Modeling, Analysis and Countermeasures for Attack Propagation in Wide Area Measurement Systems

by

Hamed Sarjan

A thesis

presented to the Lakehead University in fulfillment

of the thesis requirement for the degree of

Master of Science in

Electrical engineering

Thunderbay, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisor: Dr. Amir Ameli
Assistant Professor, Department of Electrical engineering,
Lakehead University

Committee Member (1): Dr. Mohammad Nasir Uddin
Professor, Department of Electrical engineering,
Lakehead University

Committee Member (2): Dr. Qiang Wei
Assistant Professor, Department of Electrical engineering,
Lakehead University

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Power grids are critical cyber-physical systems that employ advanced Information and Communication Technologies (ICTs), such as Wide Area Measurement Systems (WAMSs), to deliver the energy to end users reliably and efficiently. WAMSs are used to collect real-time data from Phasor Measurement Units (PMUs) to improve the operator's situational awareness, as well as to enhance real-time monitoring and control of power systems. The WAMS, however, is vulnerable to cyber-attacks due to the susceptibility of its components—such as PMUs and Phasor Data Concentrators (PDCs)—and the lack of embedded security mechanisms in its communication protocols. Some more-destructive cyber-attacks, such as malware injection, can propagate themselves into the components of a WAMS through the communication network. Thus, in such attacks, an attacker can compromise a larger number of components, resulting in more-severe consequences. Therefore, investigating the propagation of cyber-attacks in WAMSs and devising effective countermeasures for this problem are of paramount importance. On this basis, this thesis initially develops a model to analyze cyber-attack propagation in WAMS. Then, the impacts of the attacker's capability and the network operator's defensive ability on attack propagation are investigated in detail. Such a study can elucidate the required security measures and defensive strategies to prevent the spread of cyber-attacks in WAMSs. Finally, a Learning-Based Framework (LBF) is developed to estimate the attacker's capability. Furthermore, it is imperative to conduct a comprehensive examination of mitigation strategies aimed at thwarting propagable attacks given their deleterious impact on WAMSs. On this basis, this thesis develops another LBF to estimate the required defensive strategy to counter the propagation of cyber-attacks in WAMSs. Afterwards, through solving a Linear Programming (LP) problem, this thesis develops a mitigation strategy to optimally reconfigure the communication network and reduce the contamination probability for critical PMUs and PDCs while maintaining the observability of the grid. The simulation results obtained from IEEE 6-Bus and 14-Bus test systems corroborate the effectiveness of the proposed model, LBFs and communication network reconfiguration strategy in analyzing and mitigating the propagation of cyber-attacks in WAMSs.

Acknowledgments

It has been a great pleasure working with the best team throughout my Master's program. First of all, I would like to express my sincere gratitude to my supervisor, *Dr. Amir Ameli*, for believing in me and for taking me on as a Master's student as well as for providing guidance, over the course of my Master's, across both intellectual and career pursuits.

I would like to sincerely thank *Dr. Mohsen Ghafouri*, from the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, for his constantly-available help and support during my research. I could not have reached this achievement without all of his support. I also want to thank all WAMPAC research team members for their support.

I am honored that this dissertation has been examined by Dr. Mohammad Nasir Uddin, and Dr. Qiang Wei. I owe respect and thanks to them for their time and insight.

Last but not least, I am very thankful to my parents and my spouse, who have supported me through difficult and easy times. I would not be here without their unconditional encouragement.

to

My beloved Yeganeh

Table of Contents

List of Figures

Acronyms

APT Advanced Persistent Threat. 7, 8

CPS Cyber-Physical System. 22

DoS Denial of Service. 13, 14

FDIA False Data Injection Attack. 15, 20

FNN Feed-forward Neural Network. xii, xiii, 37–40, 42, 43, 50–53

GPS Global Positioning System. 9, 19, 20

IDS Intrusion Detection System. 28, 31, 36, 40, 42, 46, 67

LAN Local Area Network. 20

LBF Learning-Based Framework. x, xiii, 23, 24, 30, 36, 38–40, 42–45, 50, 52, 53, 76

LP Linear Programming. 24, 62, 76

MSE Mean Squared Error. 42, 43, 52

NIST Institute of Standards and Technology. 1, 3

NN Neural Network. 37, 76

PDC Phasor Data Concentrator. 2, 8, 10, 11, 19, 20, 22, 23, 26, 28, 29, 31, 33, 36, 39, 40, 45, 56, 60, 64–68, 70

PMU Phasor Measurement Unit. xiii, 2, 4, 8–11, 16, 19, 20, 22, 23, 26, 28–31, 33, 36, 39, 45, 46, 56, 60, 61, 64–68, 70

RF Radio Frequency. 20

SG Smart Grid. 1–8, 12, 19, 21

WAMPAC Wide Area Monitoring, Protection, and Control. 8, 19

WAMS Wide Area Measurement System. viii–x, xii, 2–4, 8–27, 30, 31, 33–35, 40, 42, 44–46, 48–50, 55, 56, 59, 62, 76

Nomenclature

Parameters:

α_{ij}	Probability of attack propagation from directly susceptible node i to node j.
β	Attack propagation rate to directly susceptible nodes.
β'	Attack propagation rate to indirectly susceptible nodes.
γ	Recovery rate.
λ_{ij}	Probability of attack propagation from indirectly susceptible node i to node j.
$\mathbb{N}_{ij}^{m}$	Cardinality of the set Ω_{ij}^{m}.
θ	The rate at which recovered nodes become susceptible.
θ'	Hardening rate.
D_{ij}^{m}	Number of routers between nodes i and j in path m.
P_j	Probability of attack propagation to node j.

W Weighting factor for the probability of attack propagation to critical PMU$_j$.

Binary variables:

a_j Infection state (equals 1 if node j is infected, and 0 otherwise).

x_{ij} Connectivity state (1 if there is a path between nodes i and j, and 0 otherwise).

z_{uv} Communication link state (equals 1 if link (u, v) is connected, and 0 otherwise).

Sets:

Ω_{PDC} Set of all PDCs.

Ω_{PMU} Set of all PMUs.

Ω_{ij}^{m} Set of all nodes between nodes i and j through path m

C Critical nodes.

E Set of all nodes.

I Infected nodes.

M_{ij} Paths between node i and j.

R Recovered nodes.

S Indirectly susceptible nodes.

S' Directly susceptible nodes.

Chapter 1

Introduction

A Smart Grid (SG) is a distributed, intelligent network for delivering electricity that allows for bidirectional power and data flows. The SGs can respond to various situations and events and provide more efficient power delivery thanks to advanced information technology. Generally speaking, the SG can adapt its methods based on events that occur anywhere in the grid, including generation, transmission, distribution, and consumption. For instance, the SG may automatically alter the power flow and restore power delivery service in the case of a medium voltage transformer failure incident in the distribution grid. More precisely, the SG is an electricity system integrating clean energy generation, transmission, substations, distribution, and consumption with cyber-secure two-way communication technologies and computational intelligence [1]. As stated by the Institute of Standards and Technology (NIST) [2], there are seven domains in an SG, including generation, transmission, distribution, customer, markets, service provider, and operations, which consists of actors and applications. This model results in a clean, secure, reliable, resilient, efficient, and sustainable system. The traditional power system has served the world for many decades, but with the emergence of new technologies, the previously-used grid has become outdated and almost unable to meet the growing electricity demand. One of these technologies is WAMS which significantly increase the efficiency and reliability of power systems.

Due to an increase in power demands that have not been matched by an increase in generation capacity, the electrical grids are currently operating near their stability limits. This can be exploited by attackers and become a serious issue as it jeopardizes the operational security and reliability of the SGs, resulting in costly blackouts and environmental damage. Large blackouts and outages in recent years, such as the 2003 North American blackout and the 2015 Ukrainian blackout, have drawn attention to the SG's flaws. Critical problems with the system's stability arise from its inability to perform automated analysis, its poor response time, and its limited situational awareness. Therefore, managing, operating, and controlling the grid effectively and safely involves complicated technological activities that can be carried out at different times and in different locations. On this basis, WAMSs are used in power networks to improve the situational awareness of the operator, as well as to facilitate real-time control and protection decisions. In WAMSs, Phasor Data Concentrator (PDC)s collect time-synchronized data of PMUs through the communication system, and direct it to the control center to be used in wide-area control and protection applications.Due to the dependence of WAMSs on information and communication technologies, cyber-attacks can target these systems and propagate through them, i.e., infect a greater number of components by accessing and controlling a few of them [3].

1.1 Smart Grids and their security challenges

This section elaborates on security requirements and security challenges of SGs. The NIST has established three criteria— i.e., confidentiality, integrity, and availability—for ensuring the safety of data in the SGs, which are explained as follows [2].

1. **Confidentiality**

   Generally, confidentiality is the maintenance of appropriate limits on the distribution and use of private information. In other words, the confidentiality criteria necessitate preventing unauthorized parties from gaining access to or disclosing sensitive information. Confidentiality is broken when information is shared without permission [4]. For example, the transmission of sensitive customer data between the customer and third parties, including the consumption of customer, and billing information, must be encrypted and secured to prevent unauthorized access, manipulation, or use.

2. **Availability**

   The term "availability" refers to the fact that information must be retrieved and used when needed. Loss of availability causes disruption of access to information in a SG, hence it is widely regarded as the most important security requirement in the SG [4]. For instance, the unavailability of information flow due to various reasons can disrupt the functioning of the control system, thus preventing the network from being accessible to the system's operators for control purposes.

3. **Integrity**

   In the context of the SG, data integrity includes preventing unauthorized access, use, disclosure, modification, or deletion. A breach of integrity occurs when information is corrupted, changed, or deleted without detection. For example, by manipulating the PMUs in WAMS, an adversary can conduct a malicious attack on the state estimator [4].

1.1.1 Security Challenges

In order to implement SGs, a robust information and communication infrastructure must be built and deployed to facilitate a higher level of situational awareness and more efficient control. This is essential for the functioning of large-scale applications and systems, such as those involved in wide-area measurement and management of electricity demand, electricity storage and transmission, and distribution automation. However, integrating cyber and physical systems raises several difficulties due to human behavior, commercial interests, regulatory policy, and even political issues. Here are some of the cyber-security challenges faced by SGs [5].

1. **Lack of standardization**

One of the most important security challenges in SGs is the lack of standardization. Many parties, including utilities, regulators, equipment manufacturers, service providers, and end users, are involved in SGs. The employment of various technologies, protocols, and standards by each stakeholder may lead to problems with interoperability and a higher risk of cyberattacks. For instance, it may be challenging to establish secure and dependable communication across different SG devices since they may employ various communication protocols [6].

2. **Interconnectedness**

Another issue with the security of SGs is their inherent interconnectedness. In order to function properly, SGs rely on a vast infrastructure of interconnected devices, sensors, and communication infrastructures. Because of this interconnection, cybercriminals have more opportunities to find weak points in systems and launch attacks that might seriously harm the functioning of the grid and the availability of electricity. The security of SGs is complicated by their interconnected nature. First, as more devices and systems become grid-connected, the attack surface grows. As a result, it may be less difficult for malicious actors to identify entry points and launch attacks. Second, grid complexity is increased through interconnection, making it more difficult to detect and counteract cyber threats. Finally, a cyber attack on one area of the grid might spread to other parts and create significant damage due to the interconnected nature of the grid. For instance, an entire region's power may go out, if a cyberattack targets just one substation. Furthermore, the security of the entire grid, including access to private information, control systems, and physical infrastructure, might be compromised by an attack on a single device or sensor [7].

3. **Complexity**

   The system's complexity is another major obstacle to SG's security. Generally, several hardware, software, and communication protocol layers are required to operate an SG. Due to the complexity of SGs, there are numerous entry points through which attackers can have access to confidential information and sensitive infrastructures that disrupt the entire grid. The security of SGs is difficult to ensure because of the system's complexity. For instance, with a vast number of sensors and systems, monitoring the SG and detecting malicious activities is more complex. Furthermore, SG's increased vulnerability to cyber-attacks is exacerbated by the fact that different communication protocols are used [8]

4. **Limited security measures**

   SGs have a huge security issue due to inadequate security measures. While SGs are intended to be more secure than conventional power grids, keeping up with new security risks can be challenging due to the complexity and quick evolution of the technology. As a result, many SG stakeholders may be susceptible to cyber-attacks due to the lack of robust security mechanisms. It's possible, for instance, that the firewalls and intrusion detection systems utilized by some utilities are insufficient. It's also possible that certain manufacturers of electronic gadgets may not use strong security measures like encryption or secure boot, leaving their products open to hacking. In addition, cybercriminals may use system flaws to steal private information or sabotage grid operations. Social engineering is another tactic they might use to get victims to provide sensitive information or even download malicious software [9].

5. Insider threats

SGs face a serious security risk from insiders. An insider threat is a person who has legitimate access to the system but uses that access to compromise its safety or functionality. Anybody with access to the system qualifies, whether they are employees, contractors, vendors, or anyone else. Insider attacks provide a serious challenge to the safety and reliability of SGs. A dissatisfied employee, for instance, may undermine the grid's operations, leak important data, or steal intellectual property. A similar scenario might occur when an employee makes an inadvertent error or falls for a phishing email and clicks on the attached link [10].

6. Advanced persistent threats

Another major issue with SG security is Advanced Persistent Threat (APT)s. APTs refer to sophisticated and targeted cyberattacks that are designed to infiltrate a system and remain undetected for an extended period. Nation-states, organized criminal groups, and other APT actors often target sensitive data, disrupt operations, and seize control of crucial infrastructure when launching an APT. In the context of SGs, APTs can significantly compromise the reliability and safety of the infrastructure. If an APT attack is successful, it can compromise the entire grid, causing outages and other problems. Moreover, APT assaults can be difficult to identify and counteract since they are frequently created to circumvent common defenses [11].

1.2 Wide Area Measurement System

Wide Area Monitoring, Protection, and Control (WAMPAC) system is a cutting-edge technology to enhance the operation of power grids in real-time, specially when local controllers and protection relays are ineffective. The WAMS is a critical component of WAMPAC applications, since it provides them with the data collected from throughout the grid. A WAMS consists of various components, such as PMUs, PDCs, communication routers and links, and a Super PDC (SPDC). In fact, a WAMS acquires the data using the measurement system, transmits it through the communication network, and processes it in the control center before sending it to control and protection applications [12, 13].

