Security of Smart Cyber-Physical Grids: A Deep Learning Approach

by
Jacob Sakhnini

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Master of Applied Science
in Engineering
with Collaborative Specialization in
Artificial Intelligence

Guelph, Ontario, Canada

ABSTRACT

SECURITY OF SMART CYBER-PHYSICAL GRIDS: A DEEP LEARNING APPROACH

Jacob Sakhnini

University of Guelph, 2020

Advisor:

Dr. Hadis Karimipour

Co-Advisor:

Dr. Ali Dehghantanha

Cyber physical systems are widely used in critical infrastructure; among the most notable applications is the smart cyber-physical grid. The smart grid technologies are accompanied with various advantages including more efficient power generation and increased integration of green energy sources. As such, many cities around the world are investing in smart cyber-physical grid technologies. The use of this technology, however, comes with great risk to cyber threats. Furthermore, current state of the art defense methods lack in robustness, scalability, and computational efficiency. This thesis presents a deep learning based solution for attack detection in cyber-physical systems, particularly in the case of the smart cyber-physical grid. The research methods implemented in this thesis focus on improving robustness, scalability, and computational efficiency of intelligent attack detection algorithms by presenting heuristic methods for feature extraction and a novel deep learning approach that proved robust to varying attack sparsity and data imbalance.

Acknowledgements

The research performed in this thesis would not have been possible without the efforts of my co-supervisors, Dr. Hadis Karimipour and Dr. Ali Dehghantanha. I am grateful for their support and constructive feedback which aided in completing the research in a timely and organized manner. I am also grateful for the opportunity provided to me by Dr. Karimipour to join her research team and for her faith in me despite my lack of background in the subject. I am also grateful for Dr. Dehghantanha for allowing me to join his research team and learn from his expertise, which was immensely helpful to my academic progress.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The advancement of society is directed in the path of interconnected devices aimed at improving every-day life. Information and communication technologies (ICTs) have played a major role in shaping economic activities and urban infrastructure. Such exponential technological growth incited substantial buzz in the topics of integrating ICTs in urban development projects such as the smart grid and smart cities. Cities and communities today have embraced ICT in their development strategies utilizing digital infrastructure for regulatory and entrepreneurial purposes [1].

The use of smart technology goes beyond the applications obvious to the general public. Networked infrastructure, smart devices, and sensors are used in various other applications ranging from health-care to energy generation. Internet of Things (IoT) is the phenomenon referring to the integration of internet in various devices; such devices are used to increase the efficiency in a number of areas, including transport, health-care, and manufacturing [2]. This integration of cyber components into physical systems is a phenomenon known as Cyber Physical Systems (CPS). CPS are systems that operate on various levels through different layers. These layers are the physical layer, which consists of the physical components of the system, a sensor and actuator layer, a network layer, and a control layer. Sensors and actuator are used to communicate information between the physical components and the network, and the control layer is to send commands to the various aspects of the system.

A pivotal set of applications CPS and smart technology is in critical infrastructure. Sensors are used along city infrastructure and buildings for data collection to be used in more efficient modeling and prediction of likely outcomes. The smart grid system consists various resources and technologies. Smart meters are incorporated to collect consumption data for more efficient power distribution. Additionally, interconnection of supervisory control and

data acquisition (SCADA) allows for more expanded centralized distribution along large geographical areas [3][4][5]. The smart grid also allows for interaction among transmission and distribution grid, building controllers, as well as various sources of energy generation.

The concepts of smart meters, smart buildings, and smart grids are often discussed as the pinnacle of smart urbanization [6]. With data flowing across a city's infrastructure, relevant information can be used in various analysis, most notably efficient energy generation. Knowledge of energy consumption along a city's infrastructure enhances the predictive analysis of control centers, which in turn allow for more efficient energy distribution. Furthermore, the increased demand for green energy calls for a smart networked infrastructure capable of efficient use of energy sources. As such, the concept of the smart grid plays a major role in shaping the technological advancement of urban areas.

The integration of digital and information technology into the smart grid and the increased complexity of the system increases the possibility of cyber-attacks and failures propagating from one system to another [7]. As such, there are many challenges accompanying cybersecurity in the smart grid. Some examples include the difficulty modeling the nonlinearities and stochasticity of the system, as well as modeling the various types of cyber-attacks that can potentially inflict the system.

Additionally, many Advanced Persistent Threat (APT) actors and hacking teams are targeting critical infrastructure and services [8] ranging from health-care [9] and safety critical systems [10] to the smart grid. IoT technology, which can be defined as a network of physical devices connected to the internet, are increasingly used in critical infrastructure. The use of such devices can help the smart grid by supporting various network functions in power generation and storage as well as provide connectivity between supplier and consumers [11]. The integration of IoT devices in the smart grid also poses additional vulnerabilities to cyber-threats [12].

1.1 Motivations

Considering the complexity of the smart grid, and its vulnerabilities to cyber threats, various methods for cyber-attack detection have been proposed in literature. Model based solutions, such as variants of state estimation techniques and statistical-based models, have been suggested [13], [14]. However, intelligent systems have shown more promise when it comes to scalability to large, stochastic, real systems [15].

While there are many works in literature discussing intelligent methods for cyber attack

detection, many of which lack in robustness, scalability, and computational efficiency. As such, the research in this thesis aims at tackling these drawbacks by introducing automated heuristic feature selection algorithms for increased computational efficiency as well as deep learning regularization methods for increased robustness and scalability.

1.2 Objectives

In this thesis, the primary research goal is:

To study how machine learning can aid in cyber-attack detection in smart cyber-physical grids as an important example of critical infrastructure, to study how to maximize performance and efficiency of attack detection algorithms, and to develop a generalized novel neural network based attack detection algorithm robust and scalable to varying attack sparsity and data imbalance.

This thesis aims to provide an effective method for cyber-attack detection in smart cyber-physical grids. This is achieved through a combination of surveying literature and experimental analysis. The research goal of this thesis is accomplished through the contributions listed in the following section.

1.3 Contributions

Contribution 1: Developing a heuristic feature selection method for dimensionality reduction

Article: *J. Sakhnini, H. Karimipour, A. Dehghantanha, Smart Grid Cyber Attacks Detection using Supervised Learning and Heuristic Feature Selection, IEEE Int. Conf. on Smart Energy Grid Engineering (SEGE), pp.1-5, Oshawa, Canada Aug.2019.*

One of the main threats facing smart grid security are False Data Injection (FDI) attacks. FDI attacks are stealthy and undetectable by traditional bad data detection schemes currently employed in the majority of critical infrastructure. As such, FDI have been widely investigated in research. While there have been various types of solutions proposed to detect FDI attacks, machine learning is among the most common and most robust. One of

the main issues in the use of machine learning for detecting FDI attacks is computational efficiency. Real power systems are very large and have many measurements. As such, it can be computationally expensive to train a machine learning algorithm on a real system. For this reason, the first contribution of this thesis is the implementation and testing of heuristic feature selection algorithms to minimize the number of features/measurements used in training while maintaining accuracy. This reduction in the dimensions of the data allows for faster training of machine learning classifiers which can be used to detect attacks such as FDI.

This contribution is achieved through testing the accuracy of classifiers with and without feature selection. This contribution is considered complete if at least one of the heuristic methods results in a reduction of number of features by a minimum of 10% while maintaining accuracy or reducing it by no more than 2%.

Contribution 2: Developing a generalized deep-learning based cyber-attack detection algorithm for smart cyber-physical grids

Article: *J. Sakhnini, H. Karimipour, A. Dehghantanha, G. Srivastava Generalized Deep Neural Network for Attack Detection in the Smart Grid, IEEE Trans. on Emerging Topics in Computational Intelligence, pp. 1-8, Jan. 2020. Under review*

Further investigation of FDI attacks reveals challenges that are yet to be addressed. Among these challenges is the detection of FDI attacks in low sparsity; which are attacks that infect very few measurements in the system. Such attacks of low sparsity are very difficult to detect even with machine learning algorithms. Therefore, as a third contribution, a deep learning algorithm is proposed that overcomes the issue of poor detection at varying sparsity. Other benefits of this algorithm include the capability of detecting attacks when trained on imbalanced data, faster learning with minimal epochs of training, as well as superior generalization to larger systems.

This contribution is deemed complete upon achieving a higher accuracy and lower training time than a similarly structured neural network that lacks the proposed methods. This accuracy testing must be done over data-sets of varying attack sparsity and data imbalance. Averaging the accuracy and F1-score over all data-sets of varying attack sparsity must yield a higher accuracy in all test systems. The accuracy is considered to have increased, thereby completing the contribution, if it is increased by 2% or more.

Minor Contribution: Survey of Security Systems in Smart Grids

Article: *J. Sakhnini, H. Karimipour, A. Dehghantanha, A. Parizi, and G. Srivastava Security aspects of internet of things aided smart grids: a bibliometric survey, Internet of Things (IoT) August 2019.*

Proposing novel security methods necessitates an investigation of other methods proposed in literature. As such, a survey of security systems in smart grids is performed as a minor contribution of this research. This survey analyzes the types of threats that can harm the smart grid as well as the defense methods used to mitigate these threats.

1.4 Organization

The remainder of this thesis is organized as follows:

Chapter 2 provides background information on cyber physical systems, the smart grid, and supervised learning. The chapter also demonstrates the mathematical model for smart power systems used in this research.

Chapter 3 is a review of literature relevant to the research performed in this thesis. The chapter surveys journal and reputable conference papers in the field of cybersecurity of power systems. It begins by identifying the types of attacks and threats existing in literature. Then it delves into state of the art of security methods at all stages of defense. Finally, the chapter identifies some of the research gaps currently in this field of literature.

Chapter 4 demonstrates the proposed frameworks in which each contribution of this thesis. It discusses the heuristic methods used for feature selection as well as the regularization methods used for the deep learning algorithm.

Chapter 5 explains the methods in which the experiments were performed. It highlights the experimental process and explains the datasets, experimental process, and evaluation methods.

Chapter 6 demonstrates the results of the experiments performed in this research. It di-

vides the experimental procedure into two main experiments; the first tests the heuristic feature selection methods and the second tests the deep learning algorithm for attack detection.

Chapter 7 concludes this thesis by summarizing its contributions and suggests future work and improvements that can be done on this research.

Chapter 2

Background

This chapter provides the necessary background to understand the experiments performed in this research. The chapter begins by defining cyber physical systems and modeling smart cyber-physical grids, then it discusses supervised learning techniques used in the experiments.

2.1 Cyber Physical Systems

The integration of cyber components into physical systems is a phenomenon known as Cyber Physical Systems (CPS). CPS are systems that operate on various levels through different layers. These layers are the physical layer, which consists of the physical components of the system including sensors and actuators, a network layer, and a control layer. Sensors and actuator are used to communicate information between the physical components and the network, and the control layer is to send commands to the various aspects of the system. These layers are illustrated in figure 2.1.

CPS can be defined by its three major components: communication, control, and computation [16]. CPS are characterized by the following actions that they perform:

- Detection and capturing events or data such as pressure, temperature, presence of an object, electrical demand, user data, etc.
- Actuators or physical components that affect a physical process within the system.
- Interactions with other CPS.
- Evaluation of saved data.

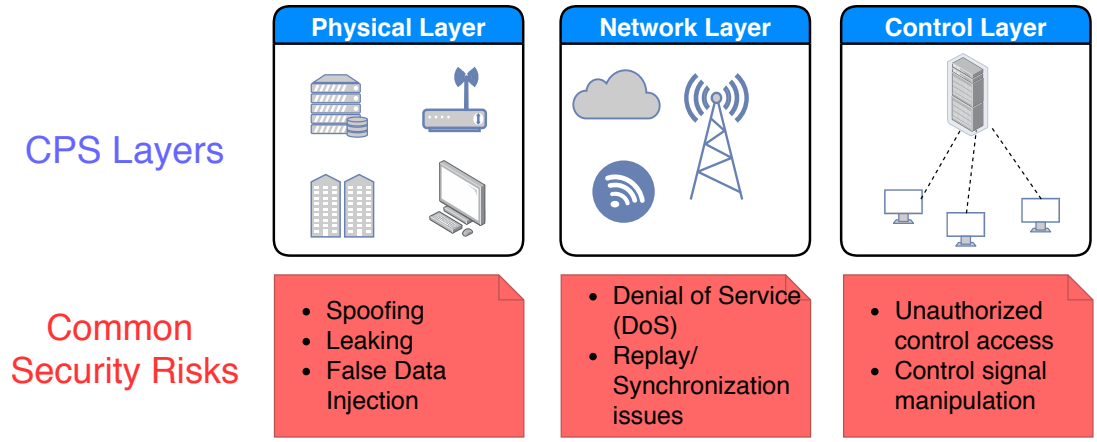

Figure 2.1: The four layers of CPS; which are communication, control, and computation

- Use of global data.
- Human machine interfaces [17].

These actions can provide great benefits for many industries. As such, CPS are used in a wide variety of applications including health-care, transportation, manufacturing, agriculture, energy generation and distribution, and other applications in critical infrastructure [18].

2.2 The Smart Cyber-Physical Grid

Among the most prominent and studied applications of CPS is the smart grid, the power systems of the next generation. The development of today's power systems is aimed towards integrating smart meters and sensors and advanced computing technologies to enhance the power generation efficiency [19]. The association of smart meters and sensors along the power grid network allows the generation centers access to real-time power demand information, which can be used to implement an efficient generation and distribution plan [20][21][5]. As such, integration of these technologies into the power system infrastructure has greatly increased the energy efficiency as well as reduced the price of electricity.

The smart grid system consists various resources and technologies. Smart meters are incorporated to collect consumption data for more efficient power distribution. smart cyber-physical grids are monitored and controlled by Supervisory Control And Data Acquisition

(SCADA) systems. The SCADA system works alongside the Advanced Metering Infrastructure (AMI) through a two-way communication that identifies detailed power consumption and distributes power accordingly. Additionally, SCADA allows for more expanded centralized distribution along large geographical areas [5][3][22]. The smart grid also allows for interaction among transmission and distribution grid, building controllers, as well as various sources of energy generation.

2.2.1 Modeling The Smart Grid

Smart grid can be modeled as a multi-agent CPS. The agents include generators, measurement devices, and control and generation agents [23][24]. The dynamic state of the system can be expressed as follows:

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\eta}) \tag{2.1}$$

where x is the system state, u is the output, η is the error term, and $f(x)$ describes the non-linear dynamic behavior of the system. Similarly, the static state of the system is represented by:

$$z = h(x, u, \varepsilon) \tag{2.2}$$

where z is the measurement vector, ϵ is an error term, and $h(x)$ represents the non-linear mapping of the system states to its measurements. Additionally, the fourth order model of generator i can be represented by [25][24]:

$$\begin{aligned}
\dot{\delta}_i =& \Omega_s \Delta\omega_i \\
\dot{\omega}_i =& \frac{\omega_s}{2H_i}\left(P_{Mi} - P_{Ei} - D_i\Delta\omega_i\right) \\
\dot{E}^{'}_{qi} =& \frac{1}{T^{'}_{di}}\left(-E^{'}_{qi} - \left(X_{di} - X^{'}_{di}\right) I_{di} + V_{fi}\right) \\
\dot{E}^{'}_{di} =& \frac{1}{T^{'}_{qi}}\left(-E^{'}_{di} + \left(X_{qi} - X^{'}_{qi}\right) I_{qi}\right) \\
E^{'}_{qi} =& V_{qi} + R_{ai} I_{qi} + X^{'}_{di} I_{di} \\
E^{'}_{di} =& V_{di} + R_{ai} I_{di} - X^{'}_{qi} I_{qi}
\end{aligned} \tag{2}$$

Table 2.1: Description of generator parameters

Parameter	Description
δ	rotor angle
$\Delta\omega$	rotor speed
Ω_S	system frequency
D	coefficient of damping
E'_d, E'_q	transient electromotive force in d-axis and q-axis
V_f	field voltage
H	machine inertia constant per unit
I_d, I_q	stator current in d-axis and q-axis
R_a	armature resistance
X_d, X_q	reactance in d-axis and q-axis
X'_d, X'_q	transient reactance in d-axis and q-axis
T'_d, T'_q	open loop time constant d-axis q-axis
P_E	electrical output torque
P_M	mechanical input torque

The electrical output for synchronous generator i can then be calculated as follows:

$$P_{Ei} = E'_{di}I_{di} + E'_{qi}I_{qi} + \left(X'_{qi} - X'_{di}\right) I_{di}I_{qi} \tag{2.3}$$

This can also be expressed in relation to other generators by:

$$P_{Ei} = \sum_{k=1}^{N} |E_i|\,|E_k| \left(G_{ik}\cos\left(\delta_i - \delta_k\right) + B_{ik}\sin\left(\delta_i - \delta_k\right)\right) \tag{4}$$

where E_i denotes the internal voltage of generator i, $G_{ik} = G_{ki}$ is the conductance between generators i and k, and $B_{ik} = B_{ki}$ is the susceptance between generators i and k. A diagram of a sample model used for the smart grid is shown in figure 2.2.

