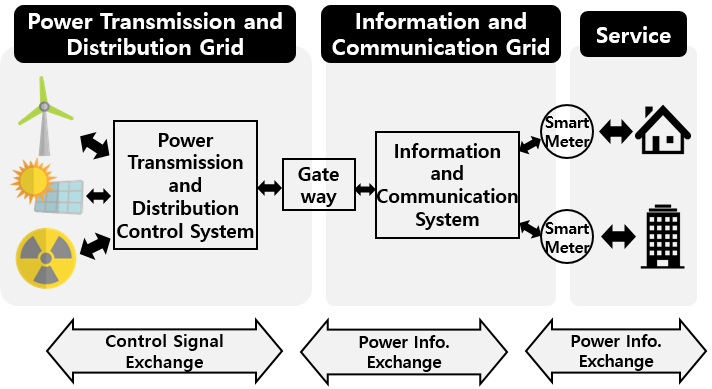
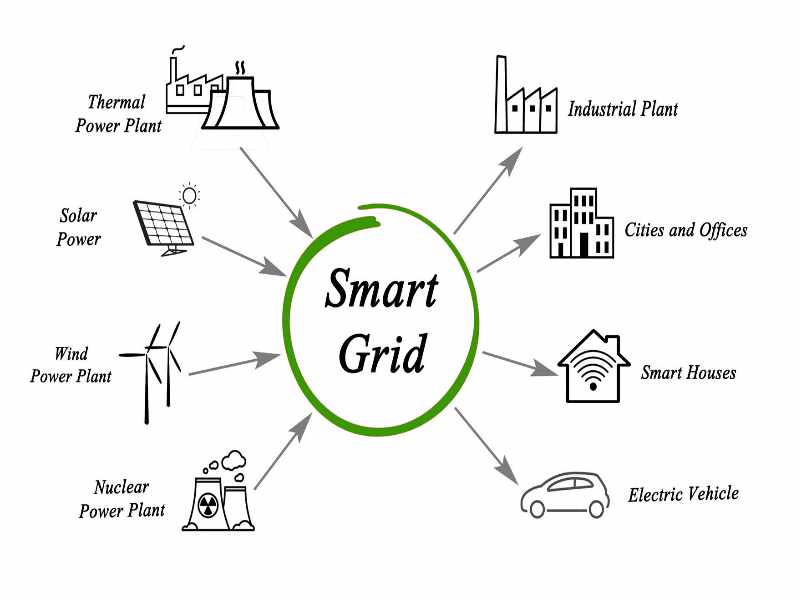
* Understanding architecture of a Smart Grid.
* Research on challenges of a Smart Grid.
* Detailed study on cybersecurity issues in a Smart Grid.
* Detailed analysis on Machine learning as a cyber-security asset to provide cybersecurity solutions.
* Replicating a Smart Grid using software - Power system modeling is done to represent a traditional grid. IEEE 5 bus is modeled to understand and represent a basic grid. After that a higher order bus consisting of 14 buses is developed which is upgraded to a smart grid. To upgrade the traditional 14 bus system to a smart grid system, the fact that a smart grid integrates renewable sources is utilized. Renewable energy sources are purposefully added to the 14 bus system to have a Smart grid environment. Solar PV system is added to bus 12, Wind energy system is added to Bus 8 and Battery Energy storage is added Bus 1 of the 14 bus system. Hence a smart grid is replicated.
* Data generation and cyber-attack-The next task is to have data generated from the smart grid to carry out the proposed work. Electrical parameters such as voltage, current, power etc. are recorded over time from software run Smart Grid to have large set of data. The project aims to provide a cyber-security ‘solution’ so the next requirement is to have a ‘problem’ to begin with, which is a scenario of cyber-attack. A scenario of attack on integrity of data is created by making subtle changes in the data set of one of the electrical parameters obtained from one of the buses of the smart 14 Bus system. Thus a set of corrupt data is achieved and thereby cyber-attack is created.
* Detection-The goal is to provide an intelligent system that detects the manipulation and identify the points of anomalous data. An unsupervised learning algorithm Isolation Forest is used for the same. The Isolation Forest algorithm needs an **anomaly score** to have an idea of how anomalous a data point is. Anomaly detection with Isolation Forest is a process composed of two main stages: in the first stage, a training data set is used to build iTrees and in the second stage, each instance in the test set is passed through these iTrees, and a proper “anomaly score” is assigned to the instance. Once all the instances in the test set have been assigned an anomaly score, it is possible to mark as “anomaly” any point whose score is greater than a predefined threshold. Based on anomaly score obtained, the manipulation in the data is detected.

**CHAPTER 2**

**SMART GRID**

**2.1 Smart Grid**

A Smart Grid is an electricity network based on digital technology that is used to supply electricity to consumers via Two-Way Digital Communication. Smart grid applications include managing demand, integrating smart meters and generated energy, storing and managing renewable resources, and utilizing systems that continuously provide and use data from an energy network [4].

  Figure 2.1 SMART GRID

**2.2 Smart Grid architecture**

Since the first grid was established in 1886, the number of power plants increased as the demand for electricity increased from the 1970s to the 1990s. This development can be seen as a result of rapid urbanization and infrastructure development worldwide. The system produces electricity from fossil fuel centrally in a power plant, where electricity and data move in one direction. Thus, the existing grid system cannot collect and interpret real-time data about the services provided to customers, and it can be considered a huge waste of capacity loss and excess capacity caused by the hierarchy. The Smart Grid provides better "situational awareness" in relation to the state of the grid, using smart information and communication technologies in the existing grid. The areas of application of smart grids include: smart meters integration, demand management, smart integration of generated energy, administration of storage and renewable resources, using systems that continuously provide and use data from an energy network. Smart grids have emerged to address the shortcomings of existing unidirectional grid systems [2].

Table 2.1 - Comparison between existing grid and Smart Grid

|  |  |
| --- | --- |
| **Existing Grid** | **Smart Grid** |
| Electromechanical | Digital |
| One-Way Communication | Two-Way Communication |
| Centralized Generation | Distributed Generation |
| Hierarchical | Network |
| Few Sensors | Sensors Throughout |
| Blind | Self-Monitoring |
| Manual Restoration | Self-Healing |
| Failures and Blackouts | Adaptive and Islanding |
| Manual Check/Test | Remote Check/Test |
| Limited Control | Pervasive Control |
| Few Customer Choices | Many Customer Choices |

**2.3 Advantages of a Smart Grid**

With the introduction of the Smart Grid, the energy sector enter a new era of dependability, availability, and efficiency that will improve both our economic situation and the environment. To make sure that the advantages we anticipate from the Smart Grid become a reality throughout the transition period, it will be crucial to conduct out testing, technology advancements, consumer education, establishment of standards and laws, and information sharing among projects. The following are some advantages of the Smart Grid:

1. Improved electrical transmission

2) Quicker electricity restoration following power outages

3) Cheaper operational and management expenses for utilities, which would eventually result in cheaper power prices for customers.