1.2.1 WAMS Components

1. **Phasor Measurement Units**

   Over the last decade, PMUs have become increasingly common in the SG transmission system, as they provide an excellent way to measure the grid's performance and improve its operation. PMUs measure voltage, current, and frequency at specific points in the grid. PMUs sample measurements hundreds of times per second and use this data to calculate phasor values, which are complex numbers that represent the magnitude and phase angle of the waveforms of voltage or current at a specific point in time. PMUs also include upgraded relay and digital fault recorders, which capture data during events such as equipment failure or generator tripping [14]. The term ”synchrophasor” refers to a phasor that has been estimated at a specific instant, known as the time tag of the synchrophasor. To obtain simultaneous measurements of phasors across a wide area of the power system, it is necessary to synchronize the time tags so that all phasor measurements that belong to the same time tag are genuinely simultaneous. Synchrophasors are essentially phasors that are synchronized to an accurate time source. PMUs are synchronized to Coordinated Universal Time (UTC), which is an internationally recognized time standard. The UTC time can be obtained through the Global Positioning System (GPS), which was created by the U.S. Department of Defense to make navigation easier and broadcast precise time and location information [15].

2. Communication Infrastructure

Phasor data and other information are transmitted from the PMUs to the control center and vice versa via a communication architecture in WAMSs. In WAMS, a mix of wired and wireless networks forms the backbone of the system's communication architecture. The PMUs and the control center depend on the communication infrastructure to reliably and promptly send and receive phasor data and other information. The communication infrastructure must be able to convey the large amounts of data produced by WAMS without any interruptions or delays [16, 14].

In addition to a communication medium, PMUs require a communication protocol to transmit synchrophasor measurements to their intended destination, such as a PDC. The capacity and latency of the communication channel are essential performance-related features. A variety of protocols have been proposed and continue to evolve, including BPA/PDCstream, IEEE Std 1344-1995 (which is discouraged), IEEE Std C37.118-2005, IEEE Std C37.118.2-2011, and IEC 61850-90-5 protocols. The two most widely used standards at present are the IEEE Std C37.118.2-2011 and the IEC 61850-90-5 protocols [16]. Safe and reliable operation of the electricity system also depends on the security of the communication infrastructure. Cyber-attacks and other risks that could interrupt the power system's operation must be prevented from damaging the communication infrastructure.

3. Phasor Data Concentrators (PDCs)

A PDC is a node in a WAMS that processes synchrophasor data from various PMUs and outputs it as a single stream to higher-level PDCs or applications. The PDC groups measurements from different PMUs that have the same timestamp into a buffer that is time-stamped. Once the buffer is full, the PDC sends the measurements to other PDCs and/or Synchrophasor applications. However, communication delays may occur due to intentional cyber attacks or unintentional communication failures, which could cause the PDC to wait for delayed measurements before forwarding them. This waiting time may violate the real-time requirements of some applications. To address this, a modification has been made that includes a timer for each time-stamped buffer. For example, in one type of waiting time, the timer starts when the first measurement with a new timestamp arrives at the PDC, and the PDC assigns a new buffer to this measurement and starts the timer. When the timer goes off, the PDC forwards the received measurements without waiting for all the measurements to arrive [17].

4. Control Center

Phasor data and other information from PMUs are gathered, processed, and analyzed at the control center of a WAMS. As it is responsible for monitoring and controlling the power system in real-time, the control center plays a crucial role in its smooth operation. Servers, interfaces for communication, databases, and programs to process, analyze, and display data are all common hardware and software components found in the control center. It should be noted that, to effectively monitor and control the power system, the control center must be secure and dependable as well as be able to handle substantial amounts of data in real-time [18].

1.2.2 Taxonomy of Cyber-Attacks Against WAMSs

Generally, attacks can exploit the vulnerabilities of WAMSs in SGs for modification, interception, or interruption of data. These vulnerabilities are mainly due to the lack of physical security, inadequate authentication, improper data protection, insufficient access control, weak programming practices, and insufficient audit mechanisms [19]. The following subsections enumerate the major families of attacks against WAMSs in SGs.

1. **Physical attacks**

   WAMSs devices and nodes are subject to physical damages, such as storage removal, firmware manipulation, tampering attacks, or information extraction using open communication ports [20]. WAMS devices are often able to communicate and change settings through communication systems. An attacker with access to the input/output ports of a WAMS object can change the parameters of devices and cause unwanted operations. Moreover, using these ports, cyber-attacks can take control of devices, manipulate their firmware, and inject codes that cause them to act maliciously or even to be destroyed [21]. The change of firmware might also include a downgrade to previous versions, where known vulnerabilities exist. In such a condition, an adversary can benefit from the known vulnerabilities and take the control of devices. For instance, attackers can remove the storage of a device to extract its data and also learn about the connections of devices in the network to plan for the next stages of an attack, or gather information about other devices that communicate with the targeted device.

2. **Firmware modification attacks**

   With physical access to a device, an attacker can replace the default firmware of the device with a malicious one [22]. This intrusion gives attackers the full control of the device, if they are present physically close to it or remotely through the communication system. In the latter case, the attack can be categorized as a threat to the network layer.

3. **Device capture/node replication attacks**

   An attacker can perform a device capture/node replication attack, in which a malicious node is added to an existing network by adapting the ID number of a legitimate node in the system [23]. With the malicious node camouflaged, the attacker can perform malicious activities, such as rerouting or dumping packets. Hence, this type of attack can compromise the functionality of the entire WAMS [24]. Due to the lack of sufficient auditing, this type of attack would not be identified easily and the operators will not notice that a legitimate node has been removed in the first place, since the power consumption remains almost unchanged. It should be mentioned that even though the malicious node has the identity of a benign node, there would be a slight imbalance in energy consumption, which can be detected if there is continuous audit of power consumption throughout the system.

4. **Denial of Service (DoS) attacks**

DoS attacks can occur in the form of firmware/software, physical, or network damage. In the case of firmware/software damage, the attack can be categorized as a threat against the cyber layer, whereas a loss of communication results in an attack on the network layer of WAMSs. DoS attacks negatively impact service availability, and occur by disabling the WAMS from performing its duties. It typically happens because of (i) a flood of requests over the service host, resulting in a full buffer in the ports of devices (i.e., routers, or servers); (ii) physical removal of a device; and (iii) interrupting the communication between devices when data transfer is required. DoS attacks are categorized as either temporary or permanent. Devices with low/no security update mechanisms may be vulnerable to malicious firmware updates, and can be used as a bot for sending floods of requests to the network to clog services. A destructive update can also disable nodes or result in their malfunction, possibly when the update targets specific parts of the memory [25].

5. **Node jamming attacks**

This attack happens when an adversary obscures network connection by interfering signals, such as jamming radio frequency signals. This type of attack disrupts the availability of WAMSs since target nodes and devices can no longer be reached or controlled [26]. Additionally, node jamming attacks make time-critical data unavailable [27]. This type of attack can be also performed to disrupt the communication system by decreasing the Signal-to-Interference-plus-Noise ratio, which is often greater than one in normal situations. To perform such an attack, the adversary must have knowledge about the frequency and the modulation technique used by the target device.

6. **False Data Injection Attack (FDIA)s**

   Compromising the integrity of data by deliberate injection of false information is categorized as an FDIA. Generally speaking, in an FDIA, the data that is gathered by WAMS devices are manipulated to portray a fake condition in the underlying system or hide an event. In this attack, an adversary can also take advantage of the limited error rate tolerance of the system, and gradually raise the effect of false data such that the attack remains unnoticed. FDIAs in cyber-controlled networks have a significant effect on the system's performance, and can result in a system failure [28]. In FDIAs, even a small portion of false data can disrupt the entire WAMS. Thus, adversaries can optimize their attacks to reach the intended goal with the minimum adversarial efforts, so keeping the attack stealthy [27].

7. **Eavesdropping**

   In this type of attack, secret information is collected from communication nodes and devices. Corrupted devices in a WAMS, including compromised nodes, may leak the systems' traffic and expose confidential information [29]. Additionally, network eavesdropping—which is often referred to as network snooping or sniffing—occurs when attackers exploit insecure or vulnerable networks to access the data transmitted between two devices. This attack is among the most common ones in wireless communication.

8. Side-channel attacks

This type of attack aims to extract private information, such as encryption keys, by recording and analyzing the Side-channel activities of WAMS devices, such as timing, power consumption, and electromagnetic radiations [30]. Secret keys, for example, can be retrieved by the statistical analysis of the timing or power consumption of cryptographic algorithm executions, or the consequences of incorrect executions. The data protected in encrypted packets can be exposed by analyzing their length and processing time. A side-channel attack is fatal when the information is extracted while a system is operating. For instance, PMU communication infrastructure is vulnerable to timing side-channel attacks, in which the Hash-based Message Authentication Code (H-MAC) algorithm can be compromised by monitoring its execution time. This attack can model some security features of the stored key, e.g., its length and processing time, to decrypt the data [31].

9. Dictionary/brute-force attacks

A dictionary attack is a brute-force technique, in which attackers bombard a device/software with a set of known credentials to guess passwords. This attack is possible when authentication mechanisms are weak, and becomes easier when factory-set credentials are still in place and not updated [22]. Therefore, not updating the users' credentials [32] and utilizing weak privacy policies [24] can enable an adversary to gain high-level access to the system and control it after performing a dictionary attack. Additionally, this attack is effective when log in attempts and user credentials are not logged, or when there are devices with the same credentials.

10. Code injection

Similar to poor/malicious updates for the physical layer of WAMSs, malign updates to applications and servers may trigger security problems, such as data leakage, data loss, and unwanted control. It is worth mentioning that this attack can also target the physical layer when the adversary physically inserts some malicious codes into a WAMS device. This can happen, for instance, by attaching a malicious gadget to the target node and, on occasion, rewriting the target's operating system. Structured Query Language (SQL) injection is a type of code injection attack to acquire administrator access to databases by exploiting vulnerabilities in the victim's network infrastructure.

11. Attacks using viruses and malware

Viruses and Worms can be injected into WAMS applications using, for instance, backdoor methods, which essentially bypass the main authorization system, embedded for developers or maintenance intentions. Primarily, default passwords and out-of-date interfaces lead to backdoor exposures [33]. In contrast to computer viruses, which need a host in order to thrive, computer worms are able to thrive on their own and propagate more quickly. A viruse can replicate itself and spread from one WAMS device to another. It infects each system by embedding itself in a variety of applications and running the code when a user starts utilizing the infected software. With the aid of this malicious application, the adversary may steal information, create botnets, and harm the host machine. A worm, however, spreads over a network by looking for a vulnerable operating system. It operates on the system to cause damage to their host networks by, for instance, overloading web servers and occupying the bandwidth [34].

12. Reverse Engineering

Attackers can gain sensitive information about a system by reverse engineering its source codes. Using this strategy, attackers can identify sensitive information left by software programmers, such as hard-coded credentials and defects, and exploit it to launch attacks. Extracted information can be used to plan future assaults against the devices or to develop and employ malicious malware for them [35].

13. Man in the Middle attack

The communication between two victim WAMS devices may be intercepted by a third agent or device that privately hands over messages between the victims without letting them know they are actually conversing with the agent. This way the agent can either eavesdrop on the conversation or inject malicious information [36]. This type of intrusion may occur mostly when there is no or a poor encryption mechanism in place [20].

14. **Spoofing**

Spoofing occurs when an attacker succeeds to pretend itself as a legitimate source and gains control over a data stream, such as GPS and network time protocol [37]. This attack is carried out by disguising the attacker's identity and pretending as a trusted source instead. This type of attack often leads to data leakage, and can be leveraged to design more sophisticated attacks.

1.2.3 Cyber-attack propagation in WAMS: Literature Review

The pronounced role of WAMPAC systems in improving the reliability and operation of power grids has persuaded operators to install more PMUs across their grids. For instance, the U.S. department of energy plans to add thousands of PMUs to the power grid over the next few years as a part of the SG initiative efforts [38]. The increased number of PMUs and the larger scale of WAMSs, however, result in a larger attack surface, which can make it easier for attackers to find vulnerabilities and gain unauthorized access to critical data and applications. Thus, WAMSs—which were initially developed to enhance the grid operation in critical situations—can now be exploited by attackers to negatively impact the integrity and stability of the grid.

In general, a WAMS can be targeted by cyber-attacks that originate from (i) the communication system, (ii) physical devices, (iii) the control center, and (iv) the time synchronization mechanism used by PMUs and PDCs. Given that the existing communication protocols for WAMSs lack effective security mechanisms [39], the measurements and communication commands can be targeted by attacks, such as Radio Frequency (RF) jamming, wireless scrambling, eavesdropping, and FDIAs [40, 41]. In addition to the communication system, the physical components of WAMSs—such as PMUs, routers, and PDCs—can be targeted physically, by supply chain attacks, or by compromising the substations where these devices are installed in [42, 43]. Moreover, an attacker can compromise the WAMS by intruding into the Local Area Network (LAN) of the control center and accessing the center's facilities [44, 45]. This vulnerability has been proven by the attacks that paralyzed the Ukrainian power system in 2015 and 2016 [46]. Finally, cyber-attacks can compromise WAMSs by targeting their time-synchronization mechanisms, including (i) space-based, such as GPS [47], and (ii) network-based, e.g., precision time protocol (PTP) [48]. For instance, GPS receiver of PMUs can be compromised by spoofing or reply attacks [49], and PTP time reference can be maliciously altered by launching deception or delay attacks on broadcasting synchronization messages [49].

Propagable cyber-attacks, such as malware, are a destructive family of intrusions capable of spreading or propagating themselves across systems (e.g., WAMSs) without requiring manual intervention [50]. For instance, BlackEnergy malware was used by hackers in December 2015 to compromise the information systems of three energy distribution companies in Ukraine, resulting in disruption of energy to consumers [51]. The vast diversity of propagable cyber-attacks and their distinct behavior in different networks make defending against these attacks a challenging task [52]. Thus, it is crucial to investigate the behavior of propagable cyber-attacks to prevent and mitigate their spread in WAMSs.

Propagation of cyber-attacks in CPSs has been extensively studied in the literature. For example, the authors of [53] examine malware injection attacks against power systems. This study also determine the best attacking time to impose the maximum damage to the system, while keeping the detection risk low. Attack propagation in the SCADA system is investigated in [54]. The researchers in this study show that not only does this family of attacks affect the physical components and authenticity of data, but also it impacts the power generation capability of SGs. Attack propagation by malware in advanced metering infrastructures is studied in [55]. This study determines the optimal time of on-site investigation and monitoring to detect malware. The authors of [56] develop an epidemic model for attack propagation in wireless sensor networks to investigate the performance of various control methods (e.g., vaccination and quarantine) on ceasing the spread of malware. Reference [57] presents an attack propagation model that considers the heterogeneity of sensor nodes in communication networks. The model also takes into account the concealment of malware and malfunctioning of sensor nodes. This reference also proposes a malware spread threshold to predict whether malware will continue to proliferate or die out. The authors of [58] develop an optimization model to prevent the spread of cyber-attacks in a computer network. This model solves a mixed integer linear programming problem and identifies which nodes should be disconnected from the network in order to maximize the number of users who can access the network resources. This optimization problem keeps the infection probabilities of connected nodes below a specific threshold. Although the above-mentioned techniques are effective for Cyber-Physical System (CPS)s for which they were originally developed, they might not be effective for WAMSs. This inefficacy is due to the inherent distinctions between WAMSs and above-mentioned CPSs (e.g., the types of components and communication protocols are different), as well as the concept of observability in power systems, which must be maintained. Thus, attack propagation in WAMSs must be modeled by considering the details of this system, and countermeasures must be developed according to the operational constraints of power grids.