2.2.2 Monitoring The Smart Grid

smart cyber-physical grids are monitored and controlled by Supervisory Control And Data Acquisition (SCADA) systems. The SCADA system works alongside the Advanced Metering Infrastructure (AMI) through a two-way communication that identifies detailed power consumption and distributes power accordingly. This two-way communication is achieved through various size networks such as Home-Area Networks (HAN) which enable communi-

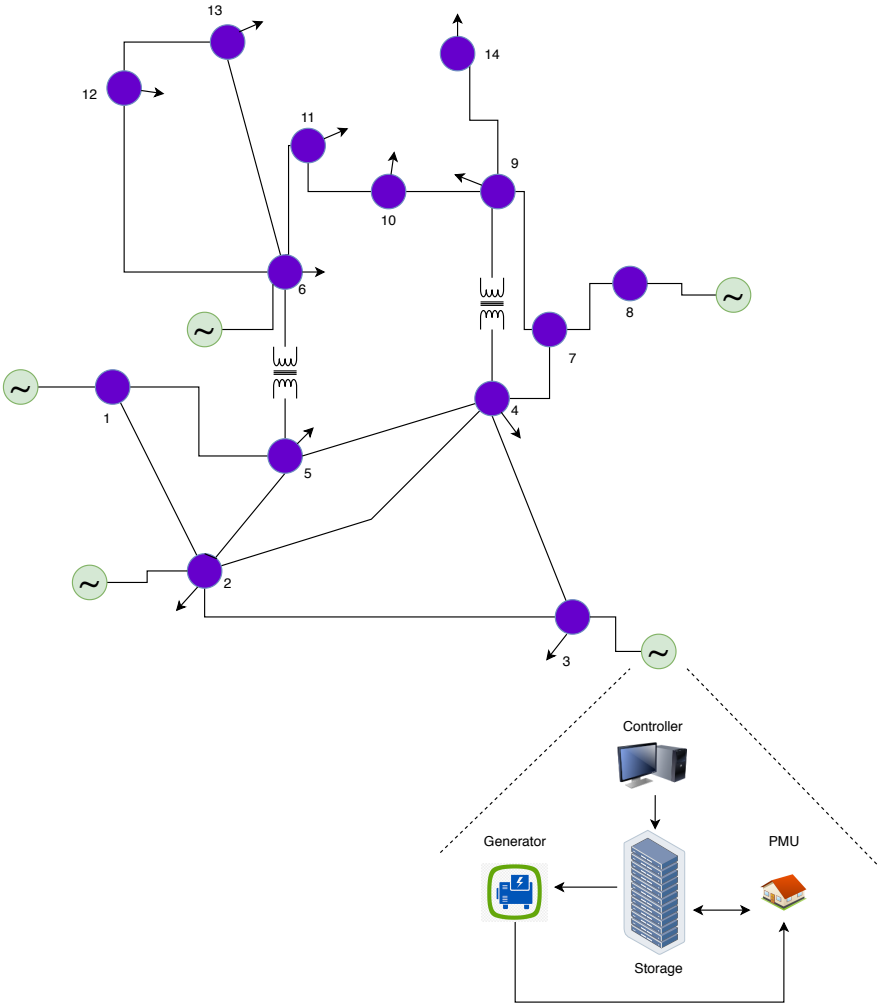

Figure 2.2: The IEEE 14-bus System

cation within a household, Neighborhood Area Networks (NAN) which enables secure flow of communication among households, and Wide-Area Networks (WAN) which connect all major components such as power stations, substations, and operation centers.

Other types of monitoring systems in smart grids include Wide-Area Situational Awareness (WASA) and Wide-Area Monitoring Systems (WAMS). These systems have the capability of real-time monitoring of power system components over large geographical areas. Furthermore, they are known to detect transient behavior not usually detected with traditional SCADA [26]. Several types of intelligent electronic devices (IEDs) are utilized in monitoring smart grids including Phasor Measurement Units (PMUs), Phasor Data Concentrators (PDCs), circuit break monitors, and more [27][28].

2.2.3 State Estimation

Power systems that employ smart grid technologies rely on state estimation to predict the state of the system which determines the optimal power generation. State Estimation is used for critical decision making, contingency analysis, and determining optimal power flow. This technique represents a relationship between the state variables of the system and the real measurements recorded along the power grid [29][5]. The measurement data consists of power flow, voltage magnitude and phase angles described as follows:

$$Z(k) = H(k)x(k) + \epsilon(k) \tag{2.4}$$

where Z represents measurement vector, x represents vector of state variables, H is the Jacobian matrix, and ϵ is the measurement error. k refers to the time step. The state estimation problem under the assumption of global observability can be formulated using the least squares method as follows:

$$\hat{x}(k+1) = \hat{x}(k) + G^{-1}(k)H(k)W^{-1}[Z(k) - H(k)\hat{x}(k)], \tag{2.5}$$

where gain matrix $G(k) = H^T(k)W^{-1}H(k)$. $\hat{x}$ is the vector of estimated states of the system. W is the co-variance matrix. To ensure optimal accuracy of the state estimation, measurement data will be checked to remove bad data [30]. Traditionally, bad data is detected through following 2-norm residual test:

$$\|z - Hx\|^2 < \varepsilon \tag{2.6}$$

where ε is the threshold for Bad Data Detection (BDD). If the residual of the measurements go above the predefined threshold bad data exist and should be removed before the next iteration. traditional BDD methods, however, fail to detect intelligent and stealthy attacks. This raises security concerns regarding monitoring the smart grid.

2.3 Security Risks in Smart Grids

Although many benefits result from the evolving smart grid technologies, the use of networked connections among these systems induces security risks. The integration of digital and information technology into the smart grid and the increased complexity of the system increases the possibility of cyber attacks and failures propagating from one system to another [7]. As such, there are many challenges accompanying cyber-security in the smart grid. Some examples include the difficulty modeling the non-linearities and stochasticity of the system, as well as modeling the various types of cyber attacks that can potentially inflict the system.

Many Advanced Persistent Threat (APT) actors and hacking teams are targeting critical infrastructure and services [31] ranging from health-care [9] and safety critical systems [10] to the smart grid. Furthermore, the rise of IoT technology can help the smart grid by supporting various network functions in power generation and storage as well as provide connectivity between supplier and consumers [11]. The integration of IoT devices in the smart grid also poses additional vulnerabilities to cyber-threats [12].

There have been several documented attacks on the electric grid attributed to cyber attacks. In January 2003, the computer network at the Davis-Besse nuclear plant in Oak Harbor, Ohio was compromised by a malware disabling its processing computer and safety monitoring system for several hours [32]. Similarly, circulation pumps at the Brown Ferry nuclear plant in Alabama failed due to excessive traffic, believed to be attributed to a DoS attack [32]. Furthermore, an investigation in 2009 revealed that hackers are able to steal power through compromising the smart meters and changing the consumption readings [33]. Phishing incidents have also been reported at electric bulk providers and malware samples were found indicating a targeted and sophisticated intrusion [32]. Additionally, in April of 2012, the FBI was asked to investigate widespread incidents of power thefts through smart meter attacks [33]. The report indicates that hackers changed the power consumption of smart meters using software available easily on the internet.

Such incidents in recent history induce various security concerns regarding critical in-

frastructure. As such, it is crucial that security of the smart grid is explored at every level including adequate situational awareness at all times. In fact, lack of situational awareness can have devastating impacts beyond cyber threats. For example, in August of 2003, a black-out occurred in the north east of the United States due to a cascading failure of the power system due to the lack of awareness of the Ohio-based electric utility company. This lack of awareness resulted in a cascading failure of 508 generators and 265 power plants across eight states and southern Ontario [33]. This clearly shows how adequate security systems can have benefits beyond mitigating cyber threats, including minimizing damage from faults or incidents.

smart cyber-physical grids can be exploited through several methods. The vulnerabilities of the smart grid are categorized based on the CPS layers as follows:

1. **Physical Layer:** The physical layer of the smart grid is vulnerable to the physical intervention from adversaries. Sensors throughout this layer are vulnerable to spoofing, leaking, and false data injection attacks.

2. **Network Layer:** Attacks on the network layer of the smart grid aim to compromise the communication channels. These attacks include replay attacks and Denial of Service (DoS) attacks. Replay attacks induce synchronization issues. Alternatively, DoS attacks jam the communication networks through numerous unauthorized request signals.

3. **Control Layer:** Attacks on the control layer typically propogate from other layers of CPS. This layer, however, is the most critical because gaining access to this layer can allow adversaries to have significant impact on the system. Furthermore, these attacks can lead to severe malfunctioning or physical destruction.

2.3.1 False Data Injection

Among the most common cyber-attacks discussed in literature are False Data Injection (FDI) attacks [15]. FDI attacks consist of malicious data injected into the measurement meters of the smart grid. FDI attacks can be performed by manipulating the measurements along the network by a linear factor of the Jacobian matrix of the system [34][35]:

$$Z_{bad} = Z + a \tag{2.7}$$

where a is an attack vector such that $a = Hc$ which results in

$$\|Z - Hx\|^2 = \|Z_{bad} - Hx_{bad}\|^2 + \Gamma \tag{2.8}$$

where Γ is an error term attributed to the state estimation that must remain within a certain threshold depending on the power system. This allows the attack to bypass the existing Bad Data Detection (BDD) methods such as Largest Normalized Residual (LNR) or a chi-square test [36]. Such a stealthy attack vector always exists even if the attacker has only partial access to the network topology [37].

In this thesis, and the majority of research in regards to FDI attacks, two assumptions are considered:

1. **Attack stealthiness:** There exists constant vectors, $\mathrm{a}_{\min}$ and $\mathrm{a}_{\max}$, where $\mathbf{a}_{\min} \preceq \mathbf{0} \preceq \mathbf{a}_{\max}$, such that the FDI attack vector a can pass the date quality check in BDD:

$$\mathbf{a} = \mathbf{Fc} \quad \text{and} \quad \mathbf{a}_{\min} \preceq \mathbf{a} \preceq \mathbf{a}_{\max} \tag{2.9}$$

   where c is an arbitrary vector and $\mathbf{x} \preceq \mathbf{y}$ means that each element of x is no greater than the corresponding element of y. It is assumed that the attacker knows F, $\mathrm{a}_{\min}$, and $\mathrm{a}_{\max}$ to construct a stealthy attack vector.

2. **Attacker's access to measurements:** It is assumed that the attacker has read access to all measurements in z and write access to a subset of the elements in z denoted by $\mathbb{W}$. Therefore, for any element j, the attack vector a is subject to

$$\mathbf{a}[j] = 0, \quad \forall j \notin \mathbb{W} \tag{2.10}$$

2.4 Supervised Learning

Supervised learning is the task of learning a function that maps inputs to outputs based on labeled training examples. This category of machine learning algorithms are extremely useful in a wide variety of applications. In cyber-security, supervised learning is widely used in threat detection. In this research, a variety of supervised learning techniques are used. These classification algorithms are used for attack detection and as cost functions to heuristic feature selection techniques. This section defines the algorithms used in this research.

2.4.1 Support Vector Machine

Support Vector Machine (SVM) is an algorithm that classifies data by constructing a set of hyper-planes in high dimensions [38]. SVMs are trained using an optimization function that relies on minimizing the hinge loss:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{n=1}^{N} \max\left\{0, 1 - y_n\left(\langle w, x_n\rangle + b\right)\right\} \tag{2.11}$$

where w is the weight vector, C is the penalty term, and x and y are the input and output respectively. The SVM optimization problem can also be expressed in the dual form in which the problem is independent of the number of features. The dual SVM is formulated as follows:

$$\begin{array}{ll} \min_{\xi,w,b} & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i \\ \text{s.t} & y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \xi_i \\ & \xi_i \geq 0; i = 1, \ldots, m \end{array} \tag{2.12}$$

To simplify the computations, kernel functions are used to represent the mapping of the data. In this study, a Gaussian kernel will be used for the SVM due to its non-linear properties and its capability of classifying data based on statistical variances with high computational efficiency. Mathematically, the Gaussian kernel is defined as follows:

$$K\left(x_i, x_{i'}\right) = \exp\left\{-\gamma\sum_{j=1}^{p}\left(x_{ij} - x_{i'j}\right)^2\right\} \tag{2.13}$$

where γ is the kernel coefficient. The SVM algorithm will be tested with varying penalty parameter, C, and kernel coefficient, γ, and cross-validated for accuracy.

2.4.2 K- Nearest Neighbor

K-Nearest Neighbor (KNN) algorithm classifies data based on its closest k neighbors. The closeness between the data is determined using the euclidean distance,

$$d_{ij} = \|\mathbf{s_i} - \mathbf{s}_j\|, \mathbf{s}_j \in S \tag{2.14}$$

where S and s correspond to labeled and unlabeled data respectively. For $k > 1$, data is classified based on majority of neighbors. In this study, various k values will be tested and

cross validated for accuracy.

2.4.3 Naive Bayes Classifier

The naive Bayes classifier is a probabilistic classifier based on Bayes' theorem. The naive Bayes acquired its name due to the strong or naive assumptions about independence among features. Using this statistical framework, the naive Bayes algorithm classifies an example $E = (x_1, x_2, \cdots, x_n)$ based on its probability of belonging to class c as follows:

$$p(c|E) = \frac{p(E|c)p(c)}{p(E)} \tag{2.15}$$

where E is classified as the class $C = +$ if and only if

$$f_b(E) = \frac{p(C = +|E)}{p(C = -|E)} \geq 1 \tag{2.16}$$

where $f_b(E)$ is called a Bayesian classifier. Additionally, the naive Bayes assumes all features are dependent, that is

$$p(E|c) = p\left(x_1, x_2, \cdots, x_n|c\right) = \prod_{i=1}^{n} p\left(x_i|c\right). \tag{2.17}$$

The resulting classifier is then:

$$f_{nb}(E) = \frac{p(C = +)}{p(C = -)} \prod_{i=1}^{n} \frac{p\left(x_i|C = +\right)}{p\left(x_i|C = -\right)} \tag{2.18}$$

where the function $f_{nb}(E)$ us called naive Bayesian classifier or simply naive Bayes (NB).

2.4.4 Artificial Neural Network

Artificial Neural Network (ANN) is an algorithm composed of interconnected elements, called neurons or nodes, which process information based on specific weights. ANNs can be constructed in various methods and architectures and typically consist of an input layer, hidden layers, and an output layer each consisting of several nodes. Each node i performs calculations represented by the transfer function f_i as follows:

$$y_i = f_i\left(\sum_{j=1}^{n} w_{ij}x_j - \theta_i\right) \tag{2.19}$$

where y_i is the output of the node i, x_j is the j^{th} input to the node, w_{ij} is the connection weight between nodes i and j, and θ_i is the bias of node i.

ANN can be constructed in various methods and architectures. In this study, the feed-forward architecture, shown in figure 2.3, is used. The feed-forward architecture typically consists of an input layer, hidden layers, and an output layer each consisting of several nodes. Each of the input nodes contains a feature of the data; these nodes are activated through various types of activation functions which process the information into the next layer of nodes. This activation process occurs in every layer until the data is classified in the output layer of the ANN.

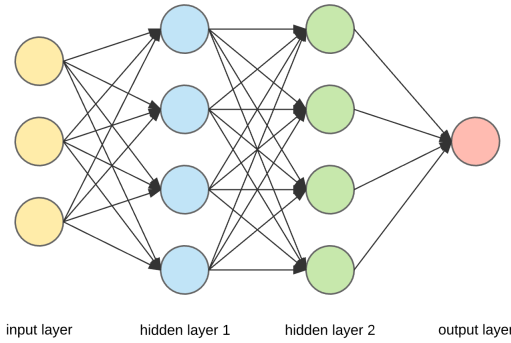

Figure 2.3: The general architecture of feed-forward ANNs

2.5 Summary

This chapter explains the background necessary to understand the remainder of the thesis. It begins by introducing cyber physical systems and the concept of the smart grid, then discusses security risks in the smart grid. The chapter also explains how the smart grid and stealthy data injection attacks are modeled in this research. Finally, the chapter discusses the supervised learning methods used in this research for the purpose of attack detection.

Chapter 3

Literature Review

This thesis proposes an attack detection solution for smart cyber-physical grids. Proposing an effective solution to this problem necessitates a thorough survey of related works. This chapter discusses the types of cyber-threats that smart cyber-physical grids face as well as the techniques to tackle these threats proposed in literature. The information discussed in this chapter is the result of a bibliometric analysis performed on all journal articles from 2010 to May 2019 gathered from Web of Science, Science Direct, and IEEE Xplore with the following search query:

("Smart Grid" AND "Cyber Security" OR "Cyber Attack" OR "Cyber Threat" OR "False Data Injection" OR "Attack Detection")

3.1 Types of Cyber Threats

Cyber threats or cyber attacks are among the most discussed and studied threats for the smart grid [39]. The wide interest in studying cyber threats in the smart grid is due to the number of significant vulnerabilities identified [40]. Furthermore, cyber attacks have the potential of leading power systems into total collapse [41]. These cyber attacks can occur for various purposes and are generally divided into two main types: Passive Attacks and Active Attacks [42]. Passive attacks include eavesdropping, spying, and traffic analysis; while active attacks include denial of service (DoS) and FDI attacks.

The various types of attacks are not equally studied in literature. Figure 3.1 shows the number of articles studying each type of attack. While there are more types of cyber threats that can compromise a network, the following sections discuss the attacks studied in the

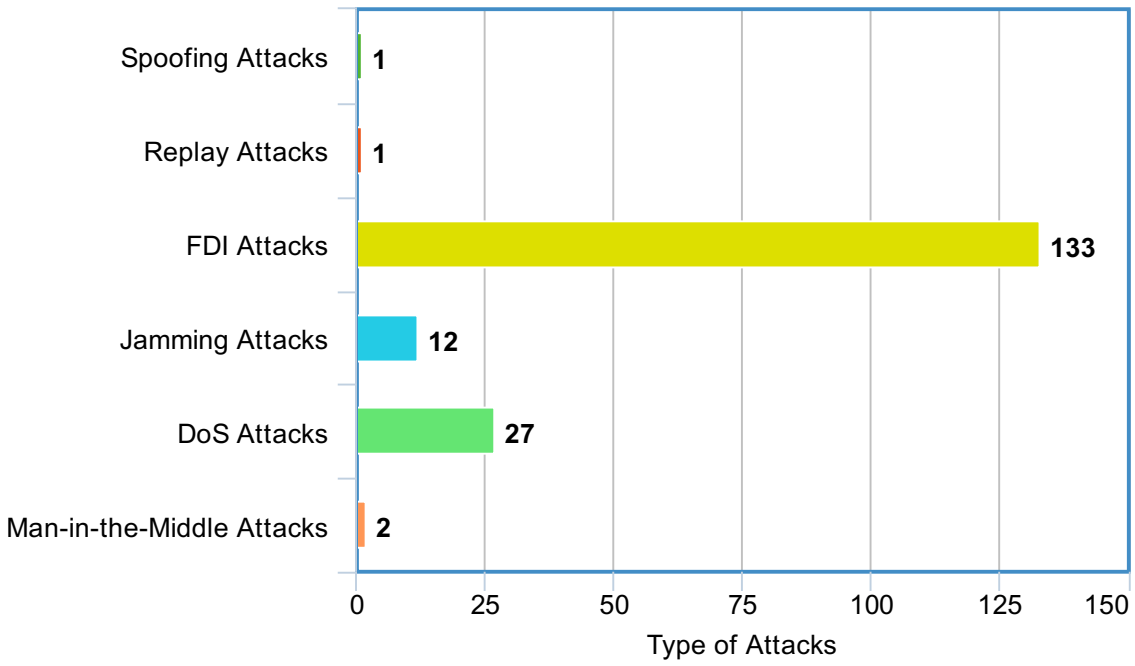

Figure 3.1: The number of journal articles studying each attack type

smart grid, which are mentioned in Figure 3.1.