4) Lower peak demand, which will also aid in bringing down electricity prices

5) More extensive incorporation of massive renewable energy systems

6) Improved customer-owner power generation system integration, particularly that of renewable energy sources

7) Greater security

These days, a blackout or other disruption of the energy supply can set off a domino effect, a chain of failures that can impact security, communications, banking, and other areas. Winter poses a special concern since homes may be left without heat. Our electric power system will become more resilient and better equipped to handle calamities like severe storms, earthquakes, huge solar flares, and terrorist attacks with the help of a smarter grid. The Smart Grid will allow for automatic rerouting when equipment breaks down or outages happen because of its two-way interactive capability. This will lessen outages and their consequences when they do occur. When there is a power loss, Smart Grid technology will be able to isolate it and contain it before it spreads to be a widespread blackout. The new technology will also make sure that, following an incident, electricity recovery resumes quickly and strategically, for example by first sending electricity to emergency services. Additionally, the Smart Grid will make greater use of customer-owned generators to provide electricity when it is not offered by utilities. A community may maintain the operation of its hospital, police force, traffic lights, phone system, and grocery store during emergencies by integrating these "distributed generation" resources. The Smart Grid is also a technique to deal with an outdated energy infrastructure that needs to be modernized or replaced. It's a strategy for addressing energy efficiency and raising consumer awareness of the link between power use and the environment. Additionally, it is a means of enhancing national security.

**2.4 Smart Grid Pilot Projects under IPDS (Integrated Power Development Scheme) in India**

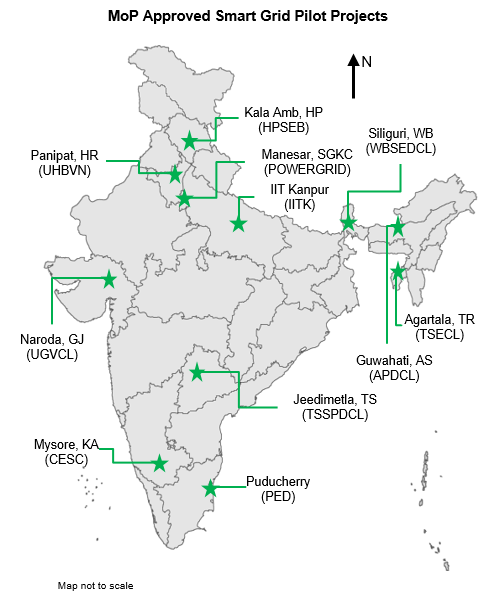


Figure 2.2 MoP approved Smart Grid Pilot Project in India.

**2.5 Smart Grid in the proposed work**

Smart grid technology is enabling the effective management and distribution of renewable energy sources such as solar, wind, and [hydrogen](https://innovationatwork.ieee.org/hydrogen-in-a-world-of-renewable-energy/). The smart grid connects a variety of [distributed energy resource assets](https://innovationatwork.ieee.org/several-states-are-embracing-distributed-energy-resources-to-achieve-clean-energy-targets/) to the power grid. In this work a Smart grid is developed using software. At first a power system modeling is done to represent a traditional grid. IEEE 5 bus is modeled to understand and represent a basic grid. After that a higher order bus consisting of 14 buses is developed which is upgraded to a smart grid. To upgrade the traditional 14 bus system to a smart grid system, the fact that a smart grid integrates renewable sources is utilized. Renewable energy sources are purposefully added to the 14 bus system to have a Smart grid environment. Solar PV system is added to bus 12, Wind energy system is added to Bus 8 and Battery Energy storage is added Bus 1 of the 14 bus system. Hence a smart grid is replicated. The next task is to have data generated from the smart grid to carry out the proposed work. Hence, electrical parameters such as voltage, current, power etc. are recorded over time from software run Smart Grid to have large set of data.

**CHAPTER 3**

**CYBER SECURITY**

**3.1 Cyber Security**

Cyber security professionals continually defend computer systems against different types of cyber threats. Cyber-attacks hit businesses and private systems every day, and the variety of attacks has increased quickly. Cybersecurity threats are acts performed by individuals with harmful intent, whose goal is to steal data, cause damage to or disrupt computing systems. Common categories of cyber threats include malware, social engineering, and man in the middle (MitM) attacks, denial of service (DoS), and injection attacks. Cyber threats can originate from a variety of sources, from hostile nation states and terrorist groups, to individual hackers, to trusted individuals like employees or contractors, who abuse their privileges to perform malicious acts.

**3.2 Cyber security Issues in a Smart grid**

A "system of systems" might be thought of as the smart grid, which will expand its existing generating, transmission, and distribution capabilities to include distributed generation, renewable energy sources, and electric vehicles. With the use of two-way information and communication technology (ICT), smart grids transfer power between producers and consumers while also enabling the control and automation of intelligent equipment. By seamlessly integrating high-speed metering and two-way networks into millions of pieces of power equipment, the Smart Grid, which blends ICT with existing power grids, intends to create a dynamic, interactive infrastructure with new energy management capabilities [6].

Smart grids, however, are susceptible to possible risks related to communication and networking systems because they largely rely on information networking. Although creating a dependable, secure, and optimized power system is the ultimate goal of the Smart Grid, doing so may paradoxically endanger the power system's ability to function.

There are various organizations that have done extensive research on the developments in cyber security objectives and requirements including Electric Power Research Institute (EPRI), National Institute of Standards and Technology (NIST), Smart Grid Interoperability Panel (SGiP) and IEEE.

The following is a description of the high-level security goals for safeguarding the smart grid's cyber infrastructure [7].

**Confidentiality:** Preventing unauthorized access by an enemy to highly secured information, such as power usage, price data, and control commands that invade customers' privacy and divulge the confidential information of utilities, is referred to as confidentiality.

**Integrity:** Integrity refers to the ability to prevent the alteration of crucial data from sensors, electronic devices (like smart meters), software, and control commands that might impair decision-making and taint data exchange of the smart grid. Unlike in case of confidentiality, integrity of software should be kept critical because an adversary might control any device or electrical equipment through compromised software.

**Availability:** Preventing an adversary from denying authorized personnel access to or control over the system is referred to as availability. Denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks can corrupt, delay, or block information, which prevents the smart grid from supplying power or exchanging information. Control command and price information must be available in this situation or there could be a loss of revenue.

Figure 3.1: Cyber security in Smart Grid

It is required to enhance the confidentiality, integrity and availability of the system by building a robust and efficient smart grid cyber infrastructure. Attack detection, mitigation, authentication and key management still remain challenging issues.

Typical Cyber Security Risks associated with the P&E Sector are:

1. Unauthorized access and breach of control systems  
   2)Interception and manipulation of control data/signals  
   3)Distributed/coordinated attack on system components  
   4)Interception and manipulation of monitoring data  
   5) Intentional and unintentional human intervention  
   6) Impairment to application software  
   7)Third-party intervention (interconnected partner, vendor)

Cyber Security is definitely a key component of an Operator’s Smart Grid deployment and of its Service Reliability strategy. The development of a Cyber Security program for the Smart Grid should not be an afterthought; it should be an integral part of the planning and design process involved with the deployment of Smart Grid initiatives. The Cyber Security program should also ensure that legacy systems receive the protection they require. A properly planned Cyber Security strategy will result in a highly secure environment that still delivers the operational flexibility and efficiency so crucial to the successful implementation of new Operational systems. Utilities should therefore implement a comprehensive, integrated, well monitored, and frequently updated Cyber Security program to ensure they derive the full benefits available from the Smart Grid.