To the best of the authors' knowledge, there are only two main studies in the literature that focus on attack propagation in WAMSs. In the first study, i.e., in [59], the authors develop a probabilistic approach to calculate the infection probability of healthy PMUs when one or more PMUs are infected. This paper also presents an optimization framework to minimize the propagation of attacks by disconnecting PMUs whose infection probabilities are high. Although this technique can prevent the spread of cyber-attacks, it negatively impacts the observability of the system by disconnecting uninfected PMUs. The second study, i.e., [60], investigates the relationship between communication network routing and cyber-attack propagation in WAMSs. To barricade the spread of cyber-attacks from PMUs to PDCs, the authors of this study propose to reroute the communication network and connect the PMUs and PDCs through longer communication trees. This technique, however, negatively affect the reliability of the communication network. Additionally, this study does not consider the probability of attack propagation through the communication links that are not parts of the longer communication trees. Additionally, both of these studies develop their countermeasures by assuming that the abilities of adversaries are known from the beginning of attacks, which is not realistic.

1.3 Research Objectives

Driven by the above-mentioned motivations and research gaps, this dissertation first develops a mathematical model for attack propagation in WAMSs and studies the impacts of the attacker's capability and the network operator's defense ability on the spread of a cyber-attack through the system. Additionally, an LBF is developed to estimate an attacker's capability, which is crucial for defending effectively against the attack. In addition, this thesis proves that attack propagation in WAMSs may not be stopped by only recovering the infected nodes, even if the recovery rate is 100%; yet, the propagation can be stopped if the attack surface of the recovered node is reduced as well. Afterwards, to address this problem, this thesis develops another LBF to determine the required defense strategy (i.e., the minimum recovery and hardening rates) of the operator based on the estimated capability of attackers. In addition, the thesis studies the impacts of communication network configuration on attack propagation in WAMSs, and presents an attack mitigation strategy that optimally reconfigures the communication network to minimize the infection probabilities of critical PMUs and PDCs (i.e., their contamination probabilities) and to maintain the observability of the system.

1.4 Dissertation Outline

This dissertation is divided into two main parts: the next two chapters, which concentrate on the modeling and analysis of cyber-attack propagation in WAMSs as well as estimating attacker's capabilities, and the subsequent two chapters, which focus on the defensive strategies and optimal responses to propagable attacks. The individual chapters are organized as follows:

Chapter 2 develops a model to analyze cyber-attack propagation in WAMS. Then, the impacts of the attacker's capability and the network operator's defensive ability on attack propagation are investigated in detail.

Chapter 3 proposes an LBF to estimate the attacker's capability.

Chapter 4 develops an LBF to estimate the required defensive strategy (i.e., the minimum recovery and hardening rates) based on the capability of an attacker.

Chapter 5 presents an optimization framework to reconfigure communication network reconfiguration in the presence of propagable attacks. This optimal reconfiguration plan can be obtained using solving a Linear Programming (LP) problem to mitigate the propagation of cyber-attacks and maintain the observability of the power system.

Chapter 6 concludes the dissertation, highlights its contributions, and suggests topics for future research.

Chapter 2

Modeling and Analysis of Cyber-Attacks Propagation in WAMSs

This chapter develops a model to analyze cyber-attack propagation in WAMS. Then, the impacts of the attacker's capability and the network operator's defensive ability on attack propagation are investigated in detail. Such a study can elucidate the required security measures and defensive strategies to prevent the spread of cyber-attacks in WAMSs. This chapter is organized as follows: Section 2.1 discusses attack propagation across WAMSs and Section 2.2 presents simulation results and discusses them.

2.1 Attack Propagation in WAMSs

Currently, there are two models in the literature to predict the propagation of cyber-attack [61], which are based on epidemiological and theoretical control techniques. By using models based on control theory, cyber-attacks are attempted to be found, and their propagation is curbed. On the other hand, the computer science community has widely examined epidemiological models, primarily concerned with the number of compromised nodes and their distributions[62].

This chapter proposes a modified epidemiological model to investigate attack propagation in WAMSs. In the proposed model, it is assumed that an infection already exists, and the aim is to quantify how quickly this infection might propagate. It should be noted that in order to comprehend the impacts of cyber-attack propagation in WAMSs, each type of component—i.e., PMUs, routers, PDCs, and SPDCs—is assumed to have a different propagation rate. In addition, it is also assumed that (i) the network is static, i.e., the nodes are immobile, (ii) the deployment layout is static and predetermined, and (iii) nodes are composed of devices with the capability and propensity to be infected by other compromised devices. These assumptions have been made to simplify the process of attack propagation modeling. In reality, however, networks may be dynamic, with nodes and communication links being added or removed. By assuming that networks are static and the deployment layout is static, we eliminate the need to update the network topology at each time step, simplifying attack propagation analysis. Furthermore, in practical scenarios, some nodes become fully secure after detecting an attack, preventing further infection and spread. However, for the sake of simplicity, we assume in this study that all nodes are capable of attack propagation.

2.1.1 Attack Propagation Model

This subsection explains the proposed model for analyzing the dynamics of cyber-attack propagation in WAMSs. The proposed model in this thesis—which is developed based on the traditional Susceptible-Infected-Recovered (SIR) epidemic model [63]—categorizes the total N nodes of the system into four groups, and any node can be in any group at any time. These groups are as follows:

- **Indirectly Susceptible (S):** The nodes in this group are not directly connected to an infected node. However, they are still vulnerable and can be infected.
- **Directly Susceptible (S'):** The nodes in this group are more susceptible to cyber-attacks than indirectly susceptible ones, since there is a direct communication link between directly susceptible and attacked nodes.
- **Infected (I):** These nodes are infected and can compromise other nodes as well.
- **Recovered (R):** The nodes in this state are recovered from the cyber-attack (e.g., by using anti-malware and patch management systems).

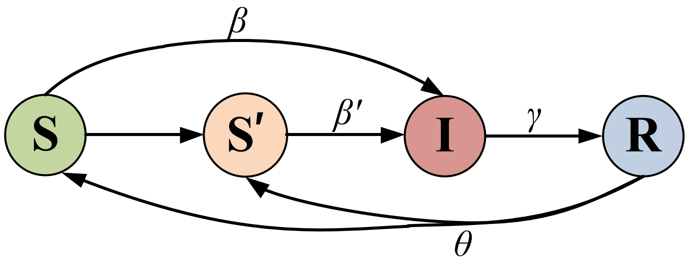


Figure 2.1: The state transition diagram of the proposed model.

Fig. 2.1 shows the state transition diagram of the proposed model. The transition of a node from one state to another happens based on the following rules:

- **Rule 1**: Cyber-attack can spread from an infected node to its directly susceptible neighbors with the rate of β, which is defined as cyber-attack propagation rate through directly susceptible nodes. This rate is different for PMUs, routers, PDCs, and SPDCs.

- **Rule 2**: When a node is infected in the network, it can infect other indirectly susceptible nodes with the rate of β', which is different for every type of component. In fact, cyber-attacks can spread from every infected node to indirectly susceptible nodes, even if they are not neighbors.

- **Rule 3**: After detecting infected nodes by an Intrusion Detection System (IDS), these nodes can be recovered with the rate of γ. This recovery action can be done by taking countermeasures, such as removing the malware or patching the vulnerabilities.

- **Rule 4**: The nodes that are recovered (e.g., only the malware is removed while the device is not patched adequately) can become vulnerable to cyber-attack again. In fact, when an infected node is recovered, it can become directly or indirectly susceptible with the rate of θ depending on its neighboring nodes.

2.1.2 Model Formulation

In this subsection, a mathematical model is developed to study the dynamics of cyber-attack propagation in accordance with the pandemic process. This model consists of two steps: (i) the initial values of $S(t), S'(t), I(t)$ and $R(t)$ at time t_0 are determined, and (ii) the parameters are updated at each time-step based on the following equations:

$$S(t+1) = S(t) - C(t+1) - D_{\mathrm{PMU}}(t+1) - D_{\mathrm{PDC}}(t+1)$$

$$-D_{\mathrm{Router}}(t+1) - D_{\mathrm{SPDC}}(t+1) \cup E'(t+1) \tag{2.1}$$

$$S'(t+1) = S'(t) \cup C(t+1) - G_{\mathrm{PMU}}(t+1) - G_{\mathrm{PDC}}(t+1)$$

$$-G_{\mathrm{Router}}(t+1) - G_{\mathrm{SPDC}}(t+1) \cup E''(t+1) \tag{2.2}$$

$$I(t+1) = I(t) \cup G_{\mathrm{PMU}}(t+1) \cup G_{\mathrm{PDC}}(t+1)$$

$$\cup G_{\mathrm{Router}}(t+1) \cup G_{\mathrm{SPDC}}(t+1) \cup D_{\mathrm{PMU}}(t+1) \cup D_{\mathrm{PDC}}(t+1)$$

$$\cup D_{\mathrm{Router}}(t+1) \cup D_{\mathrm{SPDC}}(t+1) - F(t_1) \tag{2.3}$$

$$R(t+1) = R(t) \cup F(t+1) - E(t+1) \tag{2.4}$$

$$E(t+1) = E'(t+1) \cup E''(t+1) \tag{2.5}$$

where $C(t+1)$ is the set of nodes that were indirectly susceptible at time t, but are turned to directly susceptible at time $t+1$; the set $G(t+1)$ denotes the nodes that were directly susceptible at time t, but are infected with the rate of β at time $t+1$; the set $D(t+1)$ signifies the nodes that were indirectly susceptible at time t, but are infected with the rate of β' at time t_1; the set $F(t+1)$ presents the nodes that were infected at time t, but are recovered with the rate of γ at time $t+1$; and finally, $E(t+1)$ is the set of nodes that are recovered at time t, but become either directly or indirectly susceptible at time $t+1$ with the rate of θ. The set E is divided into two subsets: (i) E', which includes the previously recovered nodes that have become indirectly susceptible, and (ii) E'', which includes the recovered nodes that have become directly susceptible.

Using the above equations, propagation of cyber-attacks in WAMSs can be analyzed by determining the transition rates (i.e., β, β', γ, and θ) as well as the initial values of each set of components. Among these parameters, β and β' are unknown and depend on the type of attacks and the abilities of attackers.

2.2 Simulation Results and Discussion

In this subsection, the performance of the proposed attack propagation model as well as and the LBF for estimating the average values of β and β' are evaluated using the IEEE 6-Bus test system shown in Fig. 2.2. PMUs are installed at buses 1, 2, 3, 4, and 6 to monitor the entire system, as suggested in [59]. These PMUs transmit their measurements to a number of clients (i.e., PDCs) using the communication network, which comprises 17 nodes in total, including five PMUs, nine routers, two PDCs, and one SPDC (Fig. 2.2). More information about this test system can be found in [60]. The main reason for selecting this case is that the IEEE 6-bus test system is relatively small, and in turn requires less computational resources and less time to simulate and analyze, making it an ideal power system for analyzing attack propagation on its WAMS. Additionally, this test case is a widely recognized and accepted benchmark within the power system research community in the area of attack propagation in WAMSs.

2.2.1 Dynamical Analysis of the Attack Propagation Model

In this subsection, a set of numerical simulations are performed to verify the dynamical behavior of the cyber-attack propagation model in a WAMS of the test system. Additionally, it studies the impacts of different rates on the average number of attacked nodes. It is assumed that PMU4 is compromised first, and this attack is detected by the IDS. Therefore, the infected PMU is disconnected. The only neighbor of PMU4, i.e., R1, is thus a directly susceptible node, and the other nodes are indirectly susceptible. The infection rates (i.e., β and β') varies from zero to one for different cyber attacks and depend on the capability of the conducted propagable attack. The higher amount of transition rates shows the higher capability of an attacker to spread the infection throughout the network. According to [59], for a specific propagable attack, it is assumed that this attack propagates to directly susceptible PMUs, routers, PDCs, and the SPDC with rates $\beta_P = 0.05$,

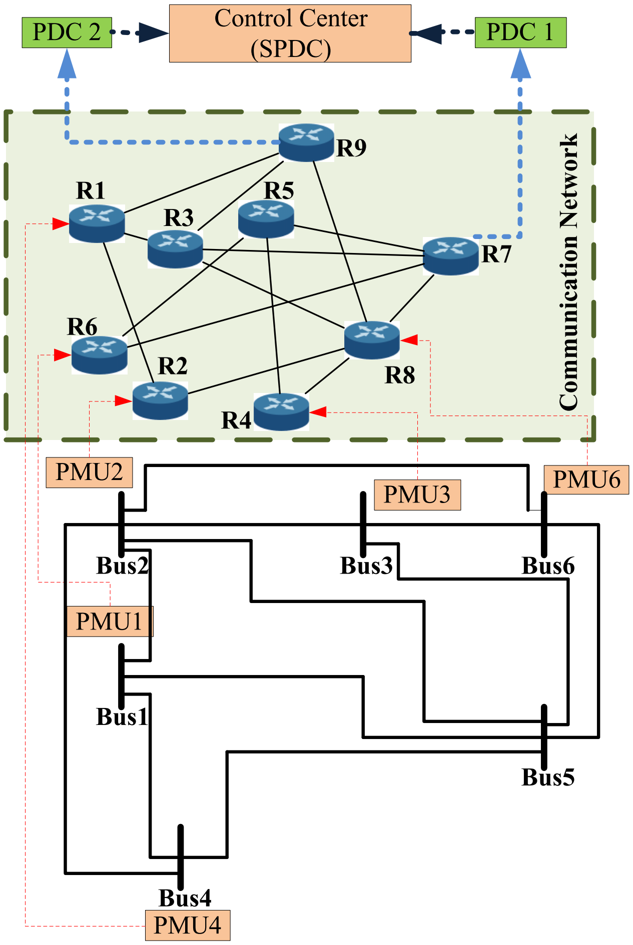
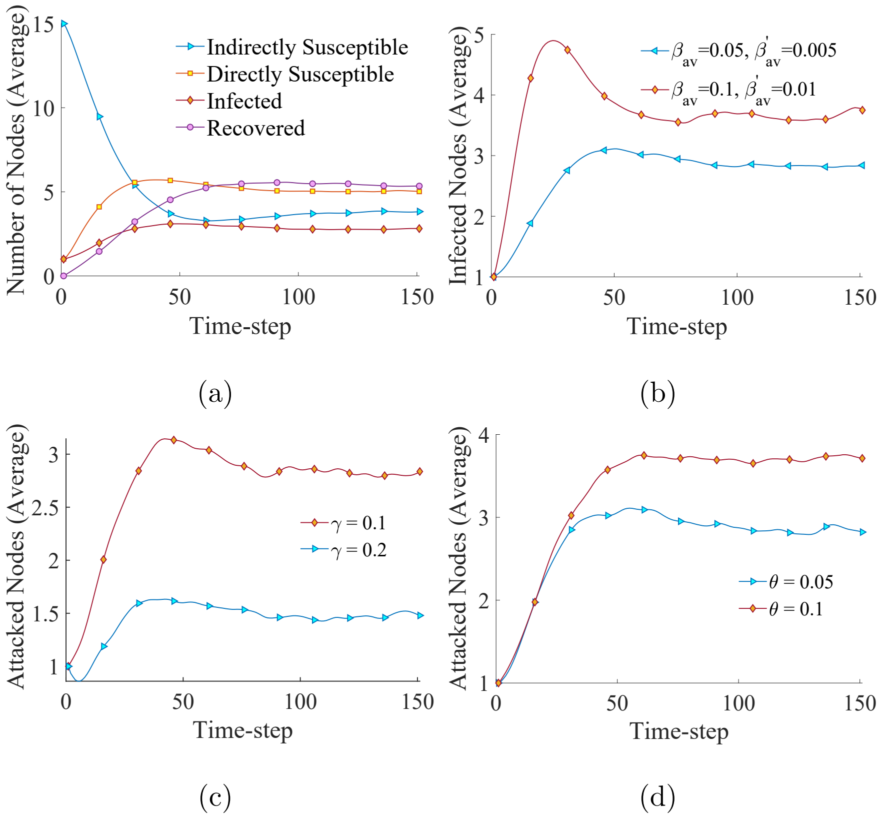


Figure 2.2: The 6-bus test system.