3.1.1 Spoofing Attacks

The main types of spoofing are GPS spoofing, ARP (Address Resolution Protocol) spoofing, and IP spoofing [43]. IP spoofing uses a modified IP to pass through security systems and is typically the first stage of a complex intrusion. GPS spoofing, however, is based on broadcasting incorrect signals of higher strength than received from satellites to deceive victims. ARP spoofing is where falsified ARP messages are used to link the attacker's MAC address with the IP address of the victim. Through this all data in the compromised system will pass through the intruder. The most common type of spoofing attacks in the smart grid is GPS spoofing due to the use of GPS receivers in the metering infrastructure. Vulnerability analysis in literature demonstrates how Phasor Measurement Units (PMUs) are susceptible to GPS spoofing attacks [44]. GPS spoofing attacks can mislead the network operator, and drastically impact subsequent corrective control actions [45].

3.1.2 Replay Attacks

Replay attacks aim to intercept authentication information. In the smart grid, replay attacks intercept the usage pattern along the varying smart meters and replay this data to carry out an undetected intrusion [46]. The integration of IoT devices in smart grid networks induces increased threat to these attacks. Furthermore, attacker can inject incorrect data to the system, which may lead to incorrect energy price or inaccurate prediction [47].

3.1.3 Man-in-the-Middle Attack

This attack makes use of ARP, which maps a protocol address to a hardware address (MAC address) [48]. The purpose of this attack is to combine the attacker's MAC address with the host's IP address triggering any traffic meant for that particular IP to be sent to the attacker instead, this is referred to as ARP spoofing [49]. This allows the attacker to capture the communication information within the SCADA system [50].

3.1.4 Smart Meter DoS Attacks

DoS attacks are typically achieved by flooding specific nodes of the system with data that prompts generating and sending large volume of reply and request packets [51]. There are various methods for generating such attacks which can cause a system blackout [52]. These attacks can also be implemented through IoT devices integrated into the smart grid. The increased integration of these IoT devices has led to increased interest in DoS attacks [53].

3.1.5 False Data Injection Attacks

FDI attacks consist of malicious data injected into measurement meters [54]. FDI attacks can be performed by manipulating the measurements along the network by a linear factor of the Jacobian matrix of the power system [55, 56]. This change in measurement is undetected by the current state estimation techniques [57]. Furthermore, these attacks can be created in various strategies with limited knowledge of power system topology [58, 59, 60]. As such, these types of attacks are widely studied in the smart grid cybersecurity field [54, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67].

3.1.6 Micro-Grid-Based Jamming Attack

This type of attack consists of jamming specific signal channels to intervene and disrupt data transmission [68]. This results in unreliable communications and decreased performance in the power system [69, 70].

3.2 Detection and Mitigation of Cyber Attacks

Security and defense against the aforementioned attacks and threats is achieved through various mechanisms. The security measures proposed in literature are divided into the "7D model" or the 7 phases of cybersecurity as given in [8]:

- Discovery
- Detection
- Denial
- Disruption
- Degradation
- Destruction

The following subsections will discuss each of the components and their proposed methods in literature.

3.2.1 Discovery

The discovery process in cybersecurity involves identifying and locating sensitive data for adequate protection. In general applications of cybersecurity, data discovery consists of auditing regulated information to ensure its protection. This is helpful because it enables context aware security, in which information within the system is protected based on its sensitivity. In smart grid security, the discovery phase mainly consists of identification of vulnerabilities within the system.

Various methods are proposed in the literature for vulnerability analysis in power grids. One paper proposes an automated binary-based vulnerability discovery method that extracts security-related features from the system [71]. This automatic discovery algorithm is tested

on real smart meter data from Korean infrastructure. Vulnerability analysis specific to certain types of environments or threats are also proposed. In another paper, the survivability of smart grid under is modeled under random and targeted attacks considering a networking islanding scheme [72]. Another paper uses Automatic Static Analysis (ASA) to detect buffer-overflow vulnerabilities of terminal devices [73]. Such automated techniques for vulnerability analysis can be useful due to their robustness and scalability to larger systems. As such, a comprehensive assessment of vulnerabilities in the smart grid from past to future has been published highlighting the various vulnerabilities and discovery techniques [74].

More specific vulnerability modeling is also proposed in literature. One such work models the vulnerabilities of the smart grid with incomplete topology information [75]. The results of this paper demonstrate the high level of threat in the smart grid by exhibiting vulnerabilities that can be exploited with limited knowledge of the system. Another paper reveals the cascading failure vulnerability in the smart grid using a novel metric, called risk graph, which shows the importance of nodes within the system as well as the relationship among them [76]. Using this method, Zhu *et al.* develop a new node attack strategy and introduce new vulnerabilities not considered before in literature.

Vulnerability analyses are also performed on specific attacks. One paper performs a vulnerability analysis of the smart grid to GPS spoofing, a type of attack capable of altering measurements to mislead network operators [44]. Another paper analyzes the vulnerability for simultaneous attacks in the smart grid [77]. Paul and Ni consider various combinations of attacks and proposes a new damage measurement matrix to quantify the loss of generation power and time to reach steady-state. Web-based threats are also considered in another paper which tackles the penetration of digital devices in the smart grid and the associated consequences [78].

Most articles assess the vulnerability of the smart grid by analyzing either substations or transmission lines. One article, however, performs a vulnerability assessment on a joint substation and transmission line system in which attacks can happen in either the substation, the transmission line, or both [79]. Another article takes into account scenarios of severe emergencies in the smart grid and SCADA network and performs a vulnerability analysis of the system under emergencies such as attacks from weapons of mass destruction (WMD) [80]. Chopade and Bikdash analyze structural vulnerabilities, which consider infrastructures topology, and functional vulnerabilities, which consider operating regimes of different infrastructures.

As demonstrated by the aforementioned articles, there is sufficient analysis on vulnerabil-

ities in the smart grid. Various attack strategies are identified and implemented in literature that demonstrate the potential of cyber threats. Detection and mitigation of some of these threats remains as a gap in this research field. Next, we discuss the detection mechanisms proposed as well as the future trend in these methods.

3.2.2 Detection of Attacks

Detection of cyber threats is typically achieved through classification using data or measurements across the power system. Measurements along various infrastructure and communication layers of the system are used to detect the presence of threats or attacks. Model-based techniques are used to detect cyber attacks through meter measurements through enhanced state-estimation techniques [13, 4, 21]. Furthermore, distributed algorithms are used to find statistical variations in cyber attack vectors [81]. Kalman filters are also used to estimate measurements along the power system along with statistical methods of finding anomalies in measurements [81, 82, 5].

Other attack detection techniques stemming from the field of control theory revolve around secure state estimation. While most utilize Kalman filters, one paper proposes a search algorithm based on Satisfiability Modulo Theory (SMT) to increase the search speed for possible sensor sets [83]. Locating the attack through control strategies has also been proposed in literature. One paper proposed a framework in which the attack location can be determined given a total number of monitoring sensor equal to twice the number of compromised sensors [84]. Another paper proposes a control system to prevent zero-dynamic attacks, which occur by compromising the actuators instead of the sensors [85]. A more robust state estimator tackling attacks in the control signal is also proposed in [84]. This method adopts the "frequentist" approach in which no known priors are assumed.

Modbus-based detection is also utilized by Hadziosmanovic et al. [86]. They demonstrated how Modbus, an industrial communication protocol, can be used to detect attacks by monitoring the state variables of the system. Another paper, however, demonstrates the vulnerability of Modbus rptocol to flooding attacks [87]. In fact, several papers demonstrate attack implementation for Modbus highlighting its vulnerability to various attacks. Chen et al. proposed a realtime cyber-physical test-bed integrating communication system and power system simulators [88]. They also demonstrated its vulnerability to cyber-attacks by successfully deploying man-in-the-middle and flooding attacks. Another paper also demonstrates these attacks in addition to replay attacks and propose a novel role-based access

control model (RBAC) for secure authorization [89].

For defense methods to be scalable to larger systems, purely model-based attack detection techniques are insufficient to guarantee the security of the smart grid [90, 55]. As such, the use of intelligent systems and machine learning for detecting cyber attacks is proposed. Supervised and unsupervised learning have been tested and compared to conclude that supervised learning approaches generally result in more accurate classification of attacks [91]. Various supervised learning algorithms have been successfully implemented [92, 93]. The results of comparing these learning algorithms demonstrate that a Gaussian-based Support Vector Machine (SVM) is more robust with more accurate classification among larger test systems [93]. Furthermore, another paper implemented the margin setting algorithm (MSA) demonstrating better results than SVM and ANN [94, 55]. Other intelligent techniques include adaboost, random forests, and common path mining method [95, 96, 97].

A critical concern in the use of intelligent systems in smart grid is computational efficiency [98, 99]. Many researchers try to tackle this issue by reducing the dimensions of the data through principal component analysis [91, 92]. One paper proposes the use of a genetic algorithm to select an ideal subset of features that can increase the computation speed while maintaining the detection accuracy of the machine learning algorithms [100]. Exploring various feature selection techniques can be effective at increasing the computational efficiency of machine learning algorithms. However, there have not been many papers exploring this subject in the area of smart grid cybersecurity. As such, deep learning techniques with automated or unsupervised feature selection methods are likely to be proposed to tackle the computational burden of larger power systems.

3.2.3 Denial of Attacks

One of the security methods in the smart grid revolves around the denial or prevention of cyber threats. Denial techniques pertaining the security of the smart grid typically take the shape of encryption methods for secure communications within the system [69, 101]. The most common encryption methods are the use of symmetric or asymmetric keys. Symmetric keys use the same key to encrypt and decrypt the messages while asymmetric keys use different keys for encryption and decryption [70, 102]. Asymmetric key encryption requires a larger computational capacity and is therefore not suitable for time-sensitive information. Symmetric key encryption does not induce significant computational delay. However, it requires a public infrastructure for key management. Therefore, it is suitable for encryption

of distribution and transmission systems [103, 104, 105].

Various encryption and key management methods have been proposed. One scheme is based on Needham-Shroeder authentication protocol and elliptic curve cryptographic algorithms for generating public keys [106]. Another scheme uses digital certificates to establish symmetric communication sessions [103]. Additionally, another authentication method is proposed that is based on S/KEY one-time password scheme aimed to provide mutual authentication between the meters and servers of the smart grid [107]. Mutual authentication between smart grid utility network and Home Area Network (HAN) smart meters is also explored through a novel key management protocol [108]. The proposed mechanism aims at preventing various attacks including Brute-force, Replay, Man-in-The-Middle, and Denial-of-Service attacks. Furthermore, encryption of specific variables and measurements is also studied, specifically pertaining to FDI attacks [109, 110].

Choosing appropriate key management schemes is done by considering the trade-off between security and computational efficiency. However, other issues pertaining denial of attacks arise from the distributed nature of smart grid systems. One paper proposes an efficient framework to read isolated smart grid devices that satisfies the hardware constraints while maintaining integrity against most typical attacks [111]. Another protocol is proposed for preserving privacy through aggregation of metering data in distributed scenarios and encryption of measurements using a secret sharing scheme [112].

Other denial techniques are proposed in literature include increasing situational awareness to prevent attacks. One paper proposes specific measures to tackle issues that lead to lack of awareness among smart grid operators. Such measures include separate networks for actuators and sensors and restricting the use of real time clocks to write-only data storage [113]. Another paper proposes a different proactive defense approach which consists of randomizing meter infrastructure configurations to lower the predictability of the system to potential adversaries [114]. While there are many approaches to deny or prevent cyber threats, further research is likely necessary due to the continuous improvement and modifications of adversarial techniques.

3.2.4 Disruption of Attacks

A critical part of the security of any system is the disruption of cyber threats once the system is infected. Disruption of attacks in the smart grid is typically tackled by game theory approaches. One paper demonstrates disruptive countermeasures to reduce the impact of

attacks based on the knowledge of non-compromised components [115]. Similarly, another paper demonstrates how informed decisions can be made in real-world scenario of attacks to mitigate or disrupt them [116]. This is done by using a sequential two-player game model that includes attacker/defender behavior. Similarly, another article attempts to achieve the same goal by making use of the Stackelberg competition, which quantitatively analyzes the game process between attacker and operator [117]. A linear game framework is also proposed with the emphasis on application to large power systems with large number of components under attack [118].

Disruption of attacks through game theory is also studied under varying circumstances. One article considers coalition attacks that can be launched by multiple adversaries [119]. A game-theoretic model is proposed to capture the interaction among the adversaries and quantify the capacity of the defender based on Iterated Public Goods Game (IPGG) model. Similarly, stochastic games for protection against coordinated attacks is also proposed in [120]. This method uses an optimal load shedding technique to quantify physical impacts of coordinated attacks which are used as input parameters to model interactions between attacker and defender. Another paper looks into specific types of attacks that exploit cyber vulnerabilities of specific meters and spread into the physical components of the system [121]. This paper also proposes game theory to analyze such attacks. Similarly, a game-theoretic perspective of data injection attacks with multiple adversaries is also studied [122].

There is also focus on the disruption of specific common attacks in the smart grid. Game theory based defense strategies against DoS attacks are proposed which use Nash Equilibrium to maintain dynamic stability in an attacked system [123, 124]. Minimizing the effects of jamming attacks is also studied through a modified version of contract network protocol (CNP) as a negotiation protocol among agents [125]. Results of this paper indicate that applying the proposed protocol can reduce the jammers illegal profit and decrease their motive. The problem with most of the proposed game theory techniques, however, is their tendency to view network interdictions as one-time events. Further research in this topic is likely to take shape as more comprehensive modeling of network interdictions occurs. There are few papers in literature that take this into consideration. One paper, however, uses zero-sum Markov games and a more comprehensive model of attacker behavior [126]. This paper also demonstrates a defender can use deception as a defense mechanism. Next, we discuss the deception techniques proposed in literature, which when combined with the aforementioned disruption techniques, can act as a comprehensive strategy for mitigating attacks.

3.2.5 Deception of Attackers

While disruption of attacks involves minimizing the damage of cyber attacks, deception focuses on altering the direction of the attack to mitigate its impact. This is done by deceiving the attacker into targeting a trap. This deception technology is an emerging field in cyber security due to its potential to detect and defend against zero-day and advanced attacks. In the security of the smart grid, however, deception technology is seldom used in literature.

A strategic honeypot game model was proposed for DoS attacks in the smart grid [127]. This paper introduces honeypots into the metering infrastructure network as a decoy system to detect and gather information. Interactions between attackers and defenders are analyzed and the existence of several Bayesian-Nash equilibrium is proved. However, this method was designed and tested for one specific type of attack. A more general honeypot system is proposed to emulate an entire smart grid field communication infrastructure in [128]. This paper claims that their honeynet system can emulate high-fidelity and realistic power grid behavior to deceive the attackers. However, evaluation of its realism and scalability are only preliminary and testing was done on a single simulated system. Another paper identifies the various types of honeypots and built a test system to emulate a device on a utility network [129]. However, similarly to the aforementioned papers, analysis regarding realism and scalability are insufficient. This is identified as a research gap in the deception strategies for smart grid security. Future research is expected to involve more comprehensive system modeling and the proposal of more versatile honeynet systems.

3.2.6 Degradation or Destruction of Attacks

Degrading or destroying the attack is the final part of the defense strategy in the smart grid and it involves minimizing or destroying the effects of the attack. An example of such mitigation techniques include defining security metrics that quantify the importance of individual substations [130]. Another proposed method uses a distinctive modeling technique with the capability to modify network topology [131]. Such a technique can be used to degrade the attack through optimizing the operation of the power system to minimize its effects. This is done through a mixed-integer nonlinear bi-level program; in the upper-level a terrorist agent maximizes the damage caused in the power system, and in the lower level the system operator minimizes the damage through optimal operation of the power system. Furthermore, the paper proposes a Benders decomposition approach to transform

the problem into a standard one-level optimization problem. Another paper, however, tackles the same problem through a genetic algorithm [132]. Alternatively, another paper proposes a different tri-level model for power network defense with the same goal of minimizing economic cost that the attacks may cause [133].

Degradation techniques are often coupled with disruption techniques in game theory approaches, as mentioned in Subsection 3.2.4. As such, defense solutions that only focus on degradation of attacks are limited. Furthermore, due to the legal implications, there are no solutions proposed that focus on destroying the attack through hostile actions towards the adversary. Therefore, most solutions in literature focus on denying, detecting, and minimizing the effect of attacks.

3.3 Summary

This chapter is a survey of literature aimed at identifying the types of existing cyber threats and defense methods to tackle them. The first part of the chapter identifies the types of cyber attacks studied in literature and identifies the most common attacks studied in regards to smart grid security. The second part of this chapter delves into the security and defense methods proposed in literature, which are summarized in figure 3.2.