It is recommended that the Cyber Security program start with a comprehensive Operational Risk Assessment. This Assessment should be specifically tailored to a P&E Operational environment as its needs vary greatly from those of a corporate Information Technology environment. The Assessment allows an Operator to identify the potential problem areas from an Operational perspective and to then formulate a strong Cyber Security strategy.

The first step in the Operational Risk Assessment is to conduct a Threat Risk Assessment (TRA) to determine the prioritization and focal areas for protection. After the TRA, the Operational Risk Assessment can include: an architecture review; an assessment of security devices, network devices, servers and workstations; a Cyber Security policy review; and a site audit, including a physical security audit. The Assessment can also be expanded to include the overall Cyber Security of an Operator, including its Information Systems, to ensure that overall productivity is protected from cyber-attacks [8].

**3.3 Artificial Intelligent techniques used for cyber-attack detection**

In early days Computer Security and AI were not connected to each other. Artificial Intelligence researchers were interested in developing programs to decrease human work, while security professionals trying to fix the outflow of information. But the two fields have grown closer over the time, when the attacks have targeted to simulate the genuine performance, not only at the human user level but also at lower system levels. CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a very good example of connection of artificial intelligence and security. This requires end-user to insert the letters of some unfair image, on some occasions with the addition of a masked sequence of letters or digits that appears on the screen. Improvements in automatic character recognition software, which can be considered to be a reasonable advance in AI technology, could motivate the field towards more refined pattern recognition. So in the practice of trying to secure properties, such as online ticket reservations, the profitable security market is in a way stimulating advances in artificial intelligence. Artificial Intelligence helps us in quickly identifying and analyzing new exploits and weaknesses to help ease further attacks and is an integral part of our solutions. Artificial Intelligence practices are the key to Interference detection and make it possible to respond even to anonymous threats before spreading itself [5].

**3.4 Machine learning for anomaly detection**

Machine learning, subset of Artificial intelligence is of rising importance in cyber security. To create intelligent security systems, it is essential to identify hidden trends and insights in network data and to build a corresponding data-driven machine learning model to stop these attacks. Machine learning can identify cyber security threats more effectively than other software-oriented methodologies, which lessens the workload on security analysts even though it cannot automate a full cyber security system. As a result, effective adaptive methods, such as various machine learning techniques, can lead to increased detection rates, decreased false alarm rates, and reasonable computation and communication costs [3].Anomaly detection is one of the most common use cases of machine learning. Finding and identifying outliers helps to prevent fraud, adversary attacks, and network intrusions. Machine learning algorithms are used to analyze the abnormal instances in a particular network. The algorithms can be trained for multiple data and can track the exploitation of a network [5].

Figure 3.2 Anomaly Detection Using Machine Learning

In this work unsupervised machine learning algorithm Isolation Forest is used for the cyber-attack detection in the communication network for the proposed Smart Grid network. The detailed methodology is described in the next chapter.

**CHAPTER 4**

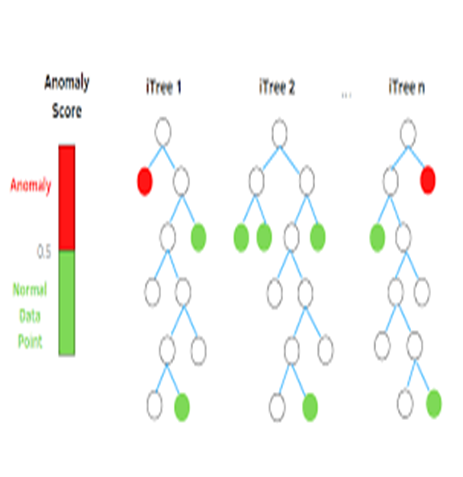
**METHODOLOGY**

**4.1 Proposed Methodology**

Isolation Forest uses random splits to recursively separate data in order to find abnormalities. Isolation Forest has grown in popularity since it was first developed as a quick algorithm for anomaly detection in a variety of industries, including cyber security, banking, and medical research etc.

Algorithm: Grows an extremely randomized tree without training labels, until every data point is isolated into a leaf of size 1.The depth of the final leaf is used as a proxy for how anomalous the data point is. The premise of the Isolation Forest algorithm is that anomalous data points are easier to separate from the rest of the sample [9].

Anomaly Score iTree 1 iTree 2 ..iTree n



Anomaly

0.5

Normal data point

Figure 4.1 Isolation Forest Tree

The *iForest* algorithm is based on the fact that anomalies are data points that are few and different from other data. Isolation Treeis a complete binary tree, with each node explicitly branched into zero or two child nodes.

## Isolation Forests- Anomaly Detection

## Isolation Forests (IF), similar to Random Forests, are build based on decision trees. And since there are no pre-defined labels here, it is an unsupervised model. In an Isolation Forest, randomly sub-sampled data is processed in a tree structure based on randomly selected features. The samples that travel deeper into the tree are less likely to be anomalies as they required more cuts to isolate them. Similarly, the samples which end up in shorter branches indicate anomalies as it was easier for the tree to separate them from other observations.

**Anomaly Score in Isolation Forest**

After an ensemble of iTrees (Isolation Forest) is created, model training is complete. During scoring, a data point is traversed through all the trees which were trained earlier. Now, an ‘anomaly score’ is assigned to each of the data points based on the depth of the tree required to arrive at that point. This score is an aggregation of the depth obtained from each of the iTrees.

Anomaly score in Isolation Forest algorithm is defined as [9]-

**S (x,n)** = **2 - E(h(x)) / c(n)**

where h(x) is the path length of observation x, c(n) is the average path length of unsuccessful search in a Binary Search Tree and n is the number of external nodes.

Each observation is given an anomaly score and the following decision can be made on its basis:

* A score close to 1 indicates anomalies
* Score much smaller than 0.5 indicates normal observations
* If all scores are close to 0.5 then the entire sample does not seem to have clearly distinct anomalies.

**Workflow:**

A Flowchart of the proposed data integrity assault detection scheme-

POWER SYSTEM SYNTHETIC DATASET

DATA PREPARATION

TEST DATA SUBSET

TRAINED MODEL

DIMENSIONALITY REDUCTION

(PREPARATION COMPONENT ANALYSIS)

OUTPUT : NORMAL OR ANOMALOUS DATA

TRAIN DATA

Subset

TRAINING

MODEL TRAINING AND TESTING PHASE

DIMENSIONALITY REDUCTION

Trained

Model

Fig. 4.2: Proposed work flow

**CHAPTER 5**

**CASE STUDY**

**5.1 Power System Modeling**

The largest technical professional society in the world is IEEE, a group committed to furthering innovation and technological excellence for the benefit of humanity. It is intended for professionals working in all facets of the science and technology fields relating to electrical, electronic, and computing as well as other fields that support modern society.