$\beta_R = 0.06$, $\beta_D = 0.04$, and $\beta_S = 0.0001$, respectively. The average value of these β rates is $\beta_{av} = 0.05$. Additionally, the attack spreads to other indirectly susceptible PMUs, routers, PDCs, and the SPDC with rates $\beta'_P = 0.005$, $\beta'_R = 0.006$, $\beta'_D = 0.004$, and $\beta'_S = 0.00001$, respectively. The average value of these β' rates is $\beta'_{av} = 0.005$. In this system, it is assumed that infected nodes are recovered with the rate of $\gamma = 0.1$, and recovered nodes become directly or indirectly susceptible with the rate of $\theta = 0.05$.

Fig. 2.3-(a) shows the average number of attacked, directly susceptible, and recovered nodes following the initiation of the propagable cyber-attack. Due to the stochastic nature of the problem, Monte Carlo simulation [64] is utilized to generate 1000 simulated cases, and Fig. 2.3-(a) shows their average. As this figure illustrates, the average number of attacked, directly susceptible, and recovered nodes increase, and the number of indirectly susceptible nodes decreases over time until they reach their steady-state values. Additionally, at steady-state, the average number of infected nodes is about 3. This means that the attack always exists in the network. Given that β and β' are indicators of the attacker's capability to infect nodes in a network, these two rates significantly impact the number of infected nodes. To study the impact of these two rates on the number of infected nodes, Fig. 2.3-(b) compares the average number of infected nodes when (i) the above-mentioned β and β' rates are used, and (ii) when these rates are doubled. As this figure shows, the number of infected nodes increases when β and β' rates grow. On the other hand, the rates γ and θ are indicators of the defender's capability to mitigate the propagation of attacks. Fig. 2.3-(c) shows the impact of γ on the propagation of cyber-attacks in WAMSs. In fact, this



(a)

(b)

(c)

(d)

Figure 2.3: (a) Attack propagation in the test WAMS, (b) the impacts of β and β' on attack propagation, (c) the impacts of γ on attack propagation, and (d) The impact of θ on attack propagation.

figure compares the average number of infected nodes when (i) the above-mentioned γ rate (i.e., 0.1) is used, and (ii) when γ is doubled. As this figure shows, the average number of infected nodes decreases significantly when γ increases. As a result, to mitigate the propagation of attacks in WAMSs, the operator should have sufficient knowledge about the ability of the system in cleaning infected nodes. Finally, Fig. 2.3-(d) demonstrates the effects of θ, which indicates how recovered nodes can be infected again, on the average number of infected nodes. To obtain this figure, one time the above-mentioned rates are used, and another time all the rates are kept the same except for θ, which is doubled. As this figure shows, the higher the θ the larger the number of infected nodes would be. Thus it is crucial for a network operator to decrease θ by increasing the security of recovered nodes against propagation of cyber-attacks. Otherwise, an attack can always exist in the network, or in the worse scenario, it might eventually infect all nodes.

Based on the above results, the system operator must improve γ and θ rates to stop the propagation of cyber-attacks. However, without having sufficient knowledge about an attacker's abilities, i.e., rates β and β', the effective values of γ and θ are difficult to obtain.

Chapter 3

Estimation of Attackers' Capability

This chapter elaborates on the proposed LBF to estimate the attacker's capabilities based on the received information from the IDS. In this regard, an LBF will be developed to estimate β and β' accurately. On this basis, Section 3.1 estimates the average values of β and β', and Section 3.2 presents simulation results and discusses them.

3.1 Estimating the Average values of β and β'

This section elaborates on the proposed LBF for estimating the attackers' capability in infecting the nodes of the system, which is modeled by rates β and β' in the proposed propagation model. The proposed LBF determines the average values of β and β' for all components (i.e., PMUs, PDCs, and SPDCs) based on the number of attacks that the IDS

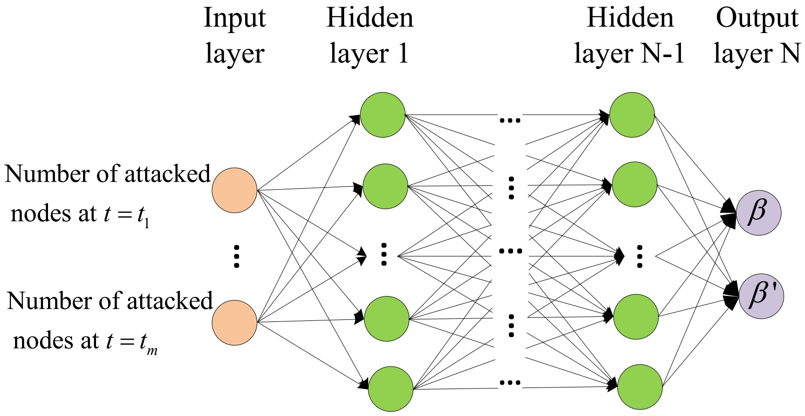

Input layer
Hidden layer 1
Hidden layer N-1
Output layer N
Number of attacked nodes at $t = t_1$
Number of attacked nodes at $t = t_m$
β
β'

Figure 3.1: The architecture of the FNN.

detects at each time interval. Among various Artificial Intelligence (AI) techniques that can be used for this purpose, Neural Network (NN)s [65] is appropriate due to its simple structure and high accuracy. Among various types of NN, this thesis leverages a FNN architecture, which is illustrated in Fig. 3.1. More information about FNN can be found in [66].

In this architecture, the input features is

$$\vec{x} = \begin{bmatrix} \mathrm{I(t_1)} & \mathrm{I(t_2)} & ... & \mathrm{I(t}_m) \end{bmatrix}^T \tag{3.1}$$

This feature vector is used to estimate the outputs of the FNN model, which are β and β'. To calculate the outputs, the input features are first fed into the model and propagate forward through the network. In this process, the input vector of each $N-1$ hidden layer is created by computing a weighted linear combination of the previous layer's outputs, as follows [67]:

$$\vec{z_1} = \vec{w_1}^T.\vec{x} + \vec{b}_1$$

$$\vec{z}_j = \vec{w}_j^T.\vec{\hat{y}}_{j-1} + \vec{b}_j \tag{3.2}$$

where $\vec{z}_1$ is the intput vector of hidden layer 1; $\vec{z}_j$ for $j \in 2, 3, ..., N-1$ is the input of j^{th} hidden layer; $\vec{w}_j$ and $\vec{b}_j$, respectively, are the weight and bias vectors of $\vec{z}_j$; and $\vec{\hat{y}}_j$ is the output of layer j, which can be obtained using

$$\vec{\hat{y}}_j = max(0, \vec{z}_j) \qquad \text{for } j \in \{1, 2, \ldots, N-1\} \tag{3.3}$$

The output of the last layer, which is the β and β' rates, are obtained using the following linear activation function

$$\vec{\hat{y}}_N = \vec{z}_N \tag{3.4}$$

The FNN model is trained for this problem using the backpropagation algorithm [68] to determine the weight and bias of each node. The objective is to minimize the difference between the expected and predicted outputs. This minimization is achieved by utilizing a loss function, which is calculated using the root mean squared error between the predicted and actual outputs:

$$J(w, b) = \sqrt{\frac{1}{C}(\vec{y} - \vec{\hat{y}}_N)^T.(\vec{y} - \vec{\hat{y}}_N)} \tag{3.5}$$

where C is the number of training cases, $\vec{y} \in R^2$ and $\vec{\hat{y}}_N \in R^2$ are the vectors of expected and predicted outputs. The Bayesian Regularization algorithm [69] is utilized in the proposed LBF to decrease the computational complexity involved in updating the model parameters.

In this study, the proposed LBF requires the number of detected attacks at the first and second time-steps to estimate β and β' (Fig. 3.1). Thus, the FNN model should be trained based on a set of attack scenarios with various β and β' values. Once the required

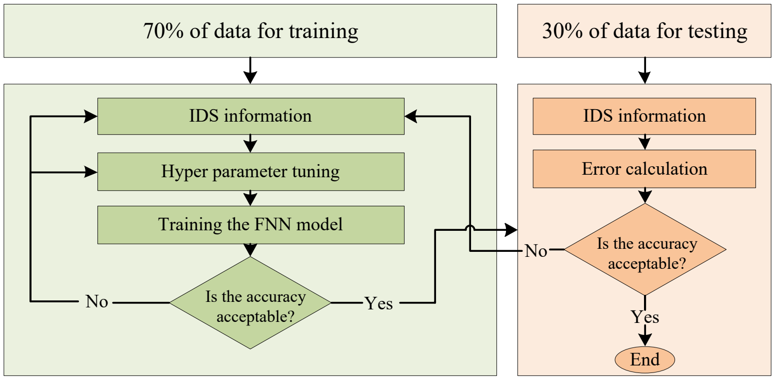


Figure 3.2: The flowchart of training and testing processes.

training data is obtained, grid search algorithm [70] is used to determine the optimal hyper parameters, i.e., the number of hidden layers and their associated nodes. Finally, the model is trained using 70% of the data, and tested based on the remaining 30%. The FNN model is ready to be implemented in the proposed LBF if the testing accuracy is satisfactory. To increase the accuracy, the number of inputs can be increased if the validation error is high. The flowchart of training and testing processes is shown in Fig. 3.2.

3.2 Simulation Results and Discussion

This part of the study assesses the effectiveness of the suggested LBF method in determining the average values of β and β'. The evaluation is conducted on the IEEE 6-Bus test system, illustrated in Figure 2.2, and PMUs are positioned at buses 1, 2, 3, 4, and 6 to supervise the entire system, as recommended in [59]. These PMUs use a communication network consisting of a total of 17 nodes, comprising five PMUs, nine routers, two PDCs,

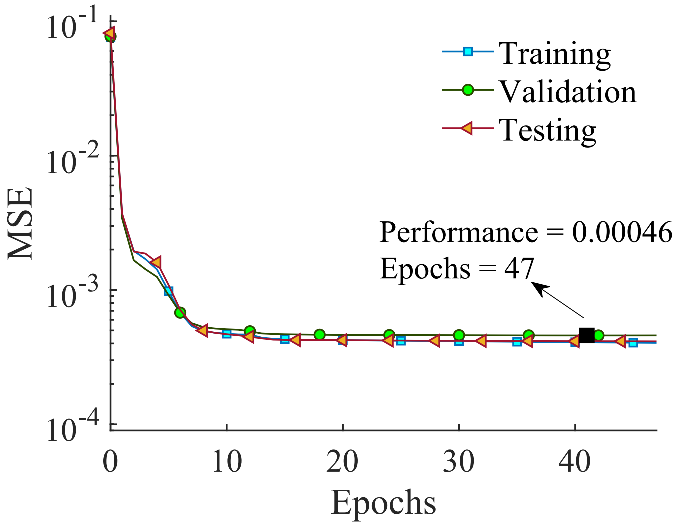
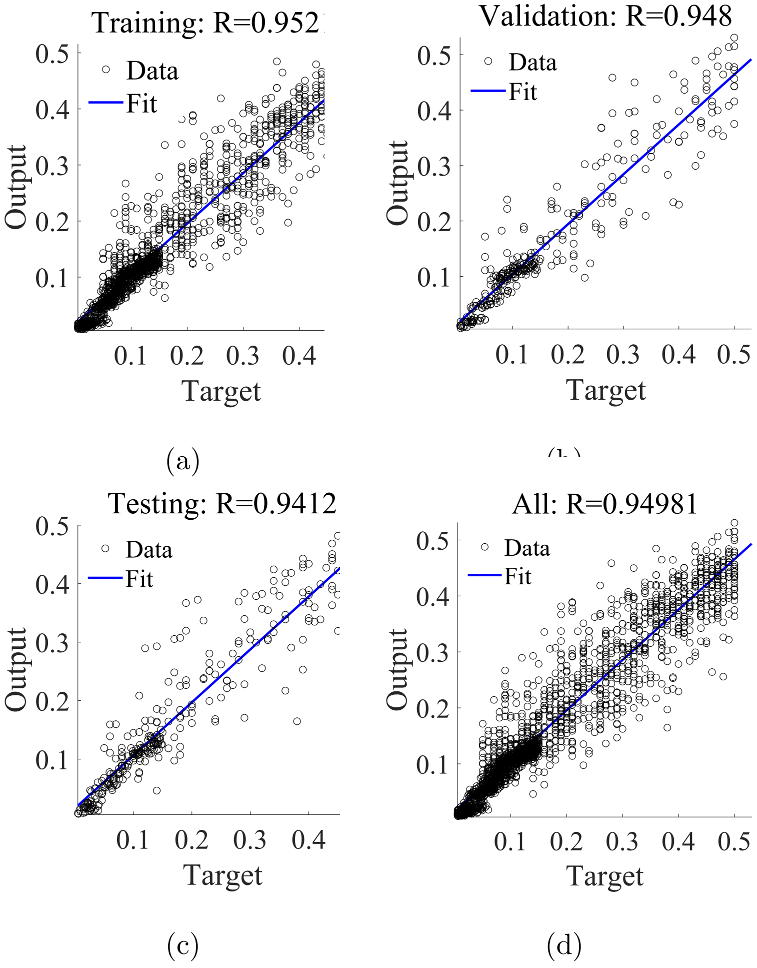


Figure 3.3: The best validation performance

and one SPDC, to transmit their measurements to a range of clients (PDCs). Additional information on this test system is available in [60].