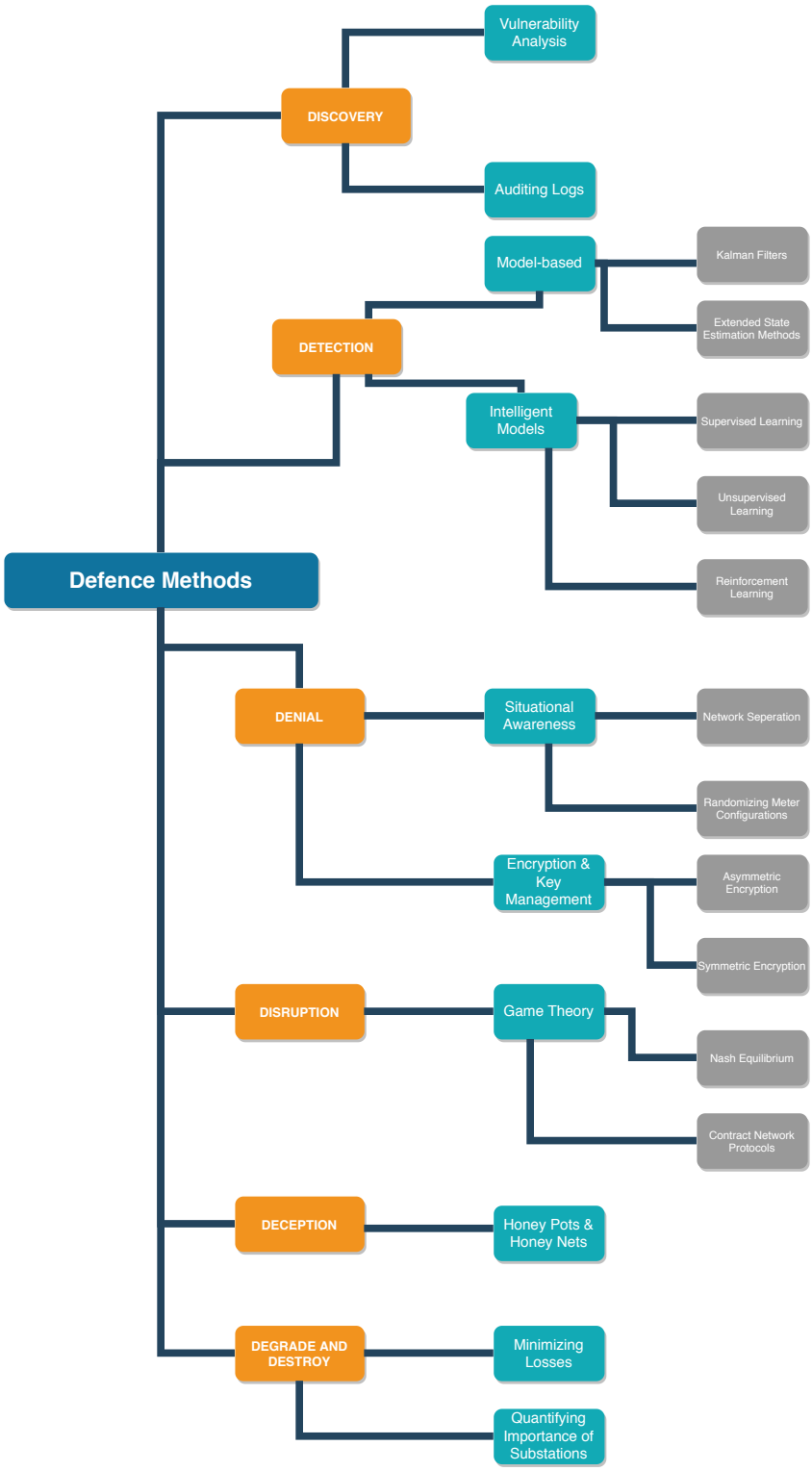

Figure 3.2: Common smart grid defence methods discussed in literature

Chapter 4

Proposed Models

This chapter discusses the proposed models used to complete the contributions of this thesis. The chapter is divided into two sections; the first demonstrates how the first contribution, heuristic feature selection, is acheived. The second discusses the second contribution, which is a deep learning algorithm robust to varying attack sparsity and data imbalance.

4.1 Model 1: Heuristic Feature Selection

Power systems are highly complex and large scale physical systems with huge number of features and measurements. Therefore, feature selection is an essential task that should be performed to optimize the computational efficiency [134]. Principal Component Analysis (PCA) has been used in previous literature for dimensionality reduction [91]. However, large-scale power systems behave somewhat non-linearly; and as such, heuristic approaches to feature selection are considered. In this paper, GA, Cuckoo Search (CS), and Particle Swarm Optimization (PSO) are used to increase the computational efficiency of the supervised learning algorithms. Each of the algorithms are aimed to obtain the most optimal subset of features that results in the best accuracy. Each solution consists of a binary vector with each index being 1 if the feature is used in this subset and 0 if it is not.

This model meets the first contribution of increased computational efficiency by utilizing heuristic algorithms to select ideal feature subsets. The three heuristic algorithms used in this experiment are explained in the following subsections.

4.1.1 Binary Cuckoo Search

BCS is a binary implementation of CS, an optimization algorithm based on the parasite behavior of some species of Cuckoo. The CS algorithm is proposed by [135] and summarized by the following three rules:

1. Each Cuckoo lays one egg at a randomly chosen nest.
2. The best nests with high quality eggs carry over to the next generation.
3. The number of available nests is fixed. And if another cuckoo egg is discovered by the host bird, the host can remove the egg or build a new nest.

Mathematically, the nests, or solutions, are updated using random walk via Lévy flights:

$$x_i^j(t) = x_i^j(t-1) + \alpha \oplus Levy(\lambda) \tag{4.1}$$

and

$$Levy \sim u = s^{-\lambda}, (1 < \lambda \leq 3) \tag{4.2}$$

where x_i^j is the j^{th} egg (feature) at nest (solution) i, s is the step size, $\alpha > 0$ is the step size scaling factor, and $\oplus$ is the entry-wise product. The Lvy flights employ a random step length which is drawn from a Lévy distribution which creates longer step length in the long run allowing more efficient search space exploration [135]. The solutions are restricted to binary values by the following equations:

$$S\left(x_i^j(t)\right) = \frac{1}{1 + e^{-x_i^j(t)}} \tag{4.3}$$

$$x_i^j(t+1) = \begin{cases} 1 & \text{if } S\left(x_i^j(t)\right) > \sigma \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

in which $\sigma \sim U(0,1)$ and $x_i^j(t)$ denotes the new egg value at time t [136]. The pseudo code for cuckoo search is shown in algorithm 1.

Algorithm 1: Cuckoo Search Optimization

Objective Function: $f(\mathbf{x}), \quad \mathbf{x} = (x_1, x_2, \ldots, x_d)$

Generate initial population of n host nests;

while $t < maxIterations$ **do**

  Get a random cuckoo and replace its solution by performing Lvy flights;
  Evaluate its fitness F_i (classification accuracy) choose a random nest, j;
  if $F_i > F_j$ **then**
    replace j with new solution;
  end
  Fraction p_a of the worst nests are abandoned and new ones are built in their place;
  Rank solutions from best to worst;
  Save the best solution for next iteration;

end

4.1.2 Genetic Algorithm

GA is an optimization technique that yields the best solution based on the evolution mechanism of living beings [137]. Following the principle of natural selection, GA chooses the best solutions based on their fitness. In each iteration, GA eliminates the solutions with the lowest fitness and retains the solutions with the highest fitness. The psuedo code for GA is shown in algorithm 2. Similarly to 4.1.1, the solution consists of a binary vector indicating the variables used as features, and the fitness of each solution is the classification accuracy of FDI attacks based on that subset of features.

Algorithm 2: Genetic Algorithm Optimization

Generate n random solutions (population);
evaluate and rank the solutions;

while $t < maxIterations$ **do**

  Select best-fit solutions for reproduction;
  create new solutions through crossover and mutation operations;
  evaluate the fitness of new solutions;
  replace least-fit solutions with the new ones;

end

4.1.3 Binary Particle Swarm Optimization

PSO is an algorithm used for solving a variety of problems. The algorithm is motivated by social behaviors in nature. The main characteristic of this algorithm is that optimization is performed through social interaction in the population where thinking is not only personal, but also social [138]. A binary implementation of Particle Swarm Optimization (BPSO) is also used as a heuristic method for feature selection.

The first step of implementing BPSO is initialization of population consisting of user defined particles; each particle represents a feasible solution. Through iterations, particles update themselves by tracking two criteria. The first criterion is the best solution of each particle. Personal best of the i^{th} particle is $pBest_{\mathrm{i}} = (pBest_i^1, pBest_i^2, \ldots, pBest_i^n)$. And the second criterion is global best solutions, $gBest = (gBest^1, gBest^2, \ldots, gBest^n)$ respectively. The pseudo code for PSO can be found in algorithm 3.

Algorithm 3: Particle Swarm Optimization

Cost Function: $f : \mathbb{R}^n \to \mathbb{R}$;
for *each particle* $i = 1, 2, ..., N$ **do**
    Initialize the particle's position with a uniformly distributed random vector;
    Initialize the particle's best known position to its initial position;
    if $f(\mathbf{p}_i) < f(\mathbf{g})$ **then**
        update the swarm's best known position $g = p_i$;
    end
    Initialize the particle's velocity;
end
while $t < maxIterations$ **do**
    for *each particle* $i = 1, 2, ..., N$ **do**
        for *each dimension* $d = 1, 2, ..., D$ **do**
            Update particle i's velocity in dimension d by a random amount;
        end
        Update the particle's position (x_i) using new velocities;
        if $f(\mathbf{x}_i) < f(\mathbf{p}_i)$ **then**
            update the swarm's best known position $p_i = x_i$;
            if $f(\mathbf{p}_i) < f(\mathbf{g})$ **then**
                update the swarm's best known position $g = p_i$;
            end
        end
    end
end

4.2 Model 2: Generalized Deep Neural Network (GDNN)

The proposed GDNN model consists of an input layer, four hidden layers of 128, 64, 32, and 16 nodes respectively, and an output layer. Each of the hidden layers employ rectified linear unit activation, commonly referred to as ReLu activation, and the output layer uses sigmoid activation for binary classification. The number of nodes and layers were selected using cross-validation of different networks and analyzing their loss history, validation accuracy, and training time. This model meets the second contribution of accurate detection with varying attack sparsity and data imbalance. This is achieved by utilizing the aforementioned regularization methods which aim to learn more generalized patterns of attacks in less data and epochs; thus outperforming other algorithms in detecting low sparsity attacks in high degrees of data imbalance.

Binary cross entropy (BCE) is used as the cost function which can be represented by:

$$J = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\left(p\left(y_i\right)\right) + \left(1 - y_i\right) \cdot \log\left(1 - p\left(y_i\right)\right) \tag{4.5}$$

where y is the label (1 for attack and 0 for normal sample), $p(y)$ is the predicted probability of the sample containing an attack, and N is the number of samples. Furthermore, L2 regularization is utilized in all four hidden layers. L2 regularization, also known as Ridge Regression, adds the squared magnitude of the weights as a penalty to the cost function as follows:

$$\begin{aligned} J = -\frac{1}{N}\sum_{i=1}^{N}[y_i \cdot \log\left(p\left(y_i\right)\right) + \left(1 - y_i\right) \\ \cdot \log\left(1 - p\left(y_i\right)\right) + \lambda \sum_{j=1}^{p} \omega_j^2] \end{aligned} \tag{4.6}$$

where λ is the regression coefficient. In this model, we use $\lambda = 0.001$. Utilizing L2 regularization helps avoid output dependencies on a specific set of parameters. As such, L2 regularization helps to avoid over-fitting and increase the generalization of a neural network model.

Additionally, dropout is used in between the hidden layers to further enhance the generalization of our model. Dropout is a method in which a certain percentage of the neuron interconnections are temporarily disabled during training. In each epoch of training, a differ-

ent set of connections are disabled. The purpose of this method is to reduce the dependency of the output on a specific set of parameters, much like the L2 regularization.

Finally, a hybrid learning rate optimizer, Adadelta, is used to train the proposed neural network. In the Adadelta optimizer, the running average of the squared gradients, $E\left[g^2\right]_t$, is computed as follows [139]:

$$E\left[g^2\right]_t = \rho E\left[g^2\right]_{t-1} + (1-\rho)g_t^2 \tag{4.7}$$

where ρ is a decay constant. The square root of the moving average is used in the parameter updates of the neural network. Therefore, the Root Mean Square (RMS) of previous gradients up to time t is computed by:

$$\text{RMS}[g]_t = \sqrt{E\left[g^2\right]_t + \beta} \tag{4.8}$$

The constant β is added to better condition the denominator. Based on the above, the resultant parameter update is represented by:

$$\Delta x_t = -\frac{\eta}{\text{RMS}[g]_t} g_t \tag{4.9}$$

The GDNN algorithm is trained on data from the smart grid system that can either be collected or simulated based on the system topology. This algorithm collects data from measurements along the system in a periodic manner, and sends a response to the control center classifying each sample of data as either normal or malicious. Upon detection of malicious data, an alarm system is triggered notifying the control center of the presence of malicious data. A diagram portraying the deployment of this model can is shown in figure 4.1.

Algorithm 4: Proposed GDNN Method

Data: power flow and load measurements from all smart meters of the system

Training Phase

for *number of training samples N* **do**
- collect measurements Z_i for $i = 1, ..., N$;
- collect associated labels y_i for $i = 1, ..., N$;

end

$Z = \frac{Z - Z_{\min}}{Z - Z_{\max}}$;

initialize learning rate $l = 1$;

initialize weight vector ω randomly;

while $validationAccuracy[k] < validationAccuracy[k + 5]$ **do**
- instructions;
- **for** *all* (Z_i, y_i) **do**
  - randomly set 30% of the weight vector ω values to 0;
  - compute y_i using feedforward;
  - compute cost function as per eq 4.6;
  - compute gradient using backpropagation;
  - accumulate gradient as per eq 4.7;
  - compute update: $\Delta\omega_t = -\frac{\mathrm{RMS}[\Delta\omega]_{t-1}^{\infty}}{\mathrm{RMS}[g]_t} gt$;
  - accumulate update: Accumulate Updates: $E\left[\Delta\omega^2\right]_t = \rho E\left[\Delta\omega^2\right]_{t-1} + (1-\rho)\Delta\omega_t^2$;
  - apply update: $\omega_{t+1} = \omega_t + \Delta\omega_t$;
- **end**

end

Attack Detection

while *system active* **do**
- collect current measurement vector Z_t;
- calculate output of GDNN y_t using feed-forward;
- **if** $y_t > 0.5$ **then**
  - activate attack alarm;
- **else**
  - continue check;
- **end**

end

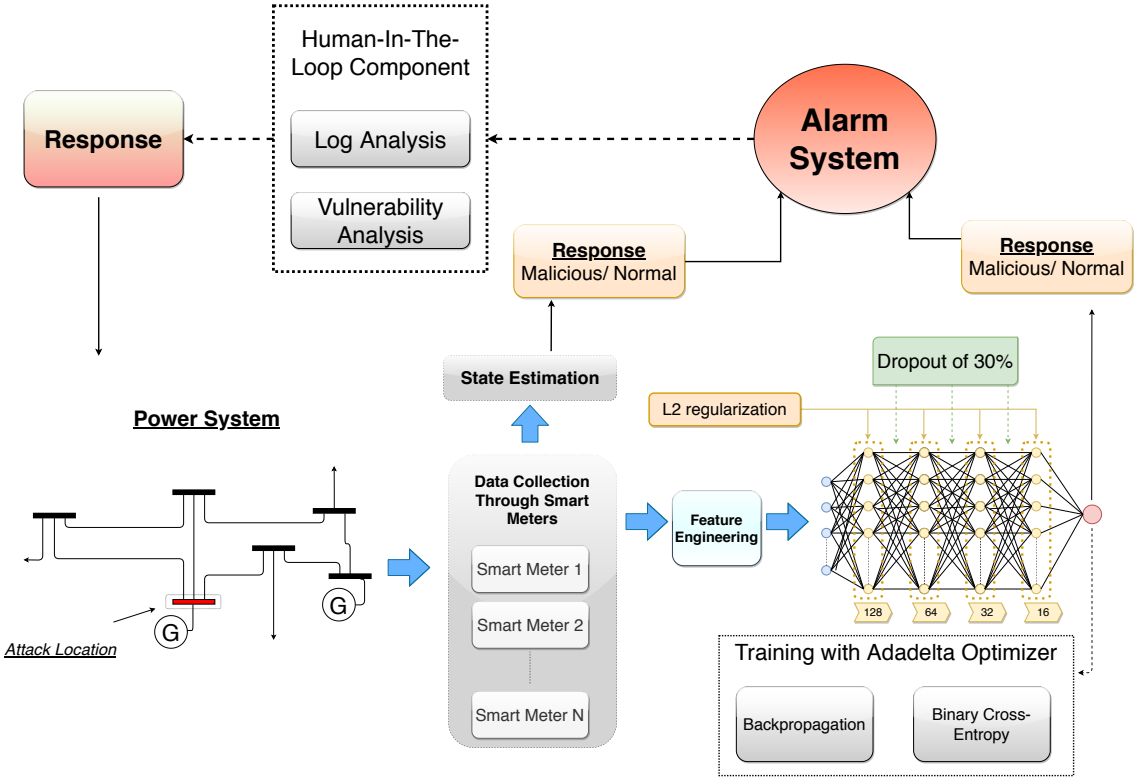

Figure 4.1: The architecture of the proposed model framework

4.3 Summary

In this chapter, the proposed frameworks used to complete each contribution are outlined and explained. The chapter discusses the three heuristic algorithms tested for feature selection to complete the first contribution. It then discusses the deep learning algorithm and the regularization techniques implemented to meet the second contribution of the thesis.

Chapter 5

Methodology

This chapter discusses the methods in which the research contributions of this thesis are met. The contributions of this study comprise of heuristic methods for increased computational efficiency, generalized deep-learning based method for attack detection that is robust to attacks of varying sparsity, as well as the minor contribution in the form of literature survey. The experimental process carried out for this research can be categorized under five steps:

1. **Literature Analysis:** The research process begins with exploring academic literature for existing issues in smart grid security as well as state of the art defense methods. This step is carried out in the form of a bibliometric analysis of journal articles in the past decade as explained in chapter 3. This survey of literature categorized the types of threats facing the smart grid as well as the defense mechanisms used in each layer of security.

2. **Proposing an Initial Framework:** After analyzing the literature for state of the art methods, an initial framework for attack detection is proposed. This framework utilizes heuristic feature selection to reduce the dimensionality of the data thereby increasing the computational efficiency of intelligent classifiers.

3. **Data Collection:** To test the proposed framework, data of smart cyber-physical grids must be collected. Since real smart grid data is scarce, particularly malicious data, a data generation framework was designed for this step. This physics-based simulation framework simulates a smart grid using standard IEEE power system structures. These systems are simulated under varying demand conditions for realistic distribution of data. Furthermore, stealthy data injection attacks are simulated using the mathematical concepts discussed in 2.3.1.

4. **Develop and Analyze Initial Framework:** In this step, the initial proposed framework, heuristic feature selection, is tested and evaluated using the data collected. Analyzing these results motivated a novel technique for attack detection.