IEEE Bus system: There are numerous distinct sections that make up the electric power system. One of these is the transmission system, in which power is delivered to consumers via transmission lines from generating plants and substations. It is crucial to collect data continuously and in real time so that it can be analyzed and used to improve the power system. Many unexpected or undesirable situations may occur in transmission lines. To make the analysis of the line and bus parameters effective and simple, it is crucial that the bus parameters be displayed in a clear and lucid manner

**5.2 Case Analysis**

The IEEE system consists of loads, capacitor banks, transmission lines, and generators .For this work at first a smaller IEEE-5 bus, a scaled-down version of the transmission system is developed consisting of 5 buses to understand the modeling of power system using software. After that on an IEEE-14 bus, a scaled-down version of the transmission system is developed consisting of 14 buses. The IEEE 5-bus network  is composed of 2 generators, G1 and G2, located at buses 1 and 2 respectively, 3 load buses located at buses 3, 4 and 5 respectively.

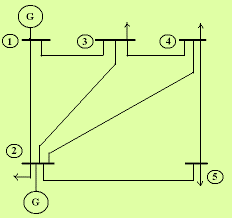
.

Figure 5.1 Single line diagram of IEEE 5 Bus System

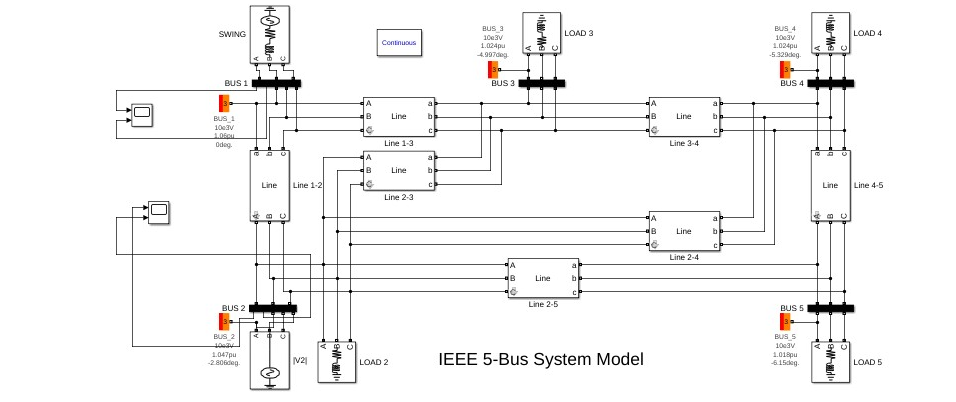


Figure 5.2 IEEE 5 Bus System Model

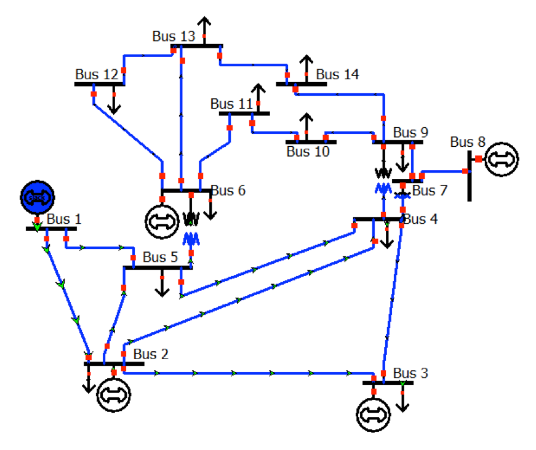
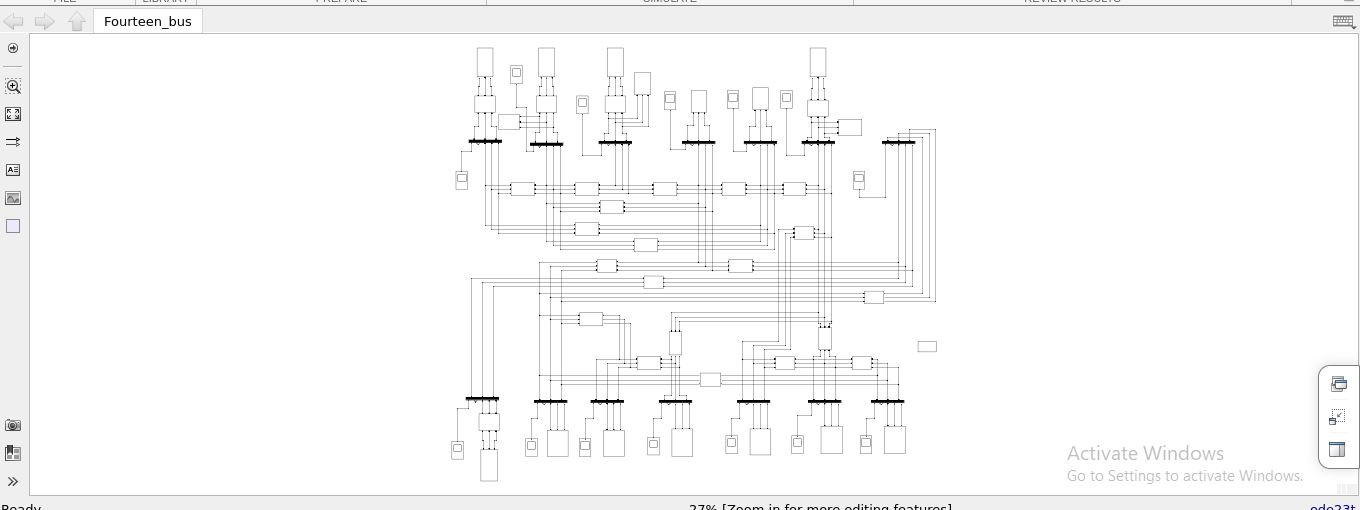


Figure 5.3 Single line diagram IEEE 14 bus system

 Figure 5.4 IEEE 14 Bus Model

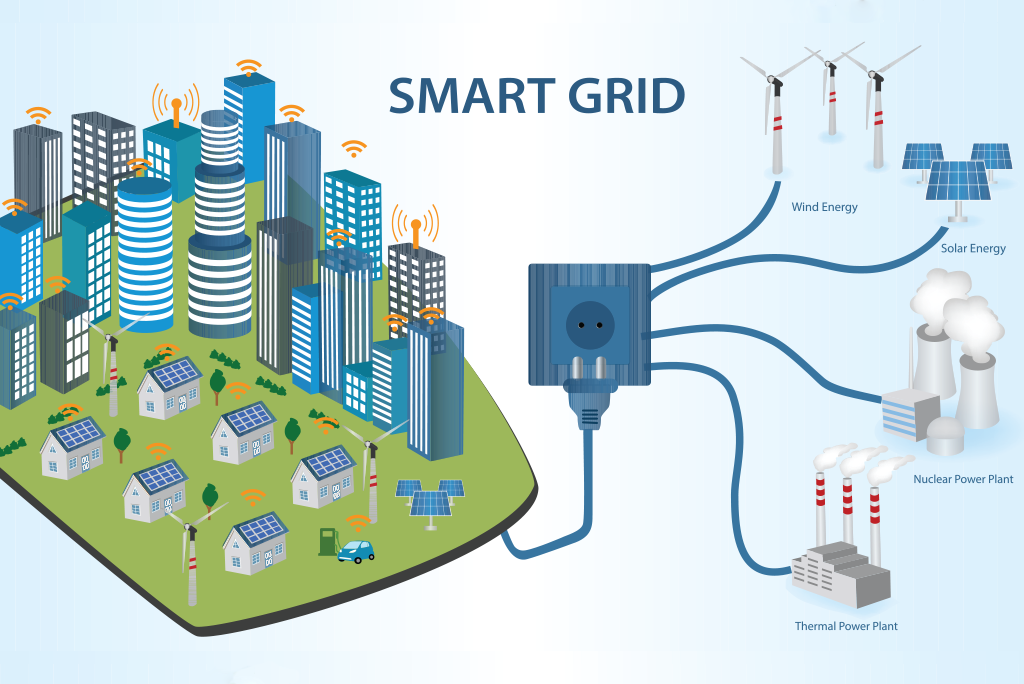


Figure 5.5 Grid with renewable energy components.