3.2.1 Estimating Parameters β and β'

In this subsection, the performance of the proposed LBF to estimate average rates β and β' is evaluated. In this LBF, the FNN model requires a set of features that results in an accurate estimation of the rates. To this aim, the proposed LBF should be designed and tailored for each WAMS so that the accuracy of the FNN model is maximized. To this aim, the training process is carried out off-line before deploying the model online. Once the model is trained, the estimation phase must be performed based on the information received from the IDS in the first two time-steps after the initiation of the attack.



(a)

(b)

(c)

(d)

Figure 3.4: Regression analysis for (a) training, (b) validation, and (c) testing phases, as well as for (d) all data.

To generate the training database for the FNN model, 1000 random values are selected for β and β', and the number of infected nodes at every time-step is obtained for each case using the proposed attack propagation model. The model is trained using 70% of the data and tested based on the remaining 30%. An acceptable accuracy is obtained when the number of features is at least two. For this obtained number of features, the optimal architecture of the FNN model has two hidden layers with 30 nodes per layer. Fig. 3.3 illustrates the best validation performance for the training and testing phases. The best Mean Squared Error (MSE) steeply decreases and converges gradually to 0.00046 in the training and testing phases. Additionally, the regression analysis—i.e., indicating the correlation between estimated and actual data using the regression coefficient—is shown in Fig. 3.4. In fact, the higher the regression coefficient, the higher the accuracy of the FNN model would be. It should be noted that the maximum value of the regression coefficient is one. As Fig. 3.4 illustrates, the regression coefficient for the proposed LBF for all data is 0.949, which is relatively high. The regression coefficient ranges between 0 and 1, where 1 denotes perfect prediction. Additionally, values more than 0.8 indicate a relatively accurate regression model, and there is no need to use non-linear models to increase the accuracy [71]. In addition, the regression coefficient for the training, testing, and validation phases are 0.952, 0.948, and 0.941, respectively. These results demonstrate that the trained FNN model can estimate rates β and β' with acceptable accuracy for an attack in WAMSs using only the information received from the IDS in the first two time-steps. Apart from being accurate, the proposed LBF is fast as well. The average estimation time of the proposed

LBF on a computer with an Intel i7-3470 CPU running at 3.20 GHz and 16 GB of RAM is 5 ms. Therefore, the time required by the proposed method to estimate an attacker's capability is acceptable.

To further corroborate the effectiveness of the proposed framework, the number of features is increased to 20 (i.e., the number of infected nodes in the first 20 time-steps is used), and the FNN model is trained and tested based on the new features. The MSE, in this case, is 0.00038, and the regression rate increases to 0.98. Thus, in comparison to the previous case (i.e., using only the number of infected nodes in the first two time-steps), the performance is slightly improved at the expense of sacrificing time for countering the attack.

Chapter 4

Determining the Optimal Defense Strategy to Stop Propagation of cyber-attack in WAMSs

This chapter proves that attack propagation in WAMSs may not be stopped by only recovering the infected nodes, even if the recovery rate is 100%; yet, the propagation can be stopped if the attack surface of the recovered node is reduced as well. Afterwards, this chapter initially develops a LBF to estimate the required defensive strategy based on the capability of an attacker. Such a study can elucidate the required security measures and defense strategies to prevent the spread of cyber-attacks in WAMSs.

This chapter is organized as follows: Section 4.1 introduces the test system that is applied in Chapters 4 and 5. Section 4.2 shows the impacts of node hardening on attack propagation in WAMSs; Section 4.3 discusses the proposed LBF to estimate an optimal defensive strategy based on the attacker's capabilities and Section 4.4 presents simulation results and discusses them.

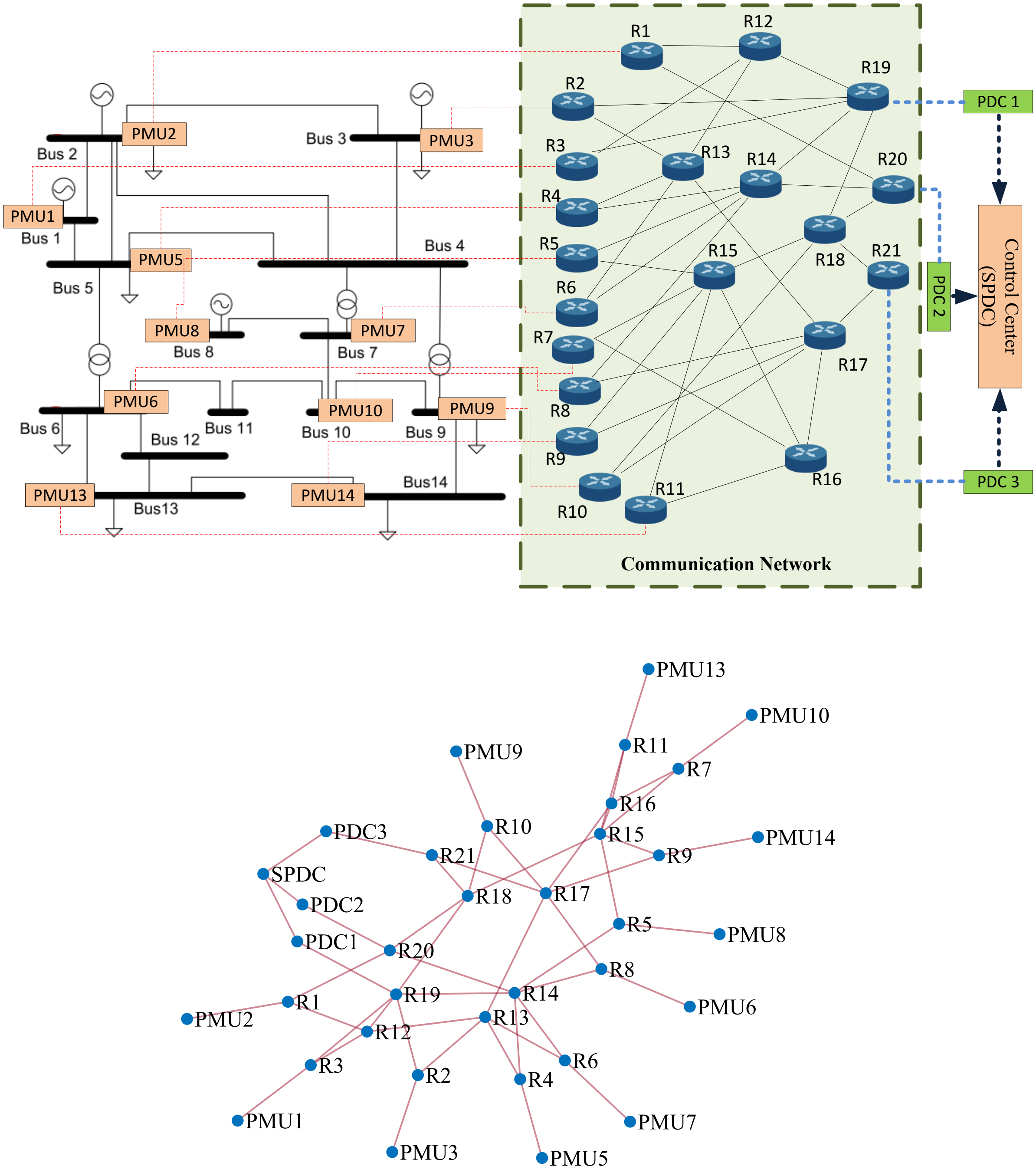
4.1 Test system

The IEEE 14-Bus test system (Fig. 4.1-(a)) is utilized in chapters 4 and 5 to investigate propagation of cyber-attacks in WAMSs. Compared to smaller test systems like the IEEE 6-bus, the IEEE 14-bus test system offers a moderate level of complexity. This enables us to study and analyze a more realistic power system and to test the scalability of the proposed mitigation methods. The parameters of the physical layer and the specifications of IEEE 14-bus test system components are provided in [72]. The physical layer consists of 14 buses, four generators, and seven loads, with a total of 20 transmission lines. This system is equipped with a WAMS, collecting real-time information throughout the system. The WAMS of this grid consists of 36 nodes comprising 11 PMUs, 21 routers, three PDCs, and one SPDC, as illustrated in Figure 4.1-(a). The equivalent graph of this test system is shown in Fig. 4.1-(b). The PMUs are installed on all buses except for Buses 4, 11, and 12, to achieve a high level of reliability and observability. To determine the communication network graph of WAMSs, all components which are connected to the network are identified, and the communication requirements of the WAMS system, such as its reliability, are specified. Afterwards, based on the requirements, the type of topology (e.g., star, mesh, ring, and hybrid) is selected. In this study, the hybrid topology is selected to develop the communication network. Additionally, it is assumed that each PMU is connected to their PDC through at least two valid communication paths. Finally, a network graph is determined to represent the network including the nodes (PMUs, PDCs, etc.) and the communication links between them [73].

4.2 Impacts of node hardening on attack propagation in WAMSs

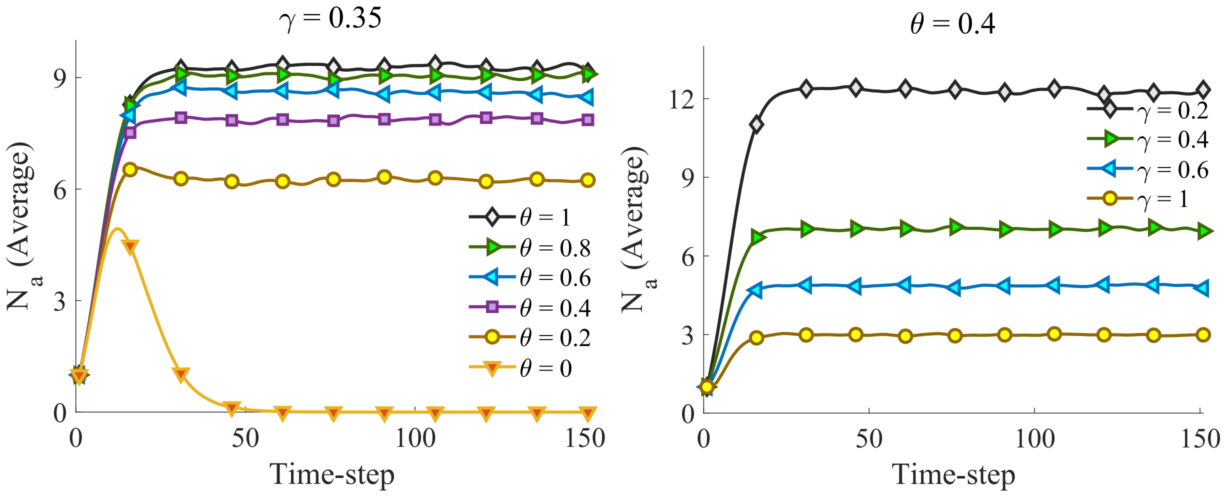
Through a case study, this subsection demonstrates that it might not be possible to stop the propagation of an attack by only recovering the infected nodes. In other words, the values of γ and θ must be selected based on attackers' capability, i.e., rates β and β'. It should be noted that the average β and β' rates can be estimated using the technique proposed in [74]. Therefore, in the rest of this chapter, it is assumed that the average values of these two rates are available, and they are assumed to be $\beta=0.15$ and $\beta'=0.015$.

In the test system, it is assumed that PMU_2 is compromised by attackers, and this intrusion is identified by the IDS. As a result, this PMU is disconnected from the network for recovery. At this moment, the only indirectly susceptible node is R_1, which is the only neighbor of PMU_2, while all other nodes are indirectly susceptible. To consider



(b)

Figure 4.1: (a) The IEEE 14-Bus Test System, and (b) its equivalent communication network graph.



(a)

(b)

Figure 4.2: (a) the impacts of θ on attack propagation, and (b) The impact of γ on attack propagation.

the stochastic nature of attack propagation in the WAMS, Monte Carlo simulation [64] is applied to generate 1000 simulated cases for each scenario.

To study the effects of θ—which indicates how recovered nodes can be infected again—on the average number of infected nodes (i.e., N_a), γ is initially set equal to 0.35, and θ is changed from 0 to 1 with steps of 0.2. Equivalently, hardening rate of recovered nodes, which equals to $\theta' = 1 - \theta$, is changed from 1 to 0 with steps of -0.2. Fig. 4.2-(a) shows the average number of infected nodes in each scenario: As the value of θ increases, fewer nodes will be hardened, and consequently the number of infected nodes will also increase. On the other hand, Fig. 4.2-(b) shows the impact of γ on propagation of cyber-attacks in WAMSs. To obtain this figure, θ is set equal to 0.4, and γ is selected as 0.2, 0.4, 0.6,

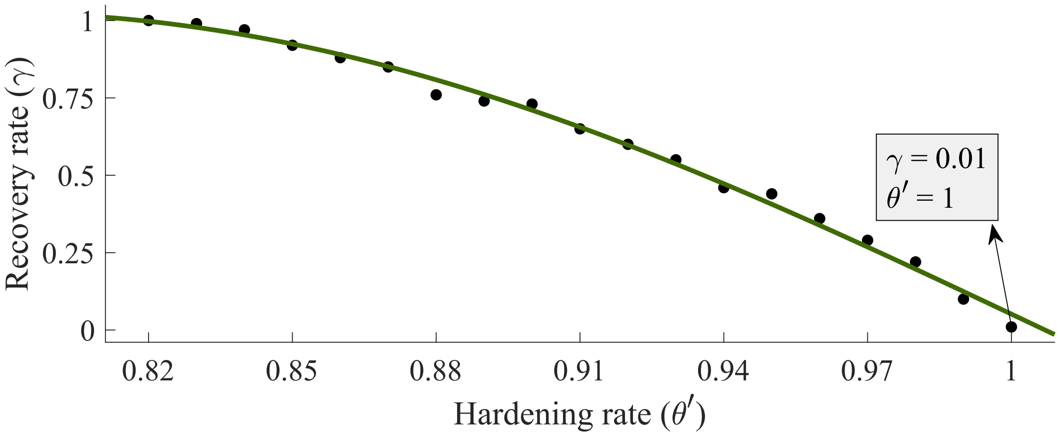


Figure 4.3: The range of recovery rate (γ) and hardening rate (θ') for stopping attack propagation.

and 1. As Fig. 4.2-(b) displays, based on these values of γ and θ, the average number of infected nodes varies between 12 and 3, indicating that attack propagation in the WAMS cannot be stopped even if all infected nodes are recovered in each time-step (i.e., $\gamma = 1$). This happens since a recovered but not hardened node is still prone to the attack, and so upon recovery it becomes either directly or indirectly susceptible, depending on the status of its neighboring nodes. Additionally, Fig. 4.3 illustrates the recovery and hardening rates required to stop the propagation of attack when β and β' are 0.15 and 0.015, respectively. As this figure shows, for hardening rates lower than 0.82, attack propagation cannot be stopped even if all infected nodes are recovered at each time-step. However, hardening rate of 1 and a very small recovery rate can eventually stop the propagation of the attack. In fact, as Fig. 4.3 shows, for any β and β', rates γ and θ should be determined together based on the operator's ability to stop the propagation of the attack.