5. **Refine Framework Based on Analysis Results:** Analyzing the results of heuristic feature selection concludes that while it is an effective mean of increasing computational efficiency of classifiers, it is not robust to varying attack sparsity and data imbalance. As such, a deep learning method for attack detection is proposed in this step.

5.1 Research Dataset

Considering the scarcity of attack data collected from real smart grids, this experiment utilizes a simulation framework to generate data. The data used in this experiment is generated using MATPOWER library [140]. This library was chosen for its convenience as well as its wide use in literature. The power systems used for testing are the IEEE 14-bus, IEEE 30-bus, IEEE 57-bus, and IEEE 118-bus. The measurement data consists of power flow of branches and buses as well as generator outputs which are mapped into the state variables, the voltage bus angles, using the Jacobian matrix. Based on the aforementioned process in section 2.3.1, samples of system data is generated under normal and attack behavior. Initially, data was generated randomly using random attack scenarios. However, for the second experiment, attack data for varying sparsity conditions were generated. In this context, sparsity refers to the percentage of measurements compromised in an attack scenario. To confirm that the attacks generated are indeed stealthy FDI attacks, we measure the L2-norm of 100 of each normal and attack samples as plotted in figure 5.1.

The data generated is divided into two halves, half of the samples are normal data, and the other half are malicious. The amount of data generated for each experiment was different based on the computational burden of the experiment. Since the first experiment was more computationally expensive, fewer samples were used. The number of samples used for each experiment are as follows:

- **Experiment 1 - Testing the initial framework**
  - *Training and Validation Data* = $10,000$
  - *Testing Data* = $1,000$
- **Experiment 2 - Testing the final framework**

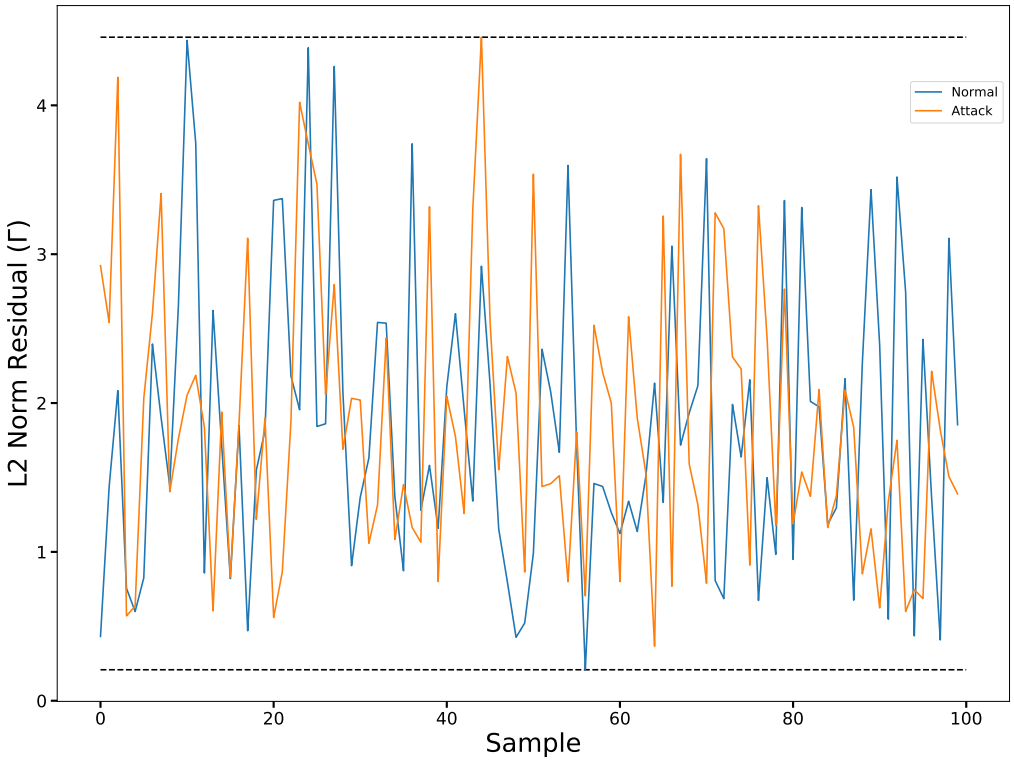

Figure 5.1: The L2-norm of attack and normal samples taken from the IEE 30-bus system

- *Training Data* = 120,000
- *Validation Data* = 30,000
- *Testing Data* = 20,000

5.2 Research Method

To achieve each of the two major contributions of this thesis, two experiments are carried out:

- **Experiment 1: Heuristic Feature Selection for Increased Computational Efficiency**
  In this experiment, various heuristic feature selection approaches are tested with dif-

ferent classification methods. The goal is to maintain classification accuracy while reducing the number of features; which in turn increases the computational efficiency of the attack detection classifier.

- **Experiment 2: Generalized Deep Neural Network for Attack Detection Under Varying Attack Sparsity Conditions**
  In this experiment, a deep neural network with additional generalization techniques is proposed. This algorithm is capable of identifying attacks regardless of sparsity, which is the percentage of measurements that are compromised. Furthermore, due to hybrid learning rate and early stop methods, this algorithm is also significantly faster to train when compared to other neural-network-based algorithms.

Testing attack detection methods can be a complex process. This testing process is approached differently based on the specific context of the application. For the purpose of attack detection in power systems, testing can be challenging due to the lack of available data. While some data of real or simulated systems can be found in public sources, the quality of such data does not allow for comprehensive testing methods. This is because there is a lack of attack data recorded from real power systems. As such, researchers tend to use physics-based simulation frameworks to simulate data from power systems and simulate the different types of attacks that have not been recorded in real systems. For this reason, data generation is an essential part for analyzing security methods for critical infrastructure.

5.3 Research Evaluation

Different benchmark and metrics are used to evaluate the performance of the attack detection algorithms:

- **True Positive (TP):** The number of samples *correctly* classified as positive (attack).
- **True Negative (TP):** The number of samples *correctly* classified as negative (no attack).
- **False Positive (TP):** The number of samples *incorrectly* classified as positive (attack).
- **False Negative (TP):** The number of samples *incorrectly* classified as negative (no attack).

Using the above core metrics,the performance of machine learning systems can be evaluated based on their test accuracy, F1-score, and Matthews Correlation Coefficient (MCC). The test accuracy refers to the percentage of correctly predicted test samples, the F1-score is a harmonic mean of the precision and recall, and MCC is the correlation between the true and predicted binary classifications [141]. MCC is returned as a value between -1 and $+1$ in which $+1$ refers to perfect prediction, -1 refers to complete disagreement between predicted and true predictions, and 0 is considered no better than random prediction [142]. The accuracy, F1 score, and MCC can be computed as follows:

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{5.1}$$

$$F_1 = 2 \cdot \frac{\text{precision recall}}{\text{precision } + \text{recall}} \tag{5.2}$$

$$\text{MCC} = \frac{T_p \times T_n - F_p \times F_n}{\sqrt{(T_p + F_p)(T_p + F_n)(T_n + F_p)(T_n + F_n)}} \tag{5.3}$$

where T_p and T_n are true positives and true negatives respectively, and F_p and F_n are false positives and false negatives respectively. Precision and recall are measures of relevance of the output of machine learning algorithms. Precision is the fraction of correctly classified positive samples to all samples classified positive. Alternatively, recall is the fraction of correctly classified positive samples to all correctly classified samples. The precision and recall can be computed as follows:

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{5.4}$$

$$\text{Recall} = \frac{T_p}{T_p + T_n} \tag{5.5}$$

The aforementioned evaluation metrics, accuracy, $F1$ score, and MCC are used to validate the two major contributions of this experiment. Contribution 1 is met when accuracy and $F1$ score are maintained for lower number of features. In other words, if the same accuracy and $F1$ score can be achieved with lower number of features, the method is thereby considered more computationally efficient. Similarly, contribution 2 is validated when all three metrics are maintained at varying attack sparsity and degrees of data imbalance.

5.4 Summary

In this chapter, the methods of this research are outlined. First, the general steps of research progress are presented. Then a framework for data generation and collection is discussed. This is followed by a thorough explanation of how each of the contributions of this research are achieved. Furthermore, the methods in which the experiments are carried out are discussed. Finally, the evaluation metrics used in this research are presented and explained.

Chapter 6

Results and Discussion

There are two interconnected experiments that comprise this research. The first experiment compares various heuristic feature selection techniques with different machine learning classifiers. The goal of this experiment is to automate the feature selection process through heuristic algorithms that choose ideal subsets of features achieving maximal computational efficiency. Next, the more complex problem of varying attack sparsity is considered. Under such methods, it was proven that all measurements in the system have equal co-variance with the output; meaning that feature selection techniques are less plausible. As such, a deep learning model is proposed. This model incorporates generalization techniques for increased scalability and improved performance on imbalanced data. This model is proposed in the second experiment, which compares the model to various machine learning models under different sparsity and data imbalance conditions.

6.1 Experiment 1: Heuristic Feature Selection for Increased Computational Efficiency

This experiment tests three machine learning classifiers with three different heuristic algorithms for feature selection. This is tested on the IEEE 14-bus, IEEE 57-bus, and IEEE 118-bus systems. Data for these standard power systems is generated as described in section 5.1. The process of this experiment can be divided into two steps: the first is finding optimal parameters to use for each machine learning classifier and the second is to test these classifiers with the chosen parameters with each feature selection (FS) technique. The three machine learning classifiers used are KNN, SVM, and ANN and the three heuristic optimiza-

tion algorithms used are Genetic Algorithm (GA), Binary Cuckoo Search (BCS) and Binary Particle Swarm Optimization (BPSO).

6.1.1 Choosing Ideal Parameters for Machine Learning Classifiers

Parameter optimization of each of the supervised learning algorithm is performed through cross-validation of varying parameters with optimal accuracy. This cross-validation test was performed on the data from the smallest system, IEEE 14-bus, due to its high computational cost. SVM is cross-validated for varying kernel coefficient and penalty parameter,γ and C respectively; this is demonstrated in figure 6.1. Based on this figure, it can be seen that optimal performance is achieved by minimizing the kernel coefficient and maximizing the penalty parameter.

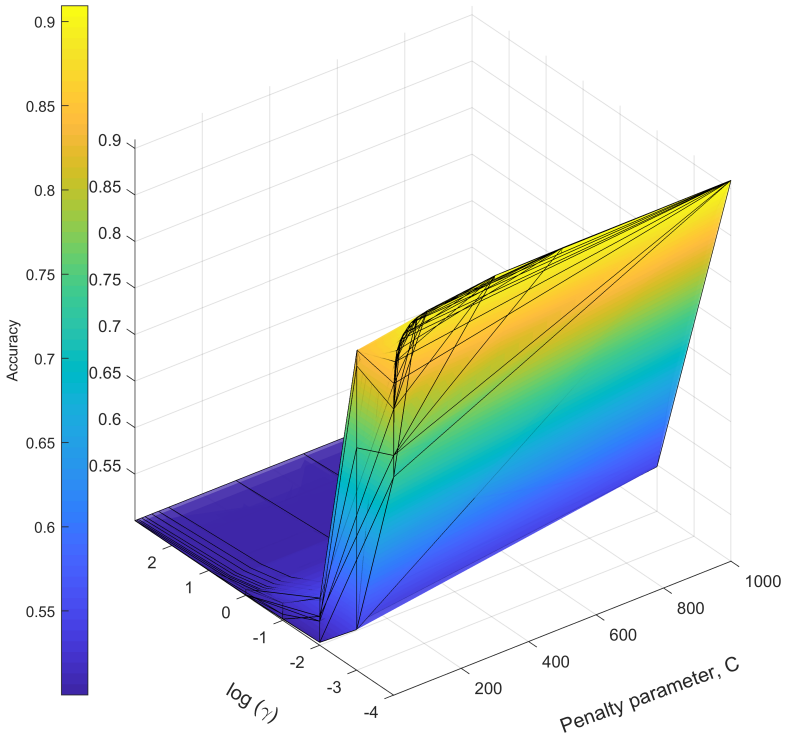

Figure 6.1: The accuracy of SVM on the IEEE 14-bus system for varying penalty parameter and kernel coefficient

KNN is cross-validated in a similar manner for varying number of neighbors, K. As demonstrated in figure 6.2, the accuracy at varying K values shows that the accuracy is

maximized in a smaller number of neighbors. As such, a value of $K = 12$ was chosen for the KNN algorithm.

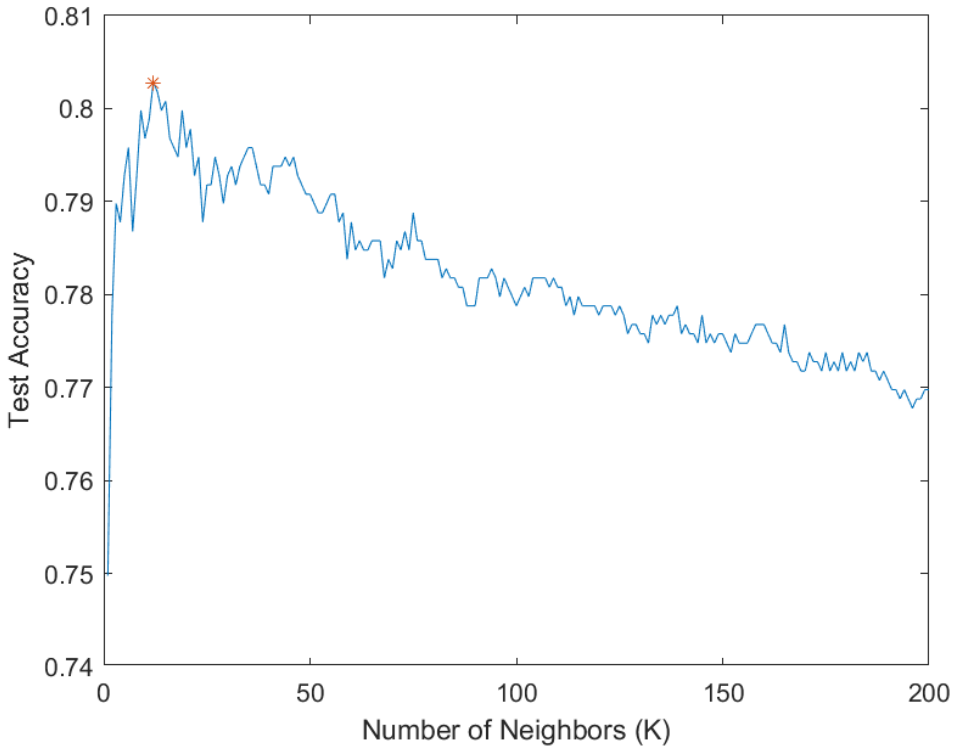

Figure 6.2: The accuracy of KNN on the IEEE 14-bus system for varying number of neighbors

Choosing optimal ANN parameters is a more complex task. There are many parameters to vary in ANN, such as the number of layers, the number of nodes in each layer, the cost function, activation functions between layers, and more. In this research, a common method for selecting an ANN architecture is used. The ANN architecture consists of 1 hidden layers of M nodes where

$$M = \left\lceil \frac{N + L}{2} \right\rceil \tag{6.1}$$

and N and L represent the number of classes and number of features respectively. Since this is a binary classification, $N = 2$. The activation functions used are rectified linear unit (ReLu) and Sigmoid activation for the output. This structure of ANN is cross-validated for varying learning rate, α. The results of this cross-validation, exhibited in figure 6.3, show that a learning rate of $10^{-}6$ is ideal for this application.

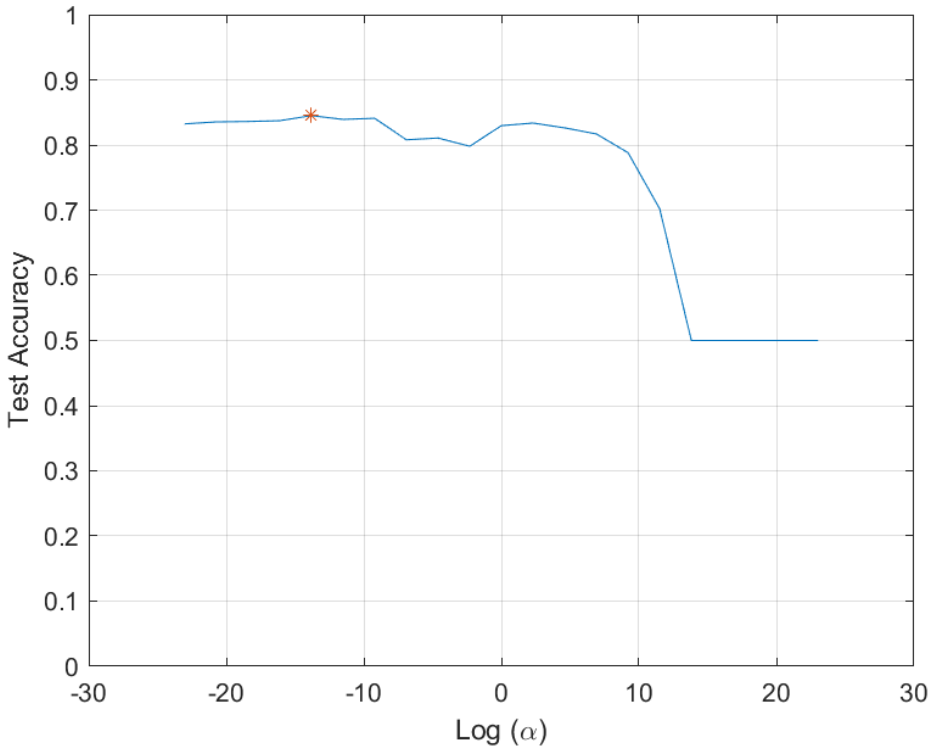

Figure 6.3: The accuracy of ANN on the IEEE 14-bus system for varying learning rates

The final parameters chosen for each of the three classifiers and the corresponding test accuracy are outlined in table 6.1. These parameters are constant throughout this experiment.