The system is upgraded to a Smart Grid. The Smart IEEE 14 bus system is modeled with subsystems -Battery Energy Storage System, Solar (PV) and Wind Energy.

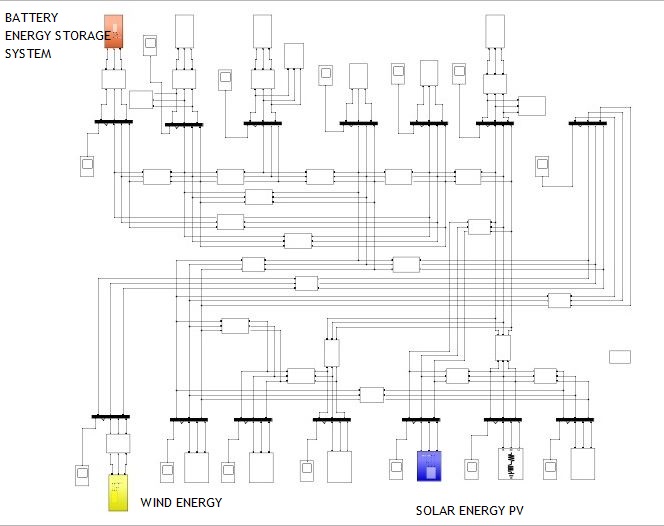


Figure 5.6: Smart Grid Model; 14 Bus System with renewable energy sources -

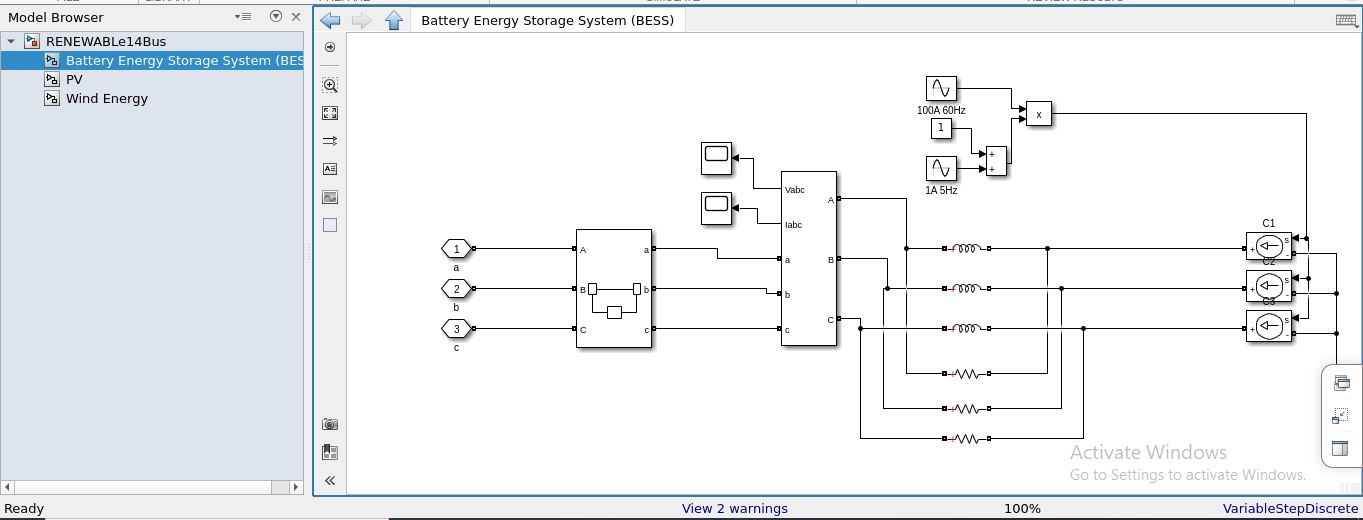


Figure: 5.7 A Battery Energy Storage System

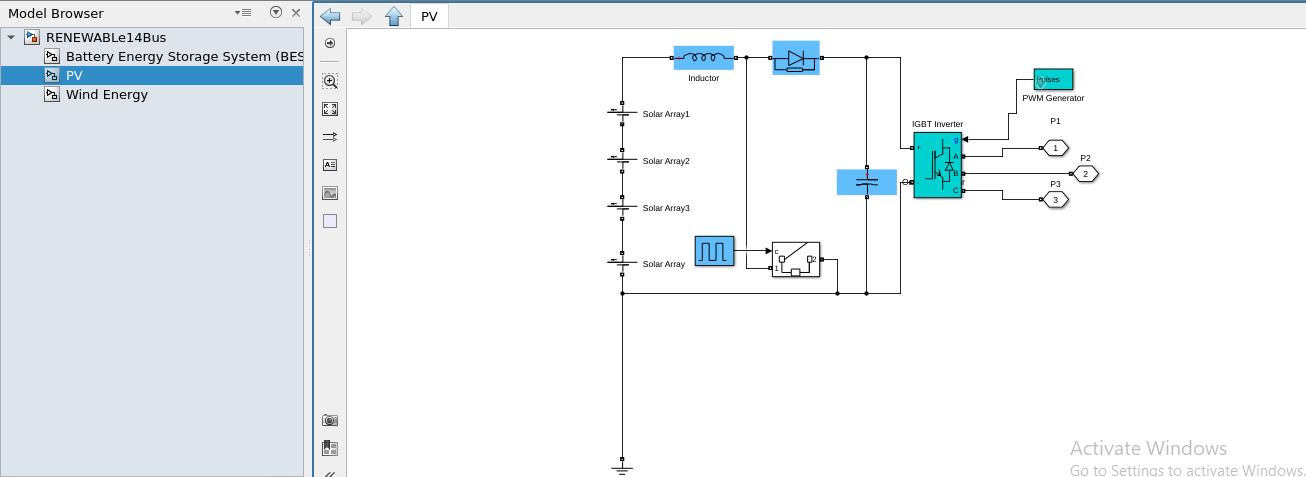


Figure: 5.8 B Solar PV

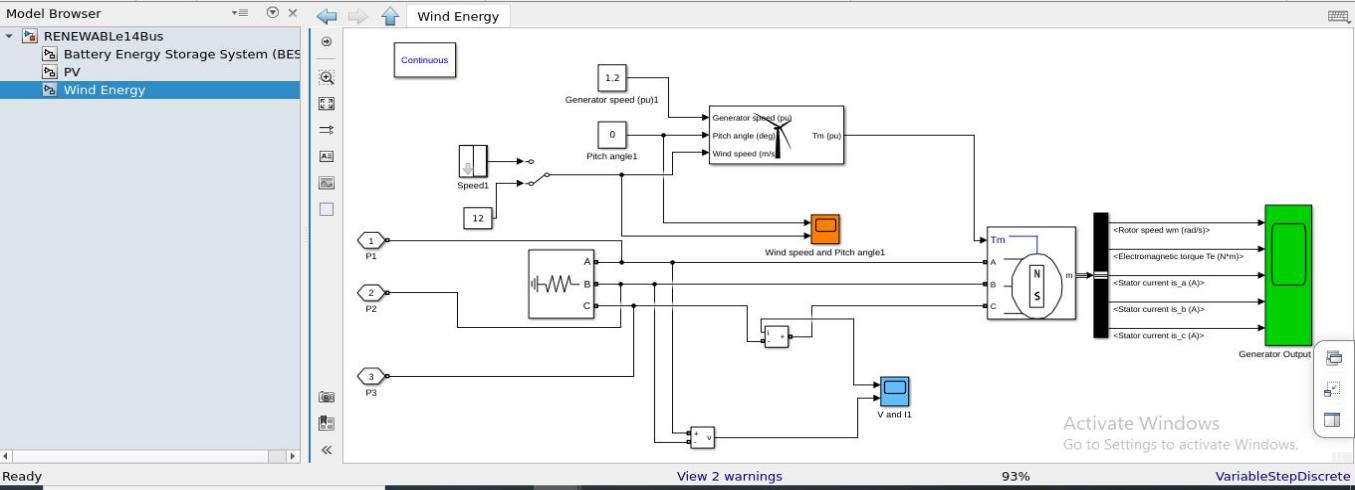


Figure 5.9 C Wind Energy

**CHAPTER 6**

**RESULTS AND DISCUSSION**

**6.1 A.GRID PARAMETERS OVER TIME**

Electrical parameters -Voltage, current, active power and reactive power are