4.3 Developing an LBF to Determine the required γ and θ Rates to Stop Attack Propagation in WAMSs

The previous section showed that when a WAMS is attacked, one way to defend against the attack is to recover the nodes that have been infected and harden them. However, since the resources and abilities of the operator are limited, it is important to divide them optimally between these two tasks to stop the propagation of the attack most effectively. In other words, an operator needs to decide how much of their available resources (such as time, money, personnel, or technical capabilities) to allocate towards recovering infected nodes and how much to allocate towards hardening the network against future attacks. Finding the optimal balance between these two tasks can help the operator stop the attack more effectively with the limited resources available. The optimal rates of recovery and hardening, however, depend on the abilities of the attackers, specifically their infection rate for directly and indirectly susceptible nodes, denoted by β and β', respectively. Considering the diverse values of these two rates, determining the optimal recovery and hardening strategy during an attack becomes a challenging task.

On this basis, this section develops an LBF to determine the optimal defense strategy that stops propagation of attacks in WAMSs. This LBF determines the optimal hardening (i.e., θ') and recovery (i.e., γ) rates based on the attacker's capabilities, i.e., β and β' rates, which can be estimated using the technique proposed in [74]. Among various learning-based techniques, FNNs [66] seems to be the most suitable one due to its simple structure

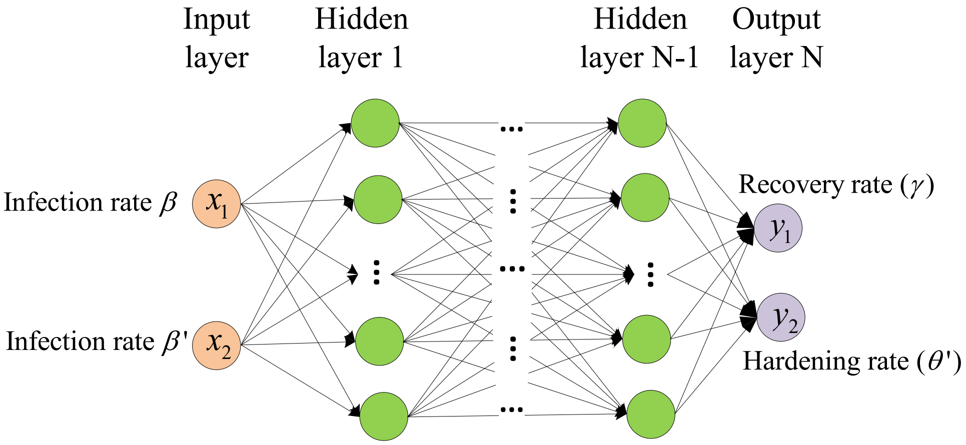


Figure 4.4: The architecture of the FNN.

and high accuracy. Fig. 4.4 shows the architecture of a FNN, which processes information in a forward direction, from the input layer through one or several hidden layers to the output layer. In this architecture, the input features is

$$\vec{x} = \begin{bmatrix} \beta & \beta' \end{bmatrix}^T \tag{4.1}$$

This feature vector is used to estimate the outputs of the FNN model, which are θ' and γ. More detail to calculate the FNN weights and outputs were discussed in chapter 3.

To prepare the training and testing data, a large number of attack scenarios with various β and β' rates are generated. Using the model presented in Chapter 2, the optimal set of θ' and γ is obtained for each scenario to stop attack propagation. Once the training data is generated, the grid search algorithm [70] is used to determine the optimal hyper-parameters, such as the number of hidden layers and their associated nodes. In this method, all possible combinations of hyper-parameters are selected, and the model is trained for each combination. The hyper-parameters that result in the highest accuracy are then selected. After training the model with a portion of the data, the remaining is used for testing. The testing accuracy and the regression rate can be used as indicators of the model's performance.

4.4 Simulation Results and Discussion

This section evaluates the proposed LBF to estimate the required γ and θ' based on the estimated values of β and β', which can be estimated using the technique elaborated in Chapter 2. To train and test the LBF model, 1000 cases of random values for β and β' are generated, and for each case the average number of infected nodes at each time-step is calculated using the attack propagation model presented in Chapter 2. Afterward, for each case, the required values of γ and θ' that prevents the propagation of the attack are obtained and saved in a database. The FNN model used in the LBF is then trained using 70% of generated cases, and the remaining 30% is used for testing and validating the model (i.e., 15% for testing and 15% for validation). To find the best trade-off between the complexity and accuracy of the model, the grid search algorithm is used, which resulted in two hidden layers with 30 nodes per layer.

Fig. 4.5 demonstrates the performance of the model during training, testing, and validation phases. In this figure, the MSE is used as a measure of accuracy, and it is observed that the MSE of validation and training phases gradually decrease and converges

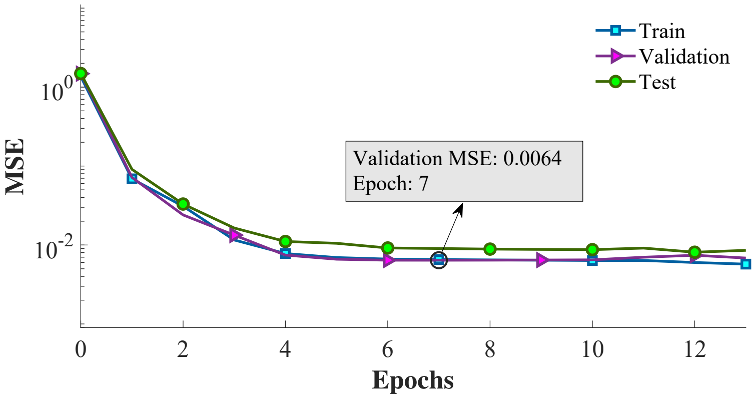


Figure 4.5: The performance of the proposed LBF in training, testing, and validation phases.

to 0.0064. Regression analysis is also performed to assess the correlation between estimated and actual data, and the regression coefficient (R) is also used to indicate the accuracy of the FNN model. The regression coefficient of the proposed LBF for all data is 0.9726 (as shown in Fig. 4.6). This coefficient for training, testing, and validation phases are 0.974, 0.964 and 0.975, respectively, indicating that the trained FNN model can estimate γ and θ' rates with a relatively high accuracy. By running the trained model on a computer with an Intel i7-3470 CPU running at 3.20 GHz and 16 GB of RAM, it is observed that the proposed LBF can estimate γ and θ' rates with an average estimation time of 9 ms.

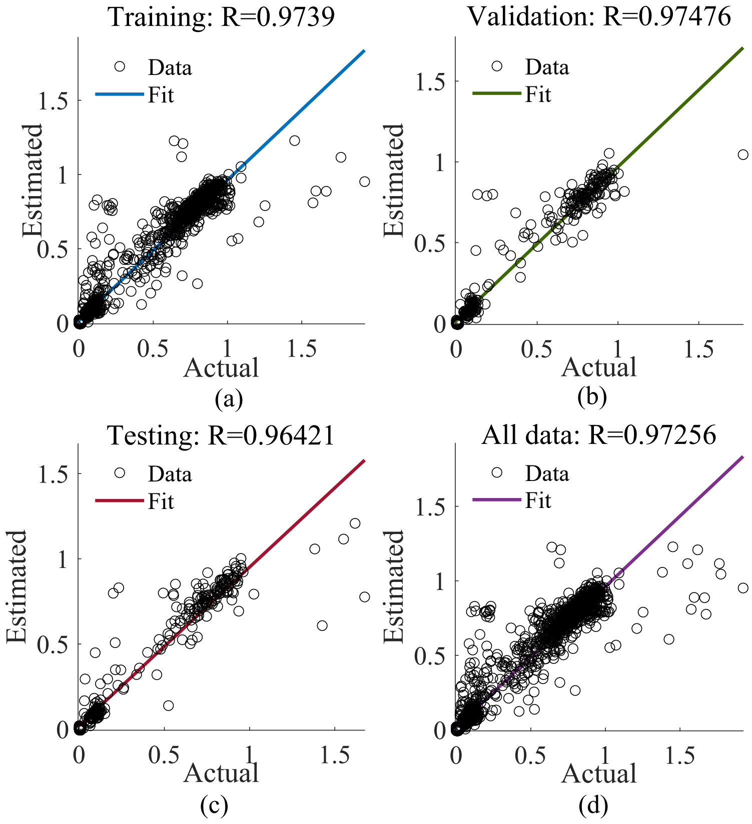


Figure 4.6: Regression analysis for (a) training, (b) validation, and (c) testing phases, as well as for (d) all data.

Chapter 5

Minimizing the Infection Probability of Critical Nodes in the Presence of Propagable Attacks

In this chapter, an optimization framework is proposed for mitigating cyber-attack propagation to WAMSs. The proposed framework minimizes the probability of infection to critical WAMS nodes by reconfiguring its communication network. Critical nodes in WAMS refer to the PMUs and PDCs that are essential for maintaining the observability of the power system. Losing even one critical node can damage the observability of the system and can make it difficult to accurately estimate the states of the power system. In this framework, the communication links that intensify the propagation of cyber-attacks throughout the network are disabled while keeping the power system observable. On this basis, first, using the attack propagation model, the infection probability of nodes in WAMSs is formulated in Section 5.1. Afterwards, Section 5.2 validates the impact of communication network on attack propagation in WAMSs as well as discusses the proposed optimization framework; Finally Section 5.3 evaluates the performance of the optimization framework on mitigating attack propagation.

5.1 Infection Probability of a Node in WAMSs

To safeguard a WAMS from propagable cyber-attacks, it is essential to identify the probability of infection for its nodes. This is because certain nodes within the system may be more critical than others and require extra protection. On this basis, this subsection uses the transition diagram presented in the previous subsection to calculate the infection probability of a node at a time-step.

Consider a typical WAMS including several nodes (e.g., PMUs, routers, PDCs and communication links). Assume a propagable attack starts at $t = 0$ and infect nodes $i \in I(t{=}0)$. One time-step later, i.e., at $t{=}\Delta t$, the infection probability of node j through compromised nodes is obtained using the following equation

$$P_j\left(a_j(\Delta t){=}1\right){=}1-(1-\beta'_j)\prod_{\substack{m\in M_{i,j}\\ i\in S'}}(1-\alpha^m_{ij})\prod_{\substack{l\in M_{k,j}\\ k\in S}}(1-\lambda^l_{kj}) \tag{5.1}$$

where $a_j(\Delta t)$ is a binary variable which equals 1 if node j is infected at time $t{=}\Delta t$, and 0 otherwise; $M_{i,j}$ is the set of all paths between nodes i and j; α_{ij}^m is the probability of attack propagation from directly susceptible node $i \in S'(t\!=\!0)$ to node j through path $m \in M_{i,j}$, and is obtained using the following equation

$$\alpha_{ij}^m = \beta_j \beta_R^{D_{ij}^m} \tag{5.2}$$

where β_j is the probability of infection for directly susceptible node j, β_R is the probability of infection for routers, and D_{ij}^m is the number of routers in path m between node i and node j [59]. Additionally, λ_{kj}^l is the probability of attack propagation from indirectly susceptible node $k \in S(t\!=\!0)$ to node j through path $l \in M_{k,j}$, and is obtained using the following equation

$$\lambda_{kj}^l = \beta'_k \beta_j \beta_R^{D_{ij}^l} \tag{5.3}$$

where β'_j is the probability of infection for indirectly susceptible node j. One time-step later, i.e., at $t\!=\!2\Delta t$, the infection probability of node j is obtained using the law of total probability, as follows:

$$\begin{aligned} P_j\left(a_j(2\Delta t)\!=\!1\right) &= P_j\left(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!1\right) P_j\left(a_j(\Delta t)\!=\!1\right) \\ &+ P_j(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!0) P_j(a_j(\Delta t)\!=\!0) \end{aligned} \tag{5.4}$$

Where $P_j\left(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!1\right)$ is the probability of node j being infected at $t = 2\Delta t$ given that it was also infected at $t\!=\!\Delta t$; and $P_j(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!0)$ is the probability of node j being infected at time $t\!=\!2\Delta t$ given that it was not infected at $t\!=\!\Delta t$. In (5.4), $P_j(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!0)$ can be written as:

$$P_j(a_j(2\Delta t)\!=\!1 | a_j(\Delta t)\!=\!0) = 1 - (1 - \beta'_j) \times$$

$$\prod_{\substack{m\in M_{i,j}\\ i\in S'}}(1-P_i(a_i(\Delta t)=1)\alpha_{ij}^m)\prod_{\substack{l\in M_{k,j}\\ k\in S}}(1-P_k(a_k(\Delta t)=1)\alpha_{kj}^l) \tag{5.5}$$

Similarly, $P_j(a_j(2\Delta t)=1|a_j(\Delta t)=1)$ can be written as:

$$P_j(a_j(2\Delta t)=1|a_j(\Delta t)=1)=\gamma\theta P_j(a_j(2\Delta t)=1|a_j(\Delta t)=0) \tag{5.6}$$

Finally, by combining equations (5.1)-(5.6), the probability of infection for node j at $t=2\Delta t$, is obtained as follows:

$$P_j\left(a_j(2\Delta t)=1\right)=\left[1-(1-\beta_j')\prod_{\substack{m\in M_{i,j}\\ i\in S'}}(1-P_i(a_i(\Delta t)=1)\alpha_{ij}^m)\times\right.$$

$$\left.\prod_{\substack{l\in M_{k,j}\\ k\in S}}(1-P_k(a_k(\Delta t)=1)\alpha_{kj}^l)\right][1-P_j(a_j(\Delta t)=1)(1-\gamma\theta)] \tag{5.7}$$

Using induction it can be shown that (5.7) can be expanded for all $n\geq 2$, and can be written in the following general form:

$$P_j\left(a_j\left((n+1)\,\Delta t\right)=1\right)=\left[1-\prod_{\substack{m\in M_{i,j}\\ i\in S'}}(1-P_i(a_i(n\Delta t)=1)\alpha_{ij}^m)\right.$$

$$\left.\times\prod_{\substack{l\in M_{k,j}\\ k\in S}}(1-P_k(a_k(n\Delta t)=1)\alpha_{kj}^l)(1-\beta_j')\right]$$

$$\times[1-P_j(a_j(n\Delta t)=1)(1-\gamma\theta)] \tag{5.8}$$

This equation quantifies the probability of infection by a propagable cyber-attack at any time-step, and will be used in the next subsections to minimize the infection probability of critical nodes.