Table 6.1: Optimal parameters of the supervised learning algorithms and their corresponding accuracy on the IEEE 14-bus system with no feature selection

Algorithm	Parameters	Accuracy
SVM	$C = 1000$, $\gamma = 0.0001$	90.93%
KNN	$K = 12$	80.82%
ANN	$\alpha = 10^{-6}$	84.50%

6.1.2 Testing Heuristic Algorithms for Feature Selection

The three FS methods, BCS, BPSO, and GA, are implemented with the parameters stated in table 6.2 which are chosen based on similar applications in literature [136][137]. The resultant subset of features selected by each algorithm are tested with the three classification algorithms, SVM, KNN, and ANN, and their classification accuracy on each of the three IEEE bus systems are recorded in tables 6.3, 6.4, and 6.5.

Table 6.2: Parameters of the heuristic FS algorithms

Algorithm	**Parameters**
BCS	$\alpha = 0.1$, $P(a) = 0.25$, population = 30, iterations = 10
BPSO	$c_1 = c_2 = 2$, $w = 0.7$, population = 30, iterations = 10
GA	mutation rate = 0.018, population = 50, iterations = 30

Table 6.3: Classification accuracy of each supervised learning algorithm with each heuristic feature selection technique on the IEEE 14-bus system

FS Method	**Num of Features**	**Classification Accuracy**		
		SVM	***KNN***	***ANN***
NO FS	34	90.79%	80.28%	81.78%
BCS	11	90.69%	81.38%	77.08%
BPSO	8	90.19%	81.68%	79.18%
GA	8	90.49%	82.28%	79.28%

Table 6.4: Classification accuracy of each supervised learning algorithm with each heuristic feature selection technique on the IEEE 57-bus system

FS Method	**Num of Features**	**Classification Accuracy**		
		SVM	***KNN***	***ANN***
NO FS	137	88.29%	83.08%	50.05%
BCS	94	88.59%	84.48%	50.15%
BPSO	130	87.39%	83.58%	48.25%
GA	56	87.39%	85.59%	50.95%

Table 6.5: Classification accuracy of each supervised learning algorithm with each heuristic feature selection technique on the IEEE 118-bus system

FS Method	**Num of Features**	**Classification Accuracy**		
		SVM	***KNN***	***ANN***
NO FS	304	84.88%	74.57%	53.05%
BCS	199	83.58%	75.48%	51.25%
BPSO	160	83.28%	76.68%	51.95%
GA	122	90.59%	78.18%	50.05%

Results show that SVM and KNN are successful at detecting FDI attacks in all three IEEE bus systems. SVM is the most versatile scoring the highest classification accuracy among all the FS methods and in all three test systems. Furthermore, all three heuristic FS methods proved successful at reducing the number of features. GA produced the most successful results among the three FS methods by achieving the highest classification accuracy with minimal number of features. ANNs with the proposed architecture were unsuccessful at detecting FDI attacks regardless of the FS method.

Overall, heuristic FS methods were successful at maintaining, and sometimes increasing, the classification accuracy with significantly lower number of features. SVM and KNN algorithms proved more accurate and versatile among the three systems when compared to the ANN implemented in this paper. However, ANNs with more complex architectures are expected to have better performance on larger systems at a higher computational cost.

FS methods were all successful at increasing accuracy or reducing the number of features, and in some cases both. Classification results conclude that GA is the most efficient heuristic FS method for power systems in terms of accuracy and number of features. SVM with GA proved to be the most accurate and versatile among the three systems.

6.2 Experiment 2: GDNN for Attack Detection Under Varying Attack Sparsity Conditions

In this experiment, we test the GDNN model, as well as other classification methods, on three IEEE standard power systems. The three systems are the IEEE 14-bus, the IEEE 30-bus, and the IEEE 57-bus system. The purpose of testing on various power systems is to ensure the robustness and scalability of our method.

The proposed GDNN model is compared to other classifiers which are all tested on the three aforementioned power systems. These classifiers are Naive Bayesian, K-nearest neighbor (KNN), a Decision Tree classifier, and a normal feed-forward Artificial Neural Network (ANN) that has similar architecture to our proposed GDNN model but lacks the proposed regularization methods.

Furthermore, the models are also tested on imbalanced data. This test was performed by removing a percentage of the attack data from the training and validation set. The accuracy of the models is recorded for varying percentage of imbalance.

6.2.1 Complexity Analysis and Feature Selection

To assess the complexity of the problem, we analyze the correlation of each feature to the output. A problem is considered easy to solve by most models if there is a high correlation of a small set of features with the output. Calculating the correlation of each feature with the output can be done by computing the information gain ratio, which is the ratio of the information gain of a feature to its intrinsic value. The information gain and the intrinsic value of each feature X in data-set D are computed as follows:

$$\text{Intrinsic Value } (D|X) = -\sum_{i=1}^{n} \frac{|D_i|}{|D|} \log_2 \frac{|D_i|}{|D|} \ , \tag{6.2}$$

$$\text{Gain}(X) = \text{Entropy}(D) - \text{Entropy}(D|X) \ , \tag{6.3}$$

where

$$\text{Entropy}(D|X) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} \text{Entropy}\,(D_i) \tag{6.4}$$

and

$$\text{Entropy}(D) = -\sum_{i=1}^{n} p_i \log_2 p_i \ . \tag{6.5}$$

The information gain ratio, is then computed as the ratio of information gain to the intrinsic value of each feature:

$$\text{Infromation Gain ratio } (A) = \frac{\text{Information Gain}(X)}{\text{Intrinsic Value}(D|X)} \tag{6.6}$$

Feature analysis is performed on the IEEE 30-bus system using data with attacks of varying sparsity. The information gain ratio of each feature of the IEEE 30-bus system are shown

in Figure 6.4. The small range and variance of IG ratio values demonstrated in this figure shows even relative importance among all features particularly in cases of high range of attack sparsity. Because attacks can be present in any subset of measurements, there is no strong correlation between one measurement and the output. As such, eliminating any feature from the data can reduce the accuracy of machine learning algorithms. For this reason, all measurements of the power system are used in training the algorithms to detect the attacks.

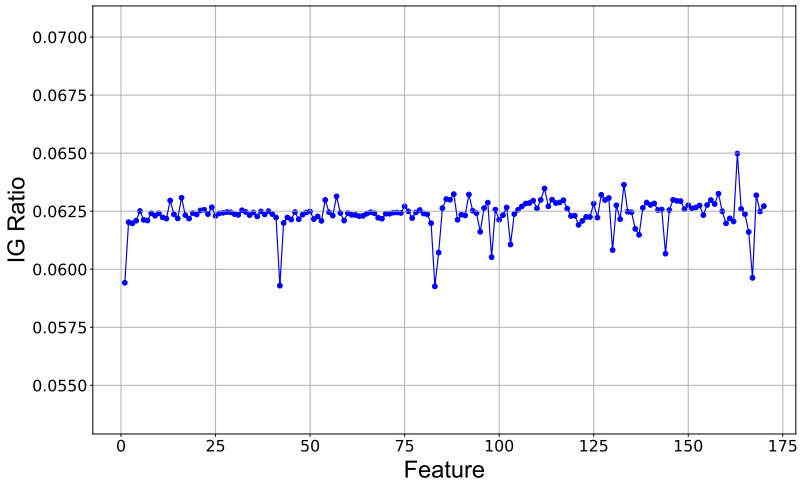

Figure 6.4: The Information Gain Ratio of Each Feature

6.2.2 Training Analysis

To ensure the algorithm is learning, the GDNN model is trained on the IEEE 30-bus system and the loss and accuracy of the model are recorded for training analysis. Figure 6.5 shows the training and validation loss of our GDNN model and a neural network with the same architecture but without the proposed regularization methods and learning rate. The loss is recorded for 100 epochs of training. Similarly, the training and validation accuracy per epoch is shown in Figure 6.6.

Based on Figures 6.5 and 6.6, we observe that our proposed model learns significantly faster than a similar non-regularized model of the same architecture. The faster learning is largely attributed to the hybrid learning rate optimizer, Adadelta. This changing learning rate, however, also results in larger fluctuations of loss and accuracy. To account for this, we

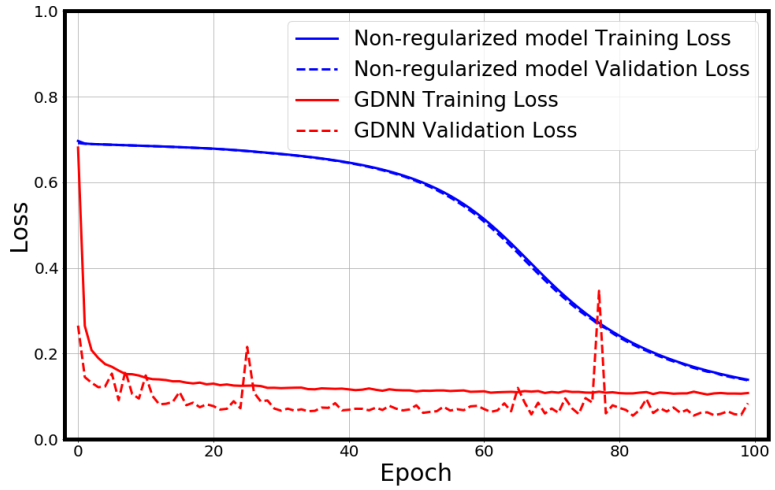

Figure 6.5: The training and validation loss of GDNN and non-regularized ANN

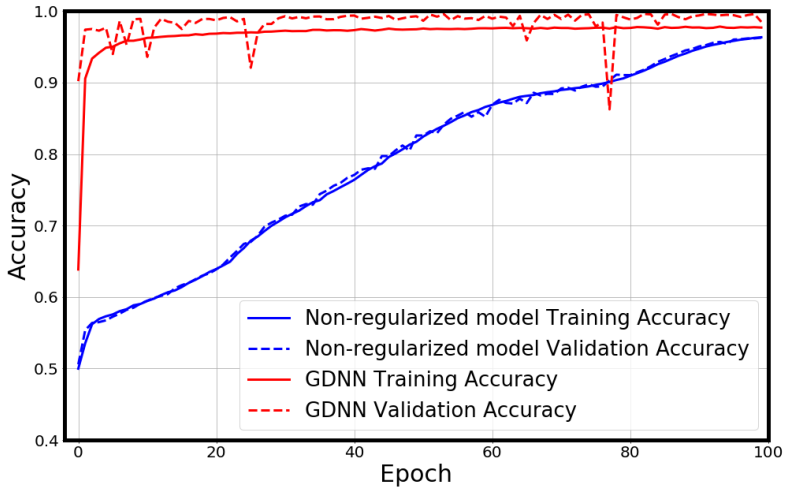

Figure 6.6: The training and validation accuracy of GDNN and non-regularized ANN

employ early stopping in which the training stops if the validation accuracy does not increase within 5 training epochs. This ensures that the model is not over-trained and maximizes its generalization on unseen data.

6.2.3 Sparsity Analysis

We compare the testing accuracy on each sparsity level of our model to the non-regularized similar model as well as other traditional machine learning classifiers. The regularization methods, dropout and l2-regularization, increase the generalization of neural networks. This is exhibited by superior accuracy and F1-score achieved by the GDNN, as shown in Figures 6.7 and 6.8. These figures demonstrate the superior performance achieved by the GDNN model under varying attack sparsity when compared to other models on the IEEE 30-bus system. This superior performance can be exhibited by the high accuracy and F1 score of GDNN at all sparsity values. In fact, the GDNN achieved the highest accuracy and F1 score at all values of sparsity except 0.1 where it was outperformed by the Naive Bayesian algorithm. This performance of the Bayesian algorithm, however, is not consistent to all sparsity values.

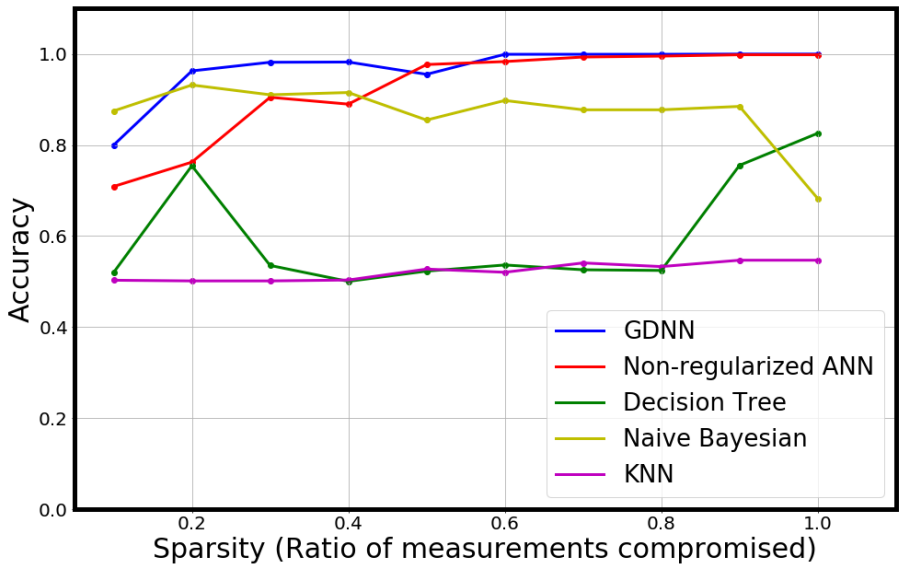

Figure 6.7: The test accuracy of all models with varying sparsity test sets on the IEEE 30-bus system

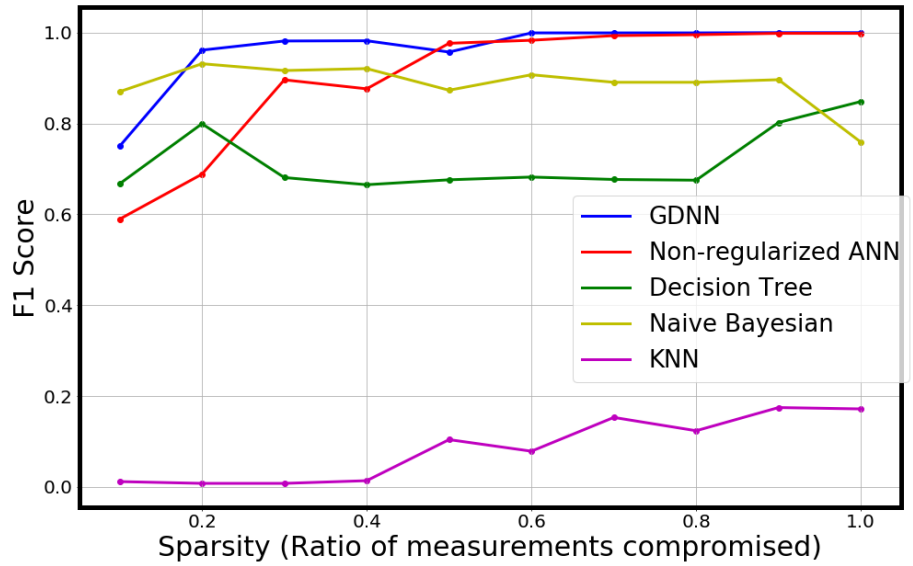

Figure 6.8: The F1-score of all models with varying sparsity test sets on the IEEE 30-bus system

6.2.4 General Model Performance

To test the scalability of the proposed algorithm, we tested on two other systems of different sizes, the IEEE 14-bus and the IEEE-57 bus systems. The average accuracy and F1-score across the entire range of sparsity is recorded for each algorithm.

As can be seen in tables 6.6, 6.7, and 6.8, GDNN, achieves superior performance on all three systems. The average accuracy and F1-score of the GDNN model is generally higher than the other machine learning models in comparison. Additionally, Table 6.9 demonstrates that training time of GDNN is significantly better than a similarly structured ANN which demonstrates computational efficiency.

Table 6.6: Performance of machine learning models on the IEEE 14-bus system

Model	Average Accuracy	Average F1-Score	Average MCC
KNN	60.66%	34.25%	0.345
Naive Bayesian	87.34%	88.15%	0.754
Decision Tree	82.80%	85.28%	0.689
ANN	90.45%	88.65%	0.824
GDNN	98.16%	98.04%	0.964
GLM [143]	95.71%	97.65%	N/A
GBM [143]	95.87%	97.74%	N/A
ELM-based [144]	95.31%	N/A	N/A

Table 6.7: Performance of machine learning models on the IEEE 30-bus system

Model	Average Accuracy	Average F1-Score	Average MCC
KNN	52.26%	8.48%	0.150
Naive Bayesian	87.08%	88.57%	0.758
Decision Tree	60.02%	71.75%	0.323
ANN	92.13%	89.97%	0.852
GDNN	96.81%	96.31%	0.937

Table 6.8: Performance of machine learning models on the IEEE 57-bus system

Model	Average Accuracy	Average F1-Score	Average MCC
KNN	93.60%	92.95%	0.875
Naive Bayesian	68.59%	78.56%	0.478
Decision Tree	55.45%	69.59%	0.296
ANN	88.21%	81.71%	0.755
GDNN	95.39%	93.87%	0.779
KPCA - Extra Trees [145]	98.20%	N/A	N/A

Table 6.9: Training Time of each machine learning algorithm (in seconds) for each power system

	IEEE 14-bus	**IEEE 30-bus**	**IEEE 57-bus**
KNN	0.83	1.58	3.15
Naive Bayesian	0.23	0.39	0.90
Decision Tree	52.63	110.36	66.77
ANN	767.0 (100 Epochs)	822.0 (100 Epochs)	1073.0 (100 Epochs)
GDNN	174.0 (15 Epochs)	252.0 (28 Epochs)	254.0 (15 Epochs)

6.2.5 Imbalance Testing

Intelligent models rely on pattern identification from data collected from the systems they are deployed on. Data of real power systems are imbalanced due to the low availability of attack data. This data imbalance can cause biased classification performance from the machine learning algorithms in which the minority class is classified with very low accuracy. As such, intelligent models must be capable of handling large imbalance in the data. Therefore, imbalance testing is performed on all algorithms for the IEEE 30-bus system. Figure 6.9 shows the accuracy of the models when trained on data with varying degree of imbalance. This figure demonstrates superior performance of the proposed GDNN algorithm on imbalanced data. The GDNN algorithm achieves higher accuracy than the other models when trained with data containing a low percentage of attack samples.