obtained for the different buses of Smart 14 Bus system. Simulation results of some buses are shown below.

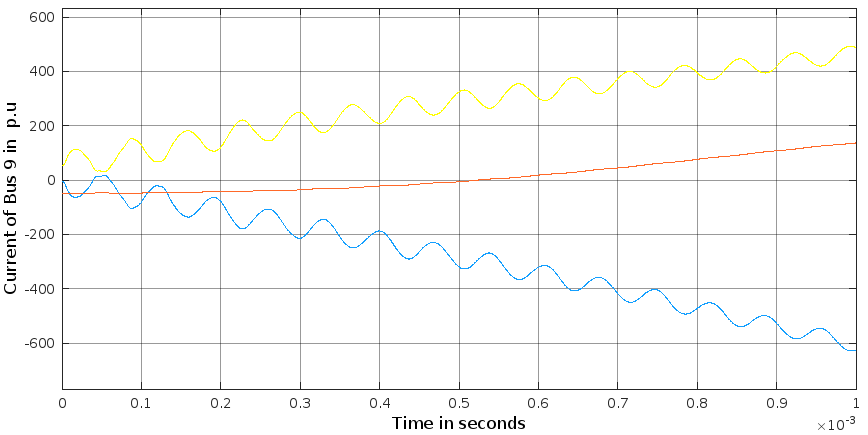


Figure 6.1: Current waveform of Bus 9

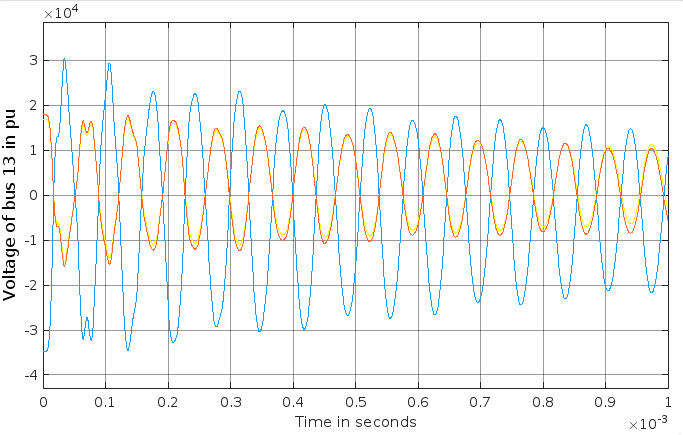


Figure 6.2: Voltage waveform of Bus 13

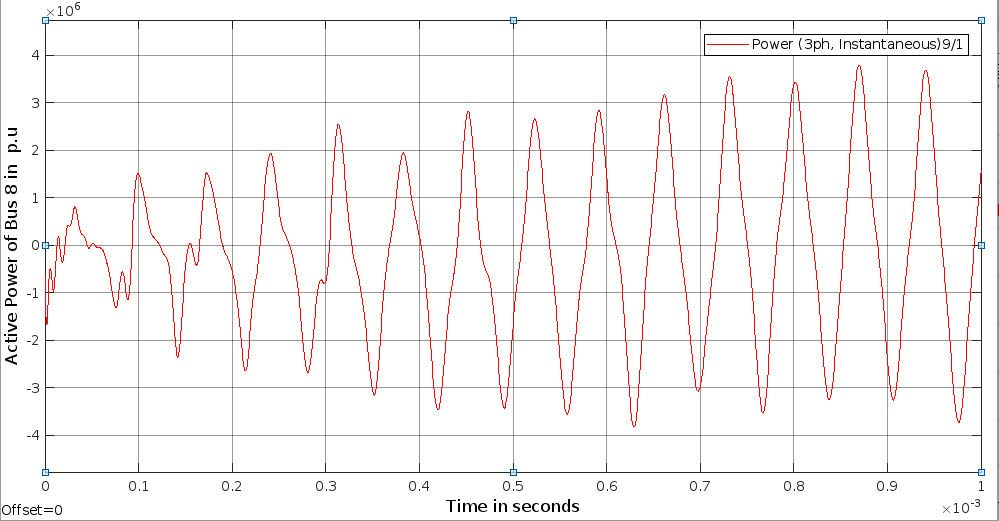
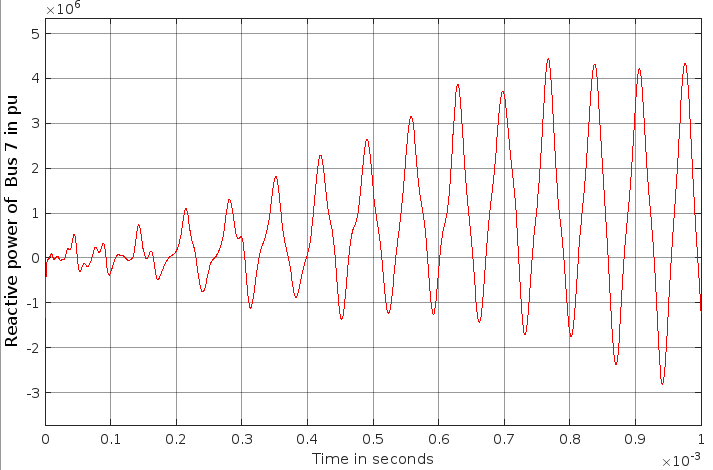


Figure 6.3: Active Power waveform of Bus 8

Figure 6.4 Reactive Power of Bus 7

Data obtained from the system is used for proposed scheme of cyber-attack detection. Subtle changes is deliberately introduced in normal data to create a scenario of cyber-attack. Data of ‘Current of Bus 9’ is used for the same.