5.2 Protecting Critical Nodes Using Communication Network Reconfiguration

This section first investigates the impacts of communication system configuration on attack propagation in WAMSs, and shows that the propagation of attacks can be slowed down and possibility of infection for critical components are reduced by reconfiguring the communication network during propagable attacks. Afterwards, it develops an optimization framework to minimize the infection probability of critical nodes in a WAMS.

5.2.1 Impacts of communication network configuration on attack propagation in WAMSs

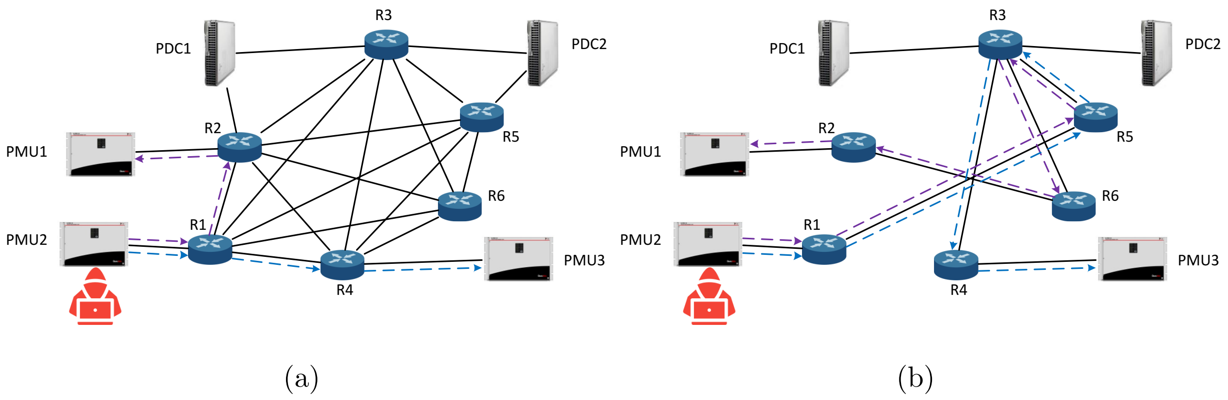
The communication network in WAMS can be reconfigured centrally using cutting edge technologies, such as software defined networking [75]. Reconfiguration in this context includes any change in the specifications of the communication network, including adding or removing devices, changing the routing of data, and increasing or decreasing the bandwidth of the network. Such changes aim to optimize the communication infrastructure to improve connectivity, increase capacity, or reduce latency.

To clearly show the impacts of communication network configuration on the infection probability of a node, this section assumes that $\beta' = 0$, meaning that only directly susceptible nodes are infected. This assumption, however, will be revisited in in the next subsection by including non-zero β' rates in the formulation of the proposed optimization framework. Using this assumption, the probability of attack propagates from node i to node j through all paths is obtained from the following equation

$$P_{ij} = 1 - \prod_{m=1}^{M_{ij}} (1 - \alpha_{ij}^{m}) \tag{5.9}$$

where (5.9) is the simplified version of (5.1). As equations (5.2) and (5.9) show, the nodal distance between two nodes increases, the probability of attack propagation from one of them to another decreases. As a result, communication network reconfiguration can be used to protect critical PMUs or PDCs by increasing their nodal distances from infected nodes.

To show this point more clearly by an example, the simple fully-connected mesh communication network shown in Fig. 5.1a is used. In this network, there are three PMUs that send their measurements to two PDCs through six routers. The probability of infection for PMUs, PDCs, and routers are assumed to be 0.05, 0.04, and 0.06, respectively [59]. In this example, it is assumed that PMU_2 is compromised, and the goal is to protect PMUs 1 and 3 from infection. To propagate the attack from PMU_2 to PMU_1, there are several paths, the shortest of which is through routers R_1 and R_2. This shortest path and



(a) (b)

Figure 5.1: (a) A fully-connected mesh communication network, (b) a partially connected mesh communication network that minimize the likelihood of PMU$_1$ and PMU$_3$ becoming infected from PMU$_2$.

other longer ones result in $P_{21} = 0.00156$. Similarly, the shortest path between PMU$_2$ and PMU$_3$ includes R$_1$ and R$_4$, resulting in $P_{23} = 0.00155$. On the other hand, Fig. 5.1b illustrates the communication network in which the distances between PMU$_1$ and PMU$_3$ are maximized relative to PMU$_2$. This configuration is selected to minimize the likelihood of PMU$_1$ and PMU$_3$ becoming infected from PMU$_2$. To propagate the attack from PMU$_2$ to PMU$_1$ in this configuration, routers R$_1$, R$_5$, R$_3$, R$_6$, and R$_2$ must be compromised first. This shortest path and other longer ones result in $P_{21} = 1.00e - 06$. The shortest path between PMU$_2$ and PMU$_3$ includes R$_1$, R$_5$ and R$_4$, resulting in $P_{23} = 9.99e - 06$. As this example showed, it is possible to benefit from communication system reconfiguration to decrease the likelihood of attack propagation from infected nodes to critical ones during a propagable cyber-attacks.

5.2.2 An optimization framework to minimize the infection probability of critical nodes in WAMSs

As shown in the previous subsection, the configuration of the communication network significantly impacts propagation of cyber-attacks in WAMSs. In this subsection, a LP optimization framework is developed to reduce the probability of attack propagation from infected nodes to critical ones during a propagable cyber-attacks. At each time-step, this optimization framework uses the status of the nodes in that time-step to determine the optimal configuration of the communication network for the next time-step. On this basis, the critical nodes must be identified first. Each operator can determine its critical nodes based on the specific needs and requirements of their system, as well as the characteristics of the components and equipment being monitored. As will be discussed in Section 5.3, this chapter identifies the critical nodes using the method proposed in [76].

To develop the LP optimization framework, and to enable connecting/disconnecting of communication links through the framework, a set binary decision variables x_{ij}^m are defined to control the connectivity of the communication links between nodes i and j. This binary variable is 1 if path m is connected between nodes i and j, and 0 otherwise. Incorporating these binary decision variables in (5.8) results in the following equation

$$P_j\left(a_j\left((n+1)\,\Delta t\right)=1\right)=\left[1-\prod_{\substack{m\in M_{i,j}\\ i\in S'}}(1-P_i(a_i(n\Delta t)=1)\alpha_{ij}^m x_{ij}^m)\right.$$

$$\times \prod_{\substack{l\in M_{k,j}\\ k\in S}} (1-P_k(a_k(n\Delta t)=1)\alpha^l_{kj}x^l_{kj})(1-\beta'_j)\Bigg]$$

$$\times[1-P_j(a_j(n\Delta t)=1)(1-\gamma\theta)] \tag{5.10}$$

Using (5.10), the optimization problem for reducing the infection probability of critical nodes through reconfiguring the communication network can be formulated as follows:

$$\min_{x^m_{ij},x^l_{kj}} \quad \sum_{j\in C} W_j P_j\left(a_j\left(\left(n+1\right)\Delta t\right)=1\right) \tag{5.11}$$

where C is the set of critical nodes, and W_j is the weighting factor of critical node j that signifies the importance of that node. In this study all critical nodes are assumed equally important, and their weighting factors are considered to be equal. This optimization problem is subjected to the following constraints:

- Connectivity constraint for directly susceptible nodes: This constraint ensures connection of directly susceptible node i and critical node j through path m, and can be formulated as shown in the following equation

$$x^m_{ij} = \prod_{u,v\in E, r\neq x} z^{ij(m)}_{uv} \qquad \forall i\in S', \forall j\in C \tag{5.12}$$

where E is the set of all nodes, and $z_{uv}^{ij(m)}$ is defined as follows:

$$z_{uv}^{ij(m)} = \begin{cases} 1 & \text{if path } m \text{ between nodes } i \text{ and } j \\ & \text{involves link (u,v)} \\ 0 & \text{if path } m \text{ between nodes } i \text{ and } j \\ & \text{does not involve link (u,v)} \end{cases} \tag{5.13}$$

• Connectivity constraint for indirectly susceptible nodes: This constraint ensures connection of indirectly susceptible node k and critical node j through path l, and is formulated as follows

$$x_{kj}^{l} = \prod_{u,v \in E, u \neq v} z_{uv}^{kj(l)} \qquad \forall k \in S, \forall j \in C \tag{5.14}$$

where $z_{uv}^{kj(l)}$ is defined similar to $z_{uv}^{kj(m)}$ in (5.13), but m must be replaced with l.

• Connectivity constraint for PMUs and PDCs: This constraint ensures that there exists at least a valid path between each PMU and its associated PDC(s). A valid path refers to a path whose communication delay is less than the waiting time of its associated PDC. This constraint can be formulated as follows:

$$\sum_{h \in M_{ij}} x_{ij}^{h} \geq 1 \tag{5.15}$$

in which, x_{ij}^{h} can be obtained using

$$x_{ij}^{h} = \prod_{u,v \in E, u \neq v} z_{uv}^{ij(h)} \qquad \forall i \in \Omega_{\text{PMU}}, \forall j \in \Omega_{\text{PDC}} \tag{5.16}$$

where, x_{ij}^{h} is a binary variable, which is one if PMU$_i$ is connected to PDC$_j$ through valid path h, and it is zero otherwise.

Since the objective function of (5.11) and its constraints nonlinear, they are reformulated to convert the the problem to an equivalent linear one. This reformulation improves the accuracy of solution and reduces the computation time. The objective function can be reformulated using the following equation:

$$\ln(1 - kx) = x \ln(1 - k) \tag{5.17}$$

which is correct when k is a constant value and x is a binary variable [59]. Using (5.17), (5.10) becomes

$$\begin{aligned} \ln\left[\frac{1}{(1-\beta'_j)} - \frac{P_j(a_j((n+1)\Delta t)=1)}{(1-P_j(a_j\,(n\Delta t)=1))(1-\gamma\theta)(1-\beta'_j)}\right] = \\ \sum_{\substack{m\in M_{i,j}\\ i\in S'}} \ln(1 - P_i(a_j\,(n\Delta t)=1)\alpha_{ij}^{m}) \times x_{ij}^{m} + \\ \sum_{\substack{l\in M_{k,j}\\ k\in S}} \ln(1 - P_k(a_j\,(n\Delta t)=1)\alpha_{kj}^{l}) \times x_{kj}^{l} \end{aligned} \tag{5.18}$$

which is a linear equation in terms of x_{ij}^{m} and x_{kj}^{l}, which are the decision variables of the proposed optimization framework. In (5.18) terms $(1 - P_j(a_j\,(n\Delta t)=1))(1-\gamma\theta)(1-\beta'_j)$ and $1/(1-\beta'_j)$ are constant values. As a result, minimizing $P_j(a_j((n+1)\Delta t) = 1)$ and maximizing (5.18) are equivalent. Additionally, constraint (5.12) can become linear by reformulating it and writing it as two linear equations [77], as follows:

$$x_{ij}^{m} \leq z_{uv}^{ij(m)} \qquad \forall i \in S', \forall j \in C, \forall (u,v) \in \Omega_{ij}^{m}, u \neq v \tag{5.19}$$

$$x_{ij}^{m} \geq \sum_{u,v \in \Omega_{ij}^{m}, u \neq v} z_{uv}^{ij(m)} - (\mathbb{N}_{ij}^{m} - 2) \qquad \forall i \in S', \forall j \in C \tag{5.20}$$

where Ω_{ij}^{m} is the set of all nodes between nodes i and j, including i and j, through path m, and $\mathbb{N}_{ij}^{m}$ is the cardinality of this set. In fact, the first constraint (i.e., (5.19)) ensures that if any of the binary variables $z_{uv}^{ij(m)}$ equals zero, then x_{ij}^{m} must also be zero. The second constraint (i.e., (5.20)) ensures that if all $z_{uv}^{ij(m)}$ are one, then x_{ij}^{m} must also be one. Similar to Constraint (5.12), Constraints (5.14) and (5.16) are also linearized using the same technique.

5.3 Performance Evaluation

This section evaluates the effectiveness of the proposed optimization framework in reconfiguring the communication network to minimize the infection rate of critical nodes. All simulations are performed using MATLAB on the IEEE 14-bus test system introduced in Chapter 4. Critical nodes are assumed to be all PDCs as well as the PMUs that influence the observability of the grid the most. To determine such PMUs, the method proposed in [76] is used, in which a node is critical if losing its measurement degrades the observability of the grid. To identify such nodes, all possible sets of essential measurements are determined first. An essential set of measurements is a group of measurements that are necessary for maintaining observability of the system, and losing any of these measurements results in a loss of observability. The essential sets of measurements for IEEE 14-bus test system are shown in Table 5.1. The nodes that are mostly repeated in these sets, i.e., PMU$_1$, PMU$_2$,

Table 5.1: Essential sets of PMUs to maintain full observability.

Essential set	PMU	Essential set	PMU	Essential set	PMU
1	**1**,**2**,**3**,**8**,**10**,13	5	**1**,**2**,**3**,7,**10**,14	9	**3**,5,7,6,9
2	**1**,**2**,**3**,**8**,**10**,14	6	**1**,**2**,**3**,7,**10**,13	10	**2**,**3**,5,**8**,6,9
3	**1**,**2**,**3**,**8**,6,9	7	**1**,**2**,5,**8**,**10**,13	11	**3**,5,7,10,13
4	**1**,**2**,**3**,7,6,9	8	**32**,,5,**8**,**10**,14	12	**32**,,5,7,**10**,14

PMU_3, PMU_8 and PMU_{10} are identified as critical, and are made bold in Table 5.1.