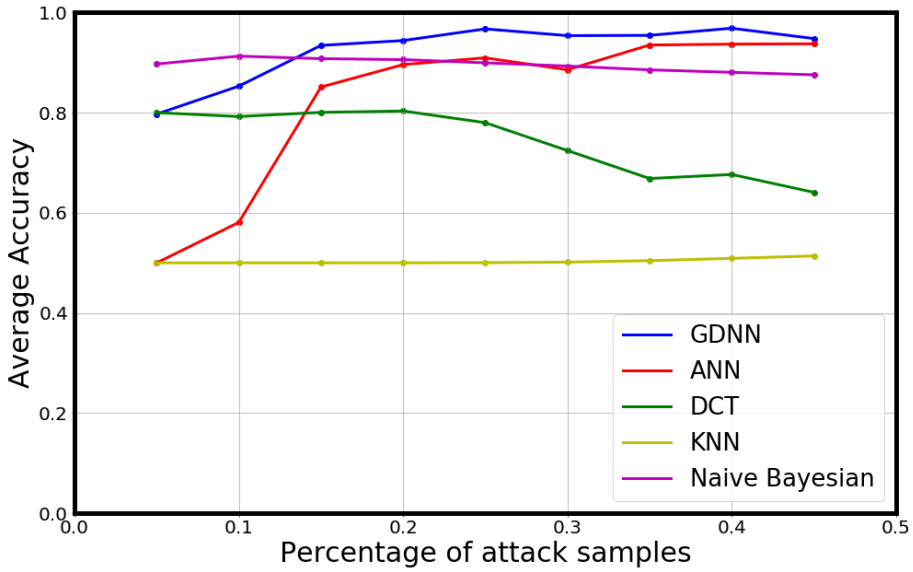

Figure 6.9: The average accuracy of machine learning models with varying degrees of imbalance in training data on the IEEE 30-bus system

6.3 Summary

In this chapter, the experimental process of demonstrating the contributions was demonstrated through two experiments. The first is to test three heuristic feature selection tech-

niques with three different machine learning classifiers, and the second is to test the GDNN method for various degrees of attack sparsity and data imbalance. Through the results demonstrated in these two experiments, it was shown that heuristic optimization algorithms are an effective technique of enhancing the computational efficiency of machine learning classifiers in the context of attack detection. Furthermore, the GDNN method proved successful at detecting attacks of various sparsity in all three test power systems with high accuracy. It was also successful at maintaining the high accuracy under various degrees of data imbalance. Such results are promising when considering scalability to larger power systems.

Chapter 7

Conclusions

In this thesis, it was demonstrated that attack detection in smart cyber-physical grids can be achieved by machine learning. Furthermore, heuristic optimization algorithms were shown to be successful at increasing the computational efficiency of machine learning classifiers through maintaining test classification accuracy with significantly lower number of features. Additionally, a deep neural network with regularization methods and hybrid learning optimizer is considered for the problem of varying attack sparsity and data imbalance. It was demonstrated that regularization methods such as dropout layers and L2-regularization are effective at detecting attacks of low sparsity, which is a challenging task for most machine learning classifiers. Furthermore, the implementation of Adadelta optimizer during the training phase proved to significantly reduce training time.

The results of this research show that the contributions of this thesis have all been met. It was found that the heuristic methods proposed have reduced the number of features by as much as 76.5% while reducing the accuracy by only 0.1% on the smallest test case. On the largest test case, the number of features were reduced by 59.87% in addition to increasing the accuracy by 5.7%. The results were successful on all test cases thereby satisfying the initial conditions of this contribution. Similarly, the second contribution was met because the proposed method yielded an increase of 7.71%, 4.68%, and 7.18% in accuracy on the IEEE 14-bus, 30-bus, and 57-bus respectively; which is more than the initial condition of the contribution. Furthermore the proposed methods increased the F1-score by 9.39%, 6.34%, and 12.16% in the IEEE 14-bus, 30-bus, and 57-bus respectively.

The research done in this thesis contributes to increasing computational efficiency, and more importantly, tackles two gaps existing in literature: detecting stealthy data injection attacks at varying sparsity, and achieving high attack detection accuracy through training on

imbalanced data. These concerns are critical to scalability and deployment to real systems; particularly due to the lack of available attack data from real systems. The robustness of the final proposed model to varying degrees of attack sparsity shows promise in detecting injection attacks at earlier stages prior to propagating to infect a large portion of the system. Furthermore, the robustness of the model to imbalance in the training data shows promise in real system application and scalability.

7.1 Future Work

Extending the research of this thesis can be aimed at multi-class attack detection in which the source of the attack as well as the attack type can be identified by the deep learning algorithm. Furthermore, testing can be done on attacks of varying magnitude, as opposed to varying sparsity. Detecting attacks of low magnitude is also challenging and can help address some of the research gaps in the field of smart grid cybersecurity.

To achieve multi-class attack detection, a more comprehensive simulation framework is necessary. Data of attacks on real systems is scarce and data with specific attack labels are extremely rare. As such, a simulation framework must be designed that incorporates multiple types of attacks occurring across various locations of the system. Such a framework can be used to generate data to train and test methods aimed at classifying attack types and source location. Furthermore, multi-view simulation methods can be proposed which integrate the operational layer with the communication layer to better simulate real-time attacks across all layers of the smart grid.

Additionally, to achieve multi-class attack detection, ensemble deep learning methods are recommended; particularly when it comes to multi-view systems. This is due to the variety of potential cyber threats that can stem from multiple layers of the smart grid and the types of features associated with each. As such, ensemble methods can be designed to account for the varying feature types as well as attack categories. Deep learning ensemble methods can be structured such that individual networks are trained for specific attack types and the ensemble algorithm makes decisions based on the outcome of each sub-network.

Further future development can be aimed towards real-time implementation of the attack detection system. Minor tuning is likely necessary to deploy the model proposed in this thesis on a real power system including scaling the number of input features to the logged measurements. Prior to deployment, testing should be done on data collected from the target power system to ensure performance is maintained.

Bibliography

[1] Elmustafa sayed ali ahmed and Zeinab Kamal Aldein Mohammed. Internet of things applications, challenges and related future technologies. *world scientific news*, 01 2017.

[2] Daniele Miorandi, Sabrina Sicari, Francesco De Pellegrini, and Imrich Chlamtac. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7):1497–1516, September 2012.

[3] The Smart Grid Interoperability PanelSmart Grid Cybersecurity Committee. Guidelines for smart grid cybersecurity. Technical Report NIST IR 7628r1, National Institute of Standards and Technology, September 2014.

[4] H. Karimipour and V. Dinavahi. Robust Massively Parallel Dynamic State Estimation of Power Systems Against Cyber-Attack. *IEEE Access*, 6:2984–2995, 2018.

[5] H. Karimipour and V. Dinavahi. Extended Kalman Filter-Based Parallel Dynamic State Estimation. *IEEE Transactions on Smart Grid*, 6(3):1539–1549, May 2015.

[6] Andrs Luque-Ayala and Simon Marvin. Developing a critical understanding of smart urbanism? *Urban Studies*, 52(12):2105–2116, September 2015.

[7] H. Khurana, M. Hadley, Ning Lu, and D.A. Frincke. Smart-grid security issues. *IEEE Security & Privacy Magazine*, 8(1):81–85, January 2010.

[8] Henry Mwiki, Tooska Dargahi, Ali Dehghantanha, and Kim-Kwang Raymond Choo. Analysis and Triage of Advanced Hacking Groups Targeting Western Countries Critical National Infrastructure: APT28, RED October, and Regin. In Dimitris Gritzalis, Marianthi Theocharidou, and George Stergiopoulos, editors, *Critical Infrastructure Security and Resilience: Theories, Methods, Tools and Technologies*, Advanced Sciences and Technologies for Security Applications, pages 221–244. Springer International Publishing, Cham, 2019.

[9] S. Walker-Roberts, M. Hammoudeh, and A. Dehghantanha. A systematic review of the availability and efficacy of countermeasures to internal threats in healthcare critical infrastructure. *IEEE Access*, 6:25167–25177, 2018.

[10] AN Jahromi, S. Hashemi, A. Dehghantanha, R. Parizi, and KKR Choo. An enhanced stacked lstm method with no random initialization for malware threat hunting in safety and time-critical systems. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2019.

[11] W. Meng, R. Ma, and H. Chen. Smart grid neighborhood area networks: a survey. *IEEE Network*, 28(1):24–32, January 2014.

[12] Y. Saleem, N. Crespi, M. H. Rehmani, and R. Copeland. Internet of things-aided smart grid: Technologies, architectures, applications, prototypes, and future research directions. *IEEE Access*, 7:62962–63003, 2019.

[13] Ali Tajer, Soummyar Kar, H. Vincent Poor, and Shuguang Cui. Distributed joint cyber attack detection and state recovery in smart grids. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 202–207, Brussels, Belgium, October 2011. IEEE.

[14] Shuguang Cui, Zhu Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer. Coordinated Data-Injection Attack and Detection in the Smart Grid: A Detailed Look at Enriching Detection Solutions. *IEEE Signal Processing Magazine*, 29(5):106–115, September 2012.

[15] Jacob Sakhnini, Hadis Karimipour, Ali Dehghantanha, Reza M. Parizi, and Gautam Srivastava. Security aspects of internet of things aided smart grids: A bibliometric survey. *Internet of Things*, page 100111, 2019.

[16] Nan Wu and Xiangdong Li. RFID Applications in Cyber-Physical System. *Deploying RFID - Challenges, Solutions, and Open Issues*, August 2011.

[17] Danda B. Rawat, Joel J. P. C. Rodrigues, and Ivan Stojmenovic. *Cyber-Physical Systems: From Theory to Practice*. CRC Press, October 2015. Google-Books-ID: _CzSCgAAQBAJ.

[18] National Academies of Sciences, Engineering, and Medicine. *A 21st Century Cyber-Physical Systems Education*. The National Academies Press, Washington, DC, 2016.

[19] H. Karimipour and V. Dinavahi. Accelerated parallel WLS state estimation for large-scale power systems on GPU. In *2013 North American Power Symposium (NAPS)*, pages 1–6, September 2013.

[20] X. Fang, S. Misra, G. Xue, and D. Yang. Smart grid the new and improved power grid: A survey. *IEEE Communications Surveys Tutorials*, 14(4):944–980, Fourth 2012.

[21] H. Karimipour and V. Dinavahi. Parallel domain decomposition based distributed state estimation for large-scale power systems. In *2015 IEEE/IAS 51st Industrial Commercial Power Systems Technical Conference (I CPS)*, pages 1–5, May 2015.

[22] Behnam Neyshabur, Srinadh Bhojanapalli, David McAllester, and Nathan Srebro. Exploring Generalization in Deep Learning. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pages 5949–5958, USA, 2017. Curran Associates Inc. event-place: Long Beach, California, USA.

[23] H. Karimipour and V. Dinavahi. Extended kalman filter-based parallel dynamic state estimation. *IEEE Transactions on Smart Grid*, 6(3):1539–1549, May 2015.

[24] H. Karimipour, A. Dehghantanha, R. M. Parizi, K. R. Choo, and H. Leung. A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. *IEEE Access*, 7:80778–80788, 2019.

[25] J Duncan Glover, Mulukutla S Sarma, and Thomas J Overbye. Power System Analysis and Design. page 853, 2010.

[26] Feng Ye. *Smart grid communication infrastructures : big data, cloud computing, and security*. John Wiley and Sons, Hoboken, New Jersey, 2018.

[27] A. G. Bruce. Reliability analysis of electric utility scada systems. In *Proceedings of the 20th International Conference on Power Industry Computer Applications*, pages 200–205, May 1997.

[28] Y. Yan, Y. Qian, H. Sharif, and D. Tipper. A survey on smart grid communication infrastructures: Motivations, requirements and challenges. *IEEE Communications Surveys Tutorials*, 15(1):5–20, First 2013.

[29] Mete Ozay, Inaki Esnaola, Fatos T. Yarman-Vural, Sanjeev R. Kulkarni, and H. Vincent Poor. Sparse attack construction and state estimation in the smart grid: Centralized and distributed models. *IEEE Journal on Selected Areas in Communications*, 31:1306–1318, 2013.

[30] H. Karimipour and V. Dinavahi. Parallel relaxation-based joint dynamic state estimation of large-scale power systems. *IET Generation, Transmission Distribution*, 10(2):452–459, 2016.

[31] Henry Mwiki, Tooska Dargahi, Ali Dehghantanha, and Kim-Kwang Raymond Choo. *Analysis and Triage of Advanced Hacking Groups Targeting Western Countries Critical National Infrastructure: APT28, RED October, and Regin: Theories, Methods, Tools and Technologies*, pages 221–244. 01 2019.

[32] United States Government Accountability Office. *Cybersecurity: Challenges in Securing the Electricity Grid*. CreateSpace Independent Publishing Platform, November 2012.

[33] Sanjay Goel and Yuan Hong. Chapter 1 Security Challenges in Smart Grid Implementation. 2015.

[34] P. Chen, S. Yang, J. A. McCann, J. Lin, and X. Yang. Detection of false data injection attacks in smart-grid systems. *IEEE Communications Magazine*, 53(2):206–213, Feb 2015.

[35] Rakesh Bobba, Katherine Davis, Qiyan Wang, Himanshu Khurana, Klara Nahrstedt, and Thomas J Overbye. Detecting false data injection attacks on dc state estimation. 01 2010.

[36] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, July 2017.

[37] Yao Liu, Peng Ning, and Michael K. Reiter. False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur.*, 14(1):13:1–13:33, June 2011.

[38] Nick Guenther and Matthias Schonlau. Support vector machines. *The Stata Journal*, 16(4):917–937, 2016.

[39] Abdulrahaman Okino Otuoze, Mohd Wazir Mustafa, and Raja Masood Larik. Smart grids security challenges: Classification by sources of threats. *Journal of Electrical Systems and Information Technology*, 5(3):468–483, December 2018.

[40] P. McDaniel and S. McLaughlin. Security and Privacy Challenges in the Smart Grid. *IEEE Security Privacy*, 7(3):75–77, May 2009.

[41] Sanjay Goel and Yuan Hong. Security Challenges in Smart Grid Implementation. In Sanjay Goel, Yuan Hong, Vagelis Papakonstantinou, and Dariusz Kloza, editors, *Smart Grid Security*, SpringerBriefs in Cybersecurity, pages 1–39. Springer London, London, 2015.

[42] V. Delgado-Gomes, J. F. Martins, C. Lima, and P. N. Borza. Smart grid security issues. In *2015 9th International Conference on Compatibility and Power Electronics (CPE)*, pages 534–538, June 2015.

[43] Paria Jokar, Nasim Arianpoo, and Victor C. M. Leung. Spoofing Detection in IEEE 802.15.4 Networks Based on Received Signal Strength. *Ad Hoc Netw.*, 11(8):2648–2660, November 2013.

[44] P. Risbud, N. Gatsis, and A. Taha. Vulnerability Analysis of Smart Grids to GPS Spoofing. In *2018 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–1, August 2018.

[45] P. Pradhan, K. Nagananda, P. Venkitasubramaniam, S. Kishore, and R. S. Blum. Gps spoofing attack characterization and detection in smart grids. In *2016 IEEE Conference on Communications and Network Security (CNS)*, pages 391–395, Oct 2016.

[46] J. Zhao, J. Wang, and L. Yin. Detection and control against replay attacks in smart grid. In *2016 12th International Conference on Computational Intelligence and Security (CIS)*, pages 624–627, Dec 2016.

[47] T. Tran, O. Shin, and J. Lee. Detection of replay attacks in smart grid systems. In *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*, pages 298–302, Jan 2013.

[48] G. Jinhua and X. Kejian. ARP spoofing detection algorithm using ICMP protocol. In *2013 International Conference on Computer Communication and Informatics*, pages 1–6, January 2013.

[49] D. Sharma, O. Khan, and N. Manchanda. Detection of ARP Spoofing: A command line execution method. In *2014 International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 861–864, March 2014.

[50] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, E. G. Im, Z. Q. Yao, B. Pranggono, and H. F. Wang. Man-in-the-middle attack test-bed investigating cyber-security vulnerabilities in Smart Grid SCADA systems. In *International Conference on Sustainable Power Generation and Supply (SUPERGEN 2012)*, pages 1–8, September 2012.

[51] P. Yi, T. Zhu, Q. Zhang, Y. Wu, and J. Li. A denial of service attack in advanced metering infrastructure network. In *2014 IEEE International Conference on Communications (ICC)*, pages 1029–1034, June 2014.

[52] Yonghe Guo, Chee-Wooi Ten, Shiyan Hu, and Wayne Weaver. Modeling Distributed Denial of Service Attack in Advanced Metering Infrastructure. February 2015.

[53] Chakib Bekara. Security issues and challenges for the iot-based smart grid. *Procedia Computer Science*, 34:532 – 537, 2014. The 9th International Conference on Future Networks and Communications (FNC'14)/The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)/Affiliated Workshops.

[54] Jiwei Tian, Buhong Wang, and Xia Li. Data-Driven and Low-Sparsity False Data Injection Attacks in Smart Grid. *Security and Communication Networks*, 2018:1–11, September 2018.

[55] Jacob Sakhnini, Hadis Karimipour, and Ali Dehghantanha. Smart Grid Cyber Attacks Detection using Supervised Learning and Heuristic Feature Selection. In *IEEE Int. Conf. on Smart Energy Grid Engineering (SEGE)*, pages 1–5, July 2019.

[56] H. Karimipour and V. Dinavahi. On false data injection attack against dynamic state estimation on smart power grids. In *2017 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, pages 388–393, August 2017.

[57] Xuan Liu and Zuyi Li. False data attack models, impact analyses and defense strategies in the electricity grid. *The Electricity Journal*, 30(4):35–42, May 2017.

[58] Xuan Liu and Zuyi Li. False Data Attacks Against AC State Estimation With Incomplete Network Information. *IEEE Trans. Smart Grid*, 8(5), September 2017.

[59] Y. Wang, M. M. Amin, J. Fu, and H. B. Moussa. A Novel Data Analytical Approach for False Data Injection Cyber-Physical Attack Mitigation in Smart Grids. *IEEE Access*, 5:26022–26033, 2017.

[60] Huixin Zhong, Dajun Du, Chuanjiang Li, and Xue Li. A Novel Sparse False Data Injection Attack Method in Smart Grids with Incomplete Power Network Information. *Complexity*, 2018:8503825–8503825, 2018.