*Attack:* A normal data set is obtained for Current of Bus 9. A small section of data is extracted .Cyber-attack is initiated by making subtle changes in the small section of output data.

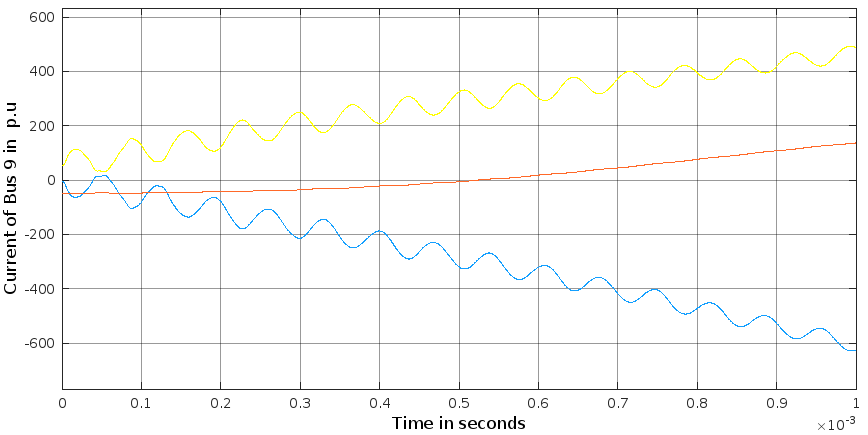
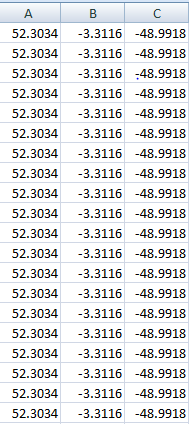
 

Figure 6.5: Current waveform of Bus 9 Figure 6.6: Normal Sample

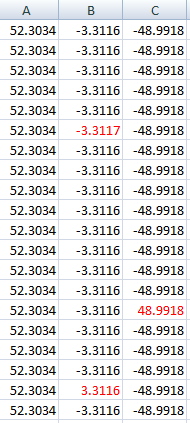


Figure 6.7 : Attacked Sample

Detection: Isolation Forest algorithm is applied to the attacked sample.

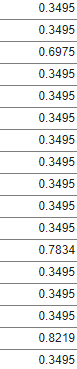
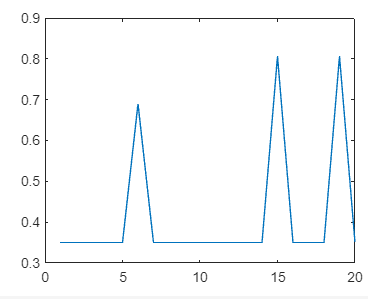
 

Figure 6.8 Anomaly scores Figure 6.9 0<score<0.5 Normal Point

obtained using Isolation Forest Algorithm. and 0<score<1 Anomalous Point

From the results obtained; three anomalous data points are detected based on the scores. A score less than 0.5 but more than 0 is considered a normal point. A score greater than 0.5 but less than 1 is considered anomalous point. Isolation Forests are computationally efficient and have been proven to be very effective in Anomaly detection. Despite its advantages, there are a few limitations such as the final anomaly score depends on the contamination parameter provided while training the model which means that an idea of what percentage of the data is anomalous beforehand is needed to get a better prediction. Also, the model suffers from a bias due to the way the branching takes place.

**CHAPTER 7**

**CONCLUSION**

In this work, vulnerabilities of the Smart Grid, different kinds of cyber security issues and attack has been discussed. Power system modeling is done using MATLAB software. Standard IEEE 14 bus has been modeled, upgraded to a smart grid by integrating renewable sources and simulation results are obtained .Subtle changes is deliberately introduced in normal data to create a scenario of cyber-attack. Effort is made to detect this manipulation using machine learning as an cyber security asset .Based on anomaly scores obtained using Isolation Forest algorithm on the corrupted data three anomalous data points are observed and hence the cyber-attack is detected .

In future, the analysis can be extended for the other electrical parameters obtained from the software based smart grid system. Moreover other artificial intelligence method can be studied and implemented for the cyber-attack detection. Different techniques can be compared and most effective method for cyber-attack detection in a smart grid can be put forward.

**APPENDIX 1**

**IEEE 5-BUS SYSTEM DATA**

Table A1.1 Bus Data for IEEE 5-Bus System

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Bus Code P** | **Assumed Bus Voltage** | **Generation** | | **Load** | |
| **Megawatts** | **Megavars** | **Megawatts** | **Megavars** |
| 1 | 1.06 + j0.0 | 0 | 0 | 0 | 0 |
| 2 | 1.0 + j0.0 | 40 | 30 | 20 | 10 |
| 3 | 1.0 + j0.0 | 0 | 0 | 45 | 15 |
| 4 | 1.0 + j0.0 | 0 | 0 | 40 | 5 |
| 5 | 1.0 + j0.0 | 0 | 0 | 60 | 10 |

Table A1.2 Line Data for IEEE 5-Bus System

|  |  |  |  |
| --- | --- | --- | --- |
| **Bus Code p – q** | **Line impedance *Zpq*** | | **Line charging**  *Ypq* / 2 |
| **R per unit** | **X per unit** |
| 1 - 2 | 0.02 | 0.06 | X per unit |
| 1 - 3 | 0.08 | 0.24 | 0.0 + j0.025 |
| 2 - 3 | 0.06 | 0.25 | 0.0 + 0.020 |
| 2 - 4 | 0.06 | 0.18 | 0.0 + j0.020 |
| 2 - 5 | 0.04 | 0.12 | 0.0 + j0.015 |
| 3 - 4 | 0.01 | 0.03 | 0.0 + j0.010 |
| 4 - 5 | 0.08 | 0.24 | 0.0 + j0.025 |

Table A1.3 MW Limits for Branches in IEEE 5 Bus System

|  |  |
| --- | --- |
| **Line** | **MW Limit (P.u)** |
| 1 - 2 | 0.8 |
| 1 - 3 | 0.3 |
| 2 - 3 | 0.2 |
| 2 - 4 | 0.2 |
| 2 - 5 | 0.6 |
| 3 - 4 | 0.1 |
| 4 - 5 | 0.1 |