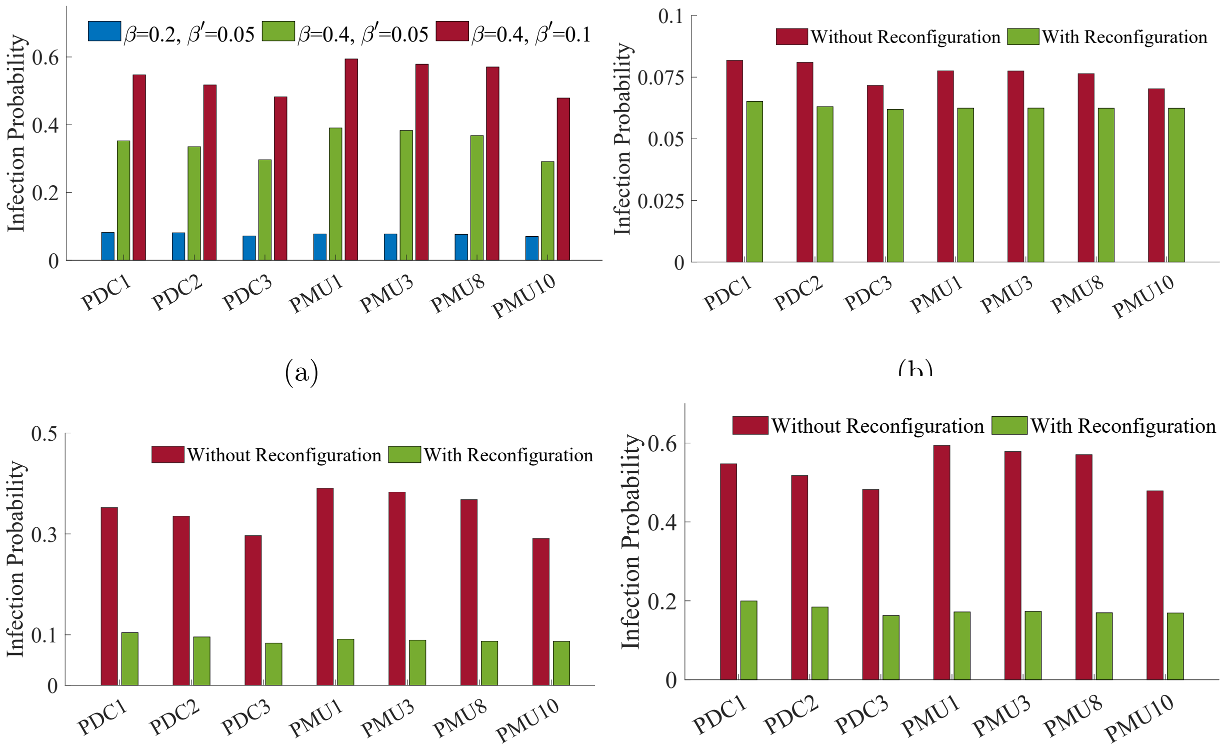
It is assumed that the system is not attacked and operates normally. At $t = 0$, the IDS detects an attack against PMU_2. Therefore, this PMU is disconnected from the network, and the remaining critical PMUs are PMU_1, PMU_3, PMU_8 and PMU_{10}. To minimize the infection probability of these PMUs, the proposed optimization framework is run. Since in these critical PMUs are essential for maintaining the observability of the grid, all weighting factors are set equal to one. All valid communication paths from PMUs to their associated PDCs are also determined and given to the framework as inputs. In this subsection, three cases are defined as follows:

- Case 1: $\beta = 0.2$, $\beta' = 0.05$
- Case 2: $\beta = 0.4$, $\beta' = 0.05$
- Case 3: $\beta = 0.4$, $\beta' = 0.1$

In all three cases, rates γ and θ are set equal to 0.1, and 0.05, respectively.

Fig. 5.2-(a) shows the infection probabilities of critical PMUs and PDCs for Cases 1-3 after targeting PMU_2 and without reconfiguring the network. As seen in the figure, the infection probabilities of critical PMUs and PDCs are low when rates β and β' are small (blue bars in 5.2-(a)). In case 2 (green bars in 5.2-(a)), the larger value of β results in a substantial rise in the infection probability of critical PMUs and PDCs. On the other hand, the larger value of β' causes a huge grow in the infection probability of critical PMUs and PDCs (red bars in 5.2-(a)). The reason is that usually, specially at the beginning of attacks, the number of indirectly susceptible nodes is larger than directly susceptible ones. Therefore, the impact of β' rate is more pronounced at the beginning of attacks.

To minimize the infection probability of critical PMUs and PDCs, the proposed optimization framework is run for all three cases and its impact on each case is investigated. Fig. 5.3 shows the reconfigured communication network for all three cases. It should be mentioned that the reason for obtaining a single graph for all three cases is that (i) the infected and critical nodes are the same in all cases, and (ii) the propagation rates in three cases are not significantly different. Comparing the obtained graph and the original one in Fig. 4.1-(b) shows that in addition to PMU_2, routers R1 and R12 are also disconnected in the reconfigured network. Router R1 is the first-hop router of PMU_2, which is a directly susceptible node with a high risk of infection. Since disconnecting this router does not violate any constraint of the problem, it is removed from the network. Additionally, Router R12 is directly connected to R1, and connects PMU_2 and R1 to several other nodes



(a)

(b)

(c)

(d)

Figure 5.2: Infection probabilities of critical nodes for (a) Cases 1-3 without reconfiguring the network, as well as for (b) Case 1, (c) Case 2, and (d) Case 3 after reconfiguring the network.

in the network. As disconnecting this router increases the nodal distance between critical nodes and PMU_2 and does not violate any of the constraints, R12 is also removed from the network. Moreover, communication links R1-PMU2, R1-R20, R1-R12, R2-R19, R3-R12, R4-R14, R5-R14, R7-R15, R9-R15, R10-R18, R11-R15, R12-R19, R12-R13, R13-R17, R16-R17, R17-R21, R18-R19, and R18-R20 are also removed from the network to minimize the connections between the nodes. However, as Fig. 5.3 shows, there is still at least one valid path between each PMU and its associated PDC(s), meaning that no PMU data will be dropped out in the communication network. As a result of this reconfiguration, the infection probabilities for critical PMUs and PDCs in all three cases reduce significantly, as shown in Figs. 5.2-(b)-(d). It is seen in these figures that the larger the values of β and β' rates, the more effective the proposed method would be in reducing the infection rates of critical nodes.

Now assume that in spite of all defensive measures, in the next time step the attack of Case 3 propagates to PMU_8 (i.e., One of critical nodes), resulting in disconnection of this device from the communication network. This attack changes the set of critical nodes to PMU_1, PMU_3, PMU_7 and PMU_{10}, as well as all PDCs. Thus, the proposed optimization framework should be run again to minimize the infection probability of critical nodes under the new situation. Red bars in Fig. 5.4-(a) show the infection probabilities for critical nodes after propagates of the attack to PMU_8. Comparing this figure with Fig. 5.2-(d) shows that the infection probabilities of Critical nodes increase after infection of PMU_8. By reconfiguring the communication network again, however, these probabilities

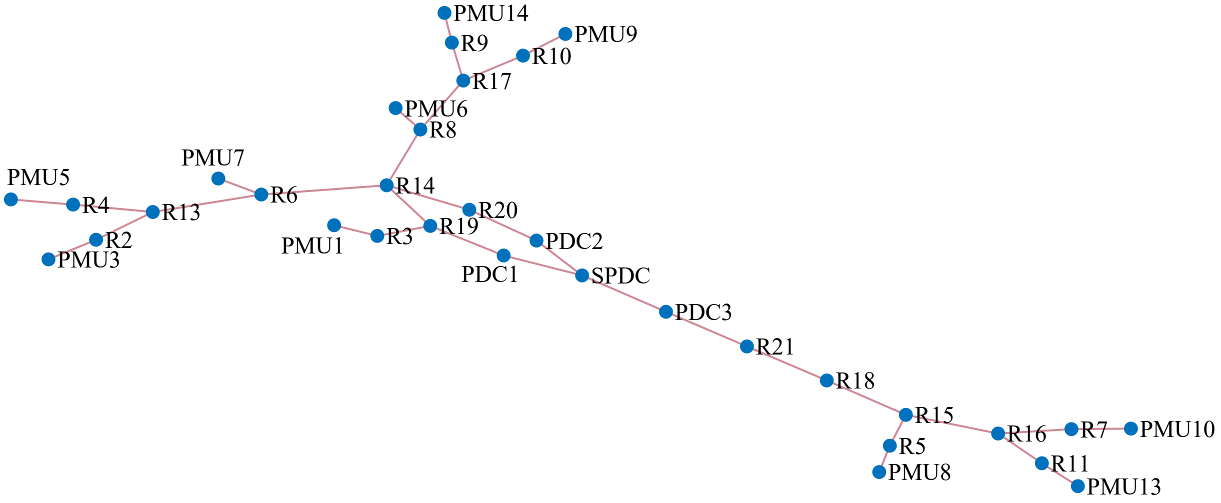


Figure 5.3: The equivalent graph of the communication system for Cases 1-3 after reconfiguration.

slightly decrease and this reduction is more pronounced for PMU_7 since it was newly added to the list of critical PMUs after disconnecting PMU_8. Fig. 5.4-(b) shows the new configuration of the communication network: compared to the configuration of the network in the previous time-step (i.e., Fig. 5.3) PMU_8 and router R5 are removed from the network. Moreover, communication links R5-PMU8, R5-R15, R6-R14, and R8-R14 are also removed from the network, while R4-R14, R9-R15, R10-R18, and R18-R20 are added to maintain the observability of the network and keeping at least one valid path between each PMU and its associated PDC(s).

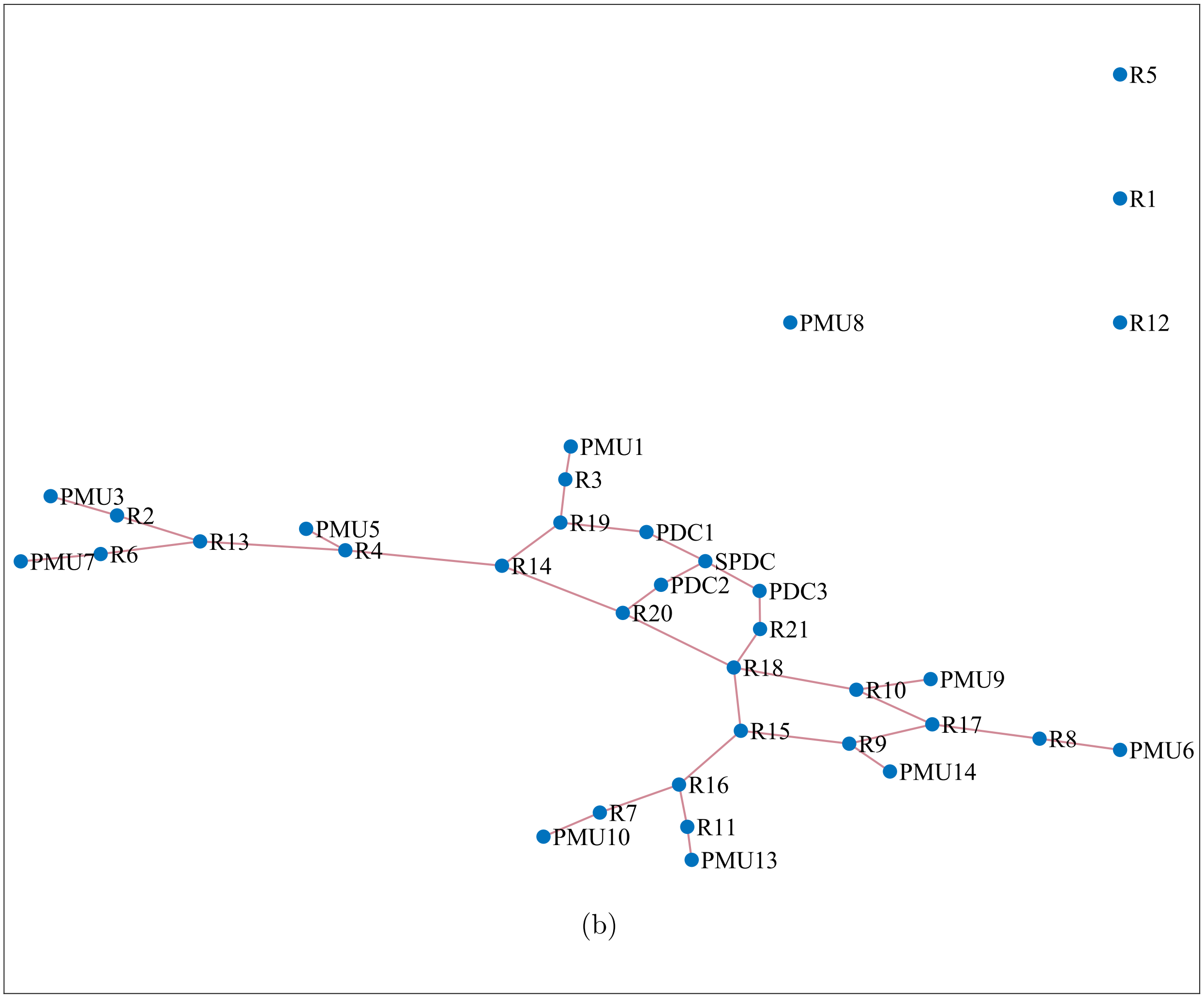


Figure 5.4: (a) Infection probabilities of critical nodes after infecting PMU_8, and (b) the equivalent graph of the communication network after reconfiguration.

Chapter 6

Conclusion and Future Works

6.1 Summery

Maintaining the observability of the power system is crucial for the efficient operation and control of the grid. However, cyber-attacks targeting the PMUs can threaten this observability. In the first part of this thesis (i.e. chapters 2, and 3), a dynamical model has been presented for analyzing the propagation of cyber-attack in WAMS. It also studied the parameters that impact the propagation of cyber-attacks in WAMSs—such as the capability of attackers and the defensive strategies of operators—and their importance in mitigating the propagation of cyber-attacks. It was also shown that depending on the attackers' and operators' abilities, an attack may always exist in the network, and it might eventually infect all nodes. The thesis also proposed an LBF for estimating the capability of attackers (i.e., rates β and β') based on the information received from the IDS in the first two time-steps after initiation of the attack. This framework, in fact, intelligently learns the relationship between the number of attacked nodes and the attacker's capability, and rationalizes it when a new attack occurs. The numerical results obtained for the 6-bus test system of this thesis showed that the proposed LBF can estimate the capability of attackers with an MSE of 0.00046 and regression rate of 0.95. Simulation results show that the data obtained in two time-steps is enough for estimating the attacker's capability with good accuracy.

The second part of this thesis (i.e. chapters 4, and 5) focused on countermeasures against propagable attacks. It has been demonstrated that only recovering the infected nodes is not enough for stopping the spread of a propagable attack in WAMSs; yet a combination of recovering and hardening strategies is required. Thus, it presented an LBF to determine the required defense strategy (i.e., recovering rate γ and hardening rate θ') based on the estimated capability for attackers (rates β and β') to stop attack propagation. Afterwards, the thesis demonstrated that the configuration of a communication network significantly impacts the propagation of cyber-attacks to its nodes. Thus it developed an LP optimization framework to change the configuration of the communication network such that the risk of infection for critical nodes is minimized. The performance of the two frameworks were evaluated on the IEEE 14-bus test system. Simulation results showed that the proposed LBF can accurately estimate the required defense strategy, and the proposed optimization framework can effectively protect the critical nodes. Although each technique is effective individually, both frameworks together are