[61] Baoyao Wang, Peidong Zhu, Yingwen Chen, Peng Xun, and Zhenyu Zhang. False Data Injection Attack Based on Hyperplane Migration of Support Vector Machine in Transmission Network of the Smart Grid. *Symmetry*, 10(5):165, May 2018.

[62] L. Lei, W. Yang, C. Yang, and H. B. Shi. False data injection attack on consensus-based distributed estimation: A TYPICAL ATTACK ON CONSENSUS-BASED DISTRIBUTED ESTIMATION. *Int. J. Robust. Nonlinear Control*, 2016.

[63] Liang Che, Xuan Liu, Zuyi Li, and Yunfeng Wen. False Data Injection Attacks Induced Sequential Outages in Power Systems. *IEEE Trans. Power Syst.*, 34(2):1513–1523, March 2019.

[64] Jeong-Won Kang, Il-Young Joo, and Dae-Hyun Choi. False Data Injection Attacks on Contingency Analysis: Attack Strategies and Impact Assessment. *IEEE Access*, 6:8841–8851, 2018.

[65] Yuzhe Li, Dawei Shi, and Tongwen Chen. False Data Injection Attacks on Networked Control Systems: A Stackelberg Game Analysis. *IEEE Trans. Automat. Contr.*, 63(10):3503–3509, October 2018.

[66] Bo Chai and Zaiyue Yang. Impacts of unreliable communication and modified regret matching based anti-jamming approach in smart microgrid. *Ad Hoc Networks*, 22:69–82, November 2014.

[67] Yuting Liu, Jinghuan Ma, Lingyang Song, and Zhu Han. Jamming Attack in Smart Grid with Dynamic Gaming Theory. 2014.

[68] K. Gai, M. Qiu, Z. Ming, H. Zhao, and L. Qiu. Spoofing-Jamming Attack Strategy Using Optimal Power Distributions in Wireless Smart Grid Networks. *IEEE Transactions on Smart Grid*, 8(5):2431–2439, September 2017.

[69] K. Tazi, F. Abdi, and M. F. Abbou. Review on cyber-physical security of the smart grid: Attacks and defense mechanisms. In *2015 3rd International Renewable and Sustainable Energy Conference (IRSEC)*, pages 1–6, December 2015.

[70] Gran N. Ericsson. Cyber Security and Power System CommunicationEssential Parts of a Smart Grid Infrastructure. *IEEE Transactions on Power Delivery*, 25(3):1501–1507, July 2010.

[71] Y. Kwon, H. K. Kim, K. M. Koumadi, Y. H. Lim, and J. I. Lim. Automated vulnerability analysis technique for smart grid infrastructure. In *2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, April 2017.

[72] P. Chopade and M. Bikdash. Modeling for survivability of Smart Power Grid when subject to severe emergencies and vulnerability. In *2012 Proceedings of IEEE Southeastcon*, pages 1–6, March 2012.

[73] H. Ying, Y. Zhang, L. Han, Y. Cheng, J. Li, X. Ji, and W. Xu. Detecting Buffer-Overflow Vulnerabilities in Smart Grid Devices via Automatic Static Analysis. pages 813–817, March 2019.

[74] G. Chen, J. Zhao, Z. Y. Dong, and S. R. Weller. Complex network theory based power grid vulnerability assessment from past to future. In *9th IET International Conference on Advances in Power System Control, Operation and Management (APSCOM 2012)*, pages 1–6, November 2012.

[75] A. Srivastava, T. Morris, T. Ernster, C. Vellaithurai, S. Pan, and U. Adhikari. Modeling Cyber-Physical Vulnerability of the Smart Grid With Incomplete Information. *IEEE Transactions on Smart Grid*, 4(1):235–244, March 2013.

[76] Y. Zhu, J. Yan, Y. Sun, and H. He. Revealing cascading failure vulnerability in power grids using risk-graph. *IEEE Transactions on Parallel and Distributed Systems*, 25(12):3274–3284, December 2014.

[77] S. Paul and Z. Ni. Vulnerability analysis for simultaneous attack in smart grid security. In *2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, April 2017.

[78] V. Dehalwar, A. Kalam, M. L. Kolhe, and A. Zayegh. Review of web-based information security threats in smart grid. In *2017 7th International Conference on Power Systems (ICPS)*, pages 849–853, December 2017.

[79] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He. Joint Substation-Transmission Line Vulnerability Assessment Against the Smart Grid. *IEEE Transactions on Information Forensics and Security*, 10(5):1010–1024, May 2015.

[80] P. Chopade and M. Bikdash. Structural and functional vulnerability analysis for survivability of Smart Grid and SCADA network under severe emergencies and WMD attacks. In *2013 IEEE International Conference on Technologies for Homeland Security (HST)*, pages 99–105, November 2013.

[81] Danda B. Rawat and Chandra Bajracharya. Detection of False Data Injection Attacks in Smart Grid Communication Systems. *IEEE Signal Processing Letters*, 22(10):1652–1656, October 2015.

[82] Mehmet Necip Kurt, Yasin Yilmaz, and Xiaodong Wang. Distributed Quickest Detection of Cyber-Attacks in Smart Grid. *IEEE Transactions on Information Forensics and Security*, 13(8):2015–2030, August 2018.

[83] Yasser Shoukry, Michelle Chong, Masashi Wakaiki, Pierluigi Nuzzo, Alberto Sangiovanni-Vincentelli, Sanjit A. Seshia, JOÃo P. Hespanha, and Paulo Tabuada. Smt-based observer design for cyber-physical systems under sensor attacks. *ACM Trans. Cyber-Phys. Syst.*, 2(1), January 2018.

[84] S. Z. Yong, M. Q. Foo, and E. Frazzoli. Robust and resilient estimation for cyber-physical systems under adversarial attacks. In *2016 American Control Conference (ACC)*, pages 308–315, July 2016.

[85] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson. Revealing stealthy attacks in control systems. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1806–1813, Oct 2012.

[86] Dina Hadžiosmanoviundefined, Robin Sommer, Emmanuele Zambon, and Pieter H. Hartel. Through the eye of the plc: Semantic security monitoring for industrial processes. In *Proceedings of the 30th Annual Computer Security Applications Conference*, ACSAC 14, page 126135, New York, NY, USA, 2014. Association for Computing Machinery.

[87] Sajal Bhatia, Nishchal Kush, Chris I. Djamaludin, Ayodeji J. Akande, and Ernest Foo. Practical modbus flooding attack and detection. In *AISC*, 2014.

[88] B. Chen, N. Pattanaik, A. Goulart, K. L. Butler-purry, and D. Kundur. Implementing attacks for modbus/tcp protocol in a real-time cyber physical system test bed. In *2015 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)*, pages 1–6, May 2015.

[89] Santiago Figueroa-Lorenzo, Javier Aorga, and Saioa Arrizabalaga. A role-based access control model in modbus scada systems. a centralized model approach. 19, October 2019.

[90] Yilin Mo, Tiffany Hyun-Jin Kim, K. Brancik, D. Dickinson, Heejo Lee, A. Perrig, and B. Sinopoli. CyberPhysical Security of a Smart Grid Infrastructure. *Proc. IEEE*, 100(1):195–209, January 2012.

[91] M. Esmalifalak, , R. Zheng, and Z. Han. Detecting stealthy false data injection using machine learning in smart grid. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 808–813, Dec 2013.

[92] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor. Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1773–1786, Aug 2016.

[93] J. Yan, B. Tang, and H. He. Detection of false data attacks in smart grid with supervised learning. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1395–1402, July 2016.

[94] Y. Wang, M. M. Amin, J. Fu, and H. B. Moussa. A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids. *IEEE Access*, 5:26022–26033, 2017.

[95] O. Kosut, L. Jia, R. J. Thomas, and L. Tong. Malicious data attacks on the smart grid. *IEEE Transactions on Smart Grid*, 2(4):645–658, Dec 2011.

[96] S. Li, Y. Ylmaz, and X. Wang. Quickest detection of false data injection attack in wide-area smart grids. *IEEE Transactions on Smart Grid*, 6(6):2725–2735, Nov 2015.

[97] S. Pan, T. Morris, and U. Adhikari. Developing a hybrid intrusion detection system using data mining for power systems. *IEEE Transactions on Smart Grid*, 6(6):3104–3113, Nov 2015.

[98] Eklas Hossain, Imtiaj Khan, Fuad Un-Noor, Sarder Shazali Sikander, and Md. Samiul Sunny. Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access*, PP:1–1, 01 2019.

[99] Sara Mohammadi, Hamid Mirvaziri, Mostafa Ghazizadeh-Ahsaee, and Hadis Karimipour. Cyber intrusion detection by combined feature selection algorithm. *Journal of Information Security and Applications*, 44:80–88, February 2019.

[100] S. Ahmed, Y. Lee, S. Hyun, and I. Koo. Feature selection based detection of covert cyber deception assaults in smart grid communications networks using machine learning. *IEEE Access*, 6:27518–27529, 2018.

[101] Lukas Malina, Gautam Srivastava, Petr Dzurenda, Jan Hajny, and Radek Fujdiak. A secure publish/subscribe protocol for internet of things. In *Proceedings of the 2019 14th International Conference on Availability, Reliability and Security (ARES 2019), Canterbury, UK*, pages 26–29, 2019.

[102] Ashutosh Dhar Dwivedi, Gautam Srivastava, Shalini Dhar, and Rajani Singh. A decentralized privacy-preserving healthcare blockchain for iot. *Sensors*, 19(2):326, 2019.

[103] A. R. Metke and R. L. Ekl. Security Technology for Smart Grid Networks. *IEEE Transactions on Smart Grid*, 1(1):99–107, June 2010.

[104] Ashutosh Dhar Dwivedi, Pawel Morawiecki, and Gautam Srivastava. Differential cryptanalysis of round-reduced speck suitable for internet of things devices. *IEEE Access*, 7:16476–16486, 2019.

[105] Ashutosh Dhar Dwivedi and Gautam Srivastava. Differential cryptanalysis of round-reduced lea. *IEEE Access*, 6:79105–79113, 2018.

[106] D. Wu and C. Zhou. Fault-Tolerant and Scalable Key Management for Smart Grid. *IEEE Transactions on Smart Grid*, 2(2):375–381, June 2011.

[107] W. Leea, T. Chen, W. Sun, and K. I. Ho. An S/Key-like One-Time Password Authentication Scheme Using Smart Cards for Smart Meter. In *2014 28th International Conference on Advanced Information Networking and Applications Workshops*, pages 281–286, May 2014.

[108] H. Nicanfar, P. Jokar, and V. C. M. Leung. Smart grid authentication and key management for unicast and multicast communications. In *2011 IEEE PES Innovative Smart Grid Technologies*, pages 1–8, November 2011.

[109] K. Khanna, B. K. Panigrahi, and A. Joshi. Feasibility and mitigation of false data injection attacks in smart grid. In *2016 IEEE 6th International Conference on Power Systems (ICPS)*, pages 1–6, March 2016.

[110] B. Li, R. Lu, G. Xiao, Z. Su, and A. Ghorbani. PAMA: A Proactive Approach to Mitigate False Data Injection Attacks in Smart Grids. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, December 2018.

[111] K. Sha, N. Alatrash, and Z. Wang. A Secure and Efficient Framework to Read Isolated Smart Grid Devices. *IEEE Transactions on Smart Grid*, 8(6):2519–2531, November 2017.

[112] C. Rottondi, M. Savi, D. Polenghi, G. Verticale, and C. Krau. Implementation of a protocol for secure distributed aggregation of smart metering data. In *2012 International Conference on Smart Grid Technology, Economics and Policies (SG-TEP)*, pages 1–4, December 2012.

[113] Y. Shovgenya, F. Skopik, and K. Theuerkauf. On demand for situational awareness for preventing attacks on the smart grid. In *2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, pages 1–4, June 2015.

[114] M. Q. Ali, E. Al-Shaer, and Q. Duan. Randomizing AMI configuration for proactive defense in smart grid. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 618–623, October 2013.

[115] P. Srikantha and D. Kundur. A DER Attack-Mitigation Differential Game for Smart Grid Security Analysis. *IEEE Transactions on Smart Grid*, 7(3):1476–1485, May 2016.

[116] R. Hewett, S. Rudrapattana, and P. Kijsanayothin. Smart Grid security: Deriving informed decisions from cyber attack game analysis. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 946–951, November 2014.

[117] M. Ni, A. K. Srivastava, R. Bo, and J. Yan. Design of A Game Theory Based Defense System for Power System Cyber Security. In *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pages 1049–1054, July 2017.

[118] M. H. Ranjbar, M. Kheradmandi, and A. Pirayesh. A Linear Game Framework for Defending Power Systems against Intelligent Physical Attacks. pages 1–1, 2019.

[119] X. Yang, X. He, J. Lin, W. Yu, and Q. Yang. A Game-Theoretic Model on Coalitional Attacks in Smart Grid. In *2016 IEEE Trustcom/BigDataSE/ISPA*, pages 435–442, August 2016.

[120] L. Wei, A. I. Sarwat, W. Saad, and S. Biswas. Stochastic Games for Power Grid Protection Against Coordinated Cyber-Physical Attacks. *IEEE Transactions on Smart Grid*, 9(2):684–694, March 2018.

[121] M. Shange, J. Lin, X. Zhang, and C. Xu. A game-theory analysis of the rat-group attack in smart grids. In *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pages 1–6, April 2014.

[122] A. Sanjab and W. Saad. Data Injection Attacks on Smart Grids With Multiple Adversaries: A Game-Theoretic Perspective. *IEEE Transactions on Smart Grid*, 7(4):2038–2049, July 2016.

[123] Sandhya Rani, Jorika Vedika, Vancha Maheshwar Reddy, Burri Sandhya Rani, and Ch Veera Reddy. Game Theory based Defense Strategy against Denial of Service Attack using Puzzles. 2013.

[124] P. Srikantha and D. Kundur. Denial of service attacks and mitigation for stability in cyber-enabled power grid. In *2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, February 2015.

[125] J. Ma, Y. Liu, L. Song, and Z. Han. Multiact Dynamic Game Strategy for Jamming Attack in Electricity Market. *IEEE Transactions on Smart Grid*, 6(5):2273–2282, September 2015.

[126] C. Y. T. Ma, D. K. Y. Yau, X. Lou, and N. S. V. Rao. Markov Game Analysis for Attack-Defense of Power Networks Under Possible Misinformation. *IEEE Transactions on Power Systems*, 28(2):1676–1686, May 2013.

[127] K. Wang, M. Du, S. Maharjan, and Y. Sun. Strategic Honeypot Game Model for Distributed Denial of Service Attacks in the Smart Grid. *IEEE Transactions on Smart Grid*, 8(5):2474–2482, September 2017.

[128] D. Mashima, B. Chen, P. Gunathilaka, and E. L. Tjiong. Towards a grid-wide, high-fidelity electrical substation honeynet. In *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 89–95, October 2017.

[129] J. Hastings, D. M. Laverty, and D. J. Morrow. Tracking smart grid hackers. In *2014 49th International Universities Power Engineering Conference (UPEC)*, pages 1–5, September 2014.

[130] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg. Network-Aware Mitigation of Data Integrity Attacks on Power System State Estimation. *IEEE Journal on Selected Areas in Communications*, 30(6):1108–1118, July 2012.

[131] A. Delgadillo, J. M. Arroyo, and N. Alguacil. Analysis of Electric Grid Interdiction With Line Switching. *IEEE Transactions on Power Systems*, 25(2):633–641, May 2010.

[132] J. M. Arroyo and F. J. Fernandez. A Genetic Algorithm Approach for the Analysis of Electric Grid Interdiction with Line Switching. In *2009 15th International Conference on Intelligent System Applications to Power Systems*, pages 1–6, November 2009.

[133] Y. Yao, T. Edmunds, D. Papageorgiou, and R. Alvarez. Trilevel Optimization in Power Network Defense. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(4):712–718, July 2007.

[134] Sara Mohammadi, H Mirvaziri, Mostafa Ghazizadeh-Ahsaee, and Hadis Karimipour. Cyber intrusion detection by combined feature selection algorithm. *Journal of Information Security and Applications*, 44:80–88, 02 2019.

[135] X. Yang and. Cuckoo search via lvy flights. In *2009 World Congress on Nature Biologically Inspired Computing (NaBIC)*, pages 210–214, Dec 2009.

[136] D. Rodrigues, L. A. M. Pereira, T. N. S. Almeida, J. P. Papa, A. N. Souza, C. C. O. Ramos, and X. Yang. Bcs: A binary cuckoo search algorithm for feature selection. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pages 465–468, May 2013.

[137] Saeed Ahmed, YoungDoo Lee, Seung Hyun, and Insoo Koo. Covert cyber assault detection in smart grid networks utilizing feature selection and euclidean distance-based machine learning. *Applied Sciences*, 8:772, 05 2018.

[138] B. Xue, M. Zhang, and W. N. Browne. Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE Transactions on Cybernetics*, 43(6):1656–1671, Dec 2013.

[139] Matthew D. Zeiler. Adadelta: An adaptive learning rate method. *ArXiv*, abs/1212.5701, 2012.

[140] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, Feb 2011.

[141] Davide Chicco and Giuseppe Jurman. The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1):6, 2020.

[142] Sabri Boughorbel, Fethi Jarray, and Mohammed El-Anbari. Optimal classifier for imbalanced data using matthews correlation coefficient metric. *PLOS ONE*, 12(6):1–17, 06 2017.

[143] M. Ashrafuzzaman, Y. Chakhchoukh, A. A. Jillepalli, P. T. Tosic, D. C. de Leon, F. T. Sheldon, and B. K. Johnson. Detecting stealthy false data injection attacks in power grids using deep learning. In *2018 14th International Wireless Communications Mobile Computing Conference (IWCMC)*, pages 219–225, June 2018.

[144] A. Ayad, H. E. Z. Farag, A. Youssef, and E. F. El-Saadany. Detection of false data injection attacks in smart grids using recurrent neural networks. In *2018 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, Feb 2018.

[145] M. R. Camana Acosta, S. Ahmed, C. E. Garcia, and I. Koo. Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks. *IEEE Access*, 8:19921–19933, 2020.