**APPENDIX 2**

**IEEE 14-BUS SYSTEM DATA**

Table A2.1 Bus Data for IEEE 14-Bus System

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Bus No.** | **Bus Voltage** | | **Generation** | | **Load** | |
| **Magnitude**  **Per Unit** | **Phase Angle**  **Degrees** | **Real MW** | **Reactive MVAR** | **Real**  **MW** | **Reactive MVAR** |
| 1 | 1.060 | 0.0 | 232.4 | - 16.9 | 0.0 | 0.0 |
| 2 | 1.045 | - 4.98 | 40.0 | 42.4 | 21.7 | 12.7 |
| 3 | 1.010 | - 12.72 | 0.0 | 23.4 | 94.2 | 19.0 |
| 4 | 1.019 | - 10.33 | 0.0 | 0.0 | 47.8 | 3.9 |
| 5 | 1.020 | - 8.78 | 0.0 | 0.0 | 7.6 | 1.6 |
| 6 | 1.070 | - 14.22 | 0.0 | 12.2 | 11.2 | 7.5 |
| 7 | 1.062 | - 13.37 | 0.0 | 0.0 | 0.0 | 0.0 |
| 8 | 1.090 | - 13.36 | 0.0 | 17.4 | 0.0 | 0.0 |
| 9 | 1.056 | - 14.94 | 0.0 | 0.0 | 29.5 | 16.6 |
| 10 | 1.051 | - 15.10 | 0.0 | 0.0 | 9.0 | 5.8 |
| 11 | 1.057 | - 14.79 | 0.0 | 0.0 | 3.5 | 1.8 |
| 12 | 1.055 | - 18.07 | 0.0 | 0.0 | 6.1 | 1.6 |
| 13 | 1.050 | - 15.16 | 0.0 | 0.0 | 13.5 | 5.8 |
| 14 | 1.036 | - 16.04 | 0.0 | 0.0 | 14.9 | 5.0 |

Table A2.2 Line Data for IEEE 14-Bus System

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Line No** | **Between**  **Buses** | **Line impedance** | | **Half Line Charging Susceptance per unit** |
| **R per unit** | **X per unit** |
| 1 | 1 – 2 | 0.01938 | 0.05917 | 0.02640 |
| 2 | 2 – 3 | 0.04699 | 0.19797 | 0.02190 |
| 3 | 2 – 4 | 0.05811 | 0.17632 | 0.01870 |
| 4 | 1 – 5 | 0.05403 | 0.22304 | 0.02460 |
| 5 | 2 – 5 | 0.05695 | 0.17388 | 0.01700 |
| 6 | 3 – 4 | 0.06701 | 0.17103 | 0.01730 |
| 7 | 4 – 5 | 0.01335 | 0.04211 | 0.0064 |
| 8 | 5 – 6 | 0.0 | 0.25202 | 0.0 |
| 9 | 4 – 7 | 0.0 | 0.20912 | 0.0 |
| 10 | 7 – 8 | 0.0 | 0.17615 | 0.0 |
| 11 | 4 – 9 | 0.0 | 0.55618 | 0.0 |
| 12 | 7 – 9 | 0.0 | 0.11001 | 0.0 |
| 13 | 9 – 10 | 0.03181 | 0.08450 | 0.0 |
| 14 | 6 – 11 | 0.09498 | 0.19890 | 0.0 |
| 15 | 6 – 12 | 0.12291 | 0.25581 | 0.0 |
| 16 | 6 – 13 | 0.06615 | 0.13027 | 0.0 |
| 17 | 9 – 14 | 0.12711 | 0.27038 | 0.0 |
| 18 | 10 – 11 | 0.8205 | 0.19207 | 0.0 |
| 19 | 12 – 13 | 0.22092 | 0.19988 | 0.0 |
| 20 | 13 – 14 | 0.17093 | 0.34802 | 0.0 |

Table A2.3 Transformer Data

|  |  |  |
| --- | --- | --- |
| **Transformer** | **Between Buses** | **Tap Setting** |
| 1 | 4 - 7 | 0.978 |
| 2 | 4 - 9 | 0.969 |
| 3 | 5 - 6 | 0.932 |

Table A2.4 Shunt Capacitor Data

|  |  |
| --- | --- |
| **Bus Number** | **Susceptance Per Unit** |
| 9 | 0.190 |

Table A2.5 Regulated Bus Data (P – V Buses)

|  |  |  |  |
| --- | --- | --- | --- |
| **Bus No** | **Voltage Magnitude per unit** | **Reactive Power Limits** | |
| **Minimum MVAR** | **Maximum MVAR** |
| 2 | 1.045 | - 40.0 | 50.0 |
| 3 | 1.010 | 0.0 | 40.0 |
| 6 | 1.070 | - 6.0 | 24.0 |
| 8 | 1.090 | - 6.0 | 24.0 |

Table A2.6 MW Limits for Branches in IEEE 14 Bus System

|  |  |
| --- | --- |
| **Line** | **MW Limit (P.u)** |
| 1 – 2 | 0.6 |
| 2 – 3 | 0.7 |
| 2 – 4 | 0.8 |
| 1 – 5 | 0.5 |
| 2 – 5 | 0.4 |
| 3 – 4 | 0.3 |
| 4 – 5 | 0.2 |
| 5 – 6 | 0.5 |
| 4 – 7 | 0.4 |
| 7 – 8 | 0.2 |
| 4 – 9 | 0.2 |
| 7 – 9 | 0.2 |
| 9 – 10 | 0.2 |
| 6 – 11 | 0.3 |
| 6 – 12 | 0.2 |
| 6 – 13 | 0.2 |
| 9 – 14 | 0.2 |
| 10 – 11 | 0.2 |
| 12 – 13 | 0.2 |
| 13 – 14 | 0.2 |

**REFERENCES:**

[1] M. Z. Gunduz and R. Das, "Analysis of cyber-attacks on smart grid applications," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-5, doi: 10.1109/IDAP.2018.8620728

[2] H. Karimipour et al.: A Deep and Scalable Unsupervised Machine Learning System for Cyber-Attack Detection in Large-Scale Smart Grids,publication May 31, 2019

[3] E. Hossain et al.: Application of Big Data and Machine Learning in SG, and Associated Security Concerns ,January 24, 2019

[4]T. Halder, "A smart grid," 2014 6th IEEE Power India International Conference (PIICON), Delhi, India, 2014, pp. 1-6, doi: 10.1109/POWERI.2014.7117674.

[5]Arockia Panimalar.S1, Giri Pai.U2, Salman Khan.K3,Mar-2018,’Artificial intelligence techniques for cyber security’ *International Research Journal of Engineering and Technology (IRJET),*Volume: 05 Issue: 03 ,e-ISSN: 2395-0056 ,p-ISSN: 2395-0072

[6] M. M. Pour, A. Anzalchi and A. Sarwat, "A review on cyber security issues and mitigation methods in smart grid systems," SoutheastCon 2017, Concord, NC, USA, 2017, pp. 1-4

[7] Ding, Jianguo & Qammar, Attia & Zhang, Zhimin & Karim, Ahmad & Ning, Huansheng. (2022). Cyber Threats to Smart Grids: Review, Taxonomy, Potential Solutions, and Future Directions. Energies. 15. 6799. 10.3390/en15186799.

[8] R. V. Yohanandhan *et al.*: CPPS: A Review on Modeling, Simulation, and Analysis With Cyber Security Applications

[9] S. Ahmed, Y. Lee, S. -H. Hyun and I. Koo, "Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest," in IEEE Transactions on Information Forensics and Security, vol. 14, no. 10, pp. 2765-2777,Oct.2019, doi: 10.1109/TIFS.2019.2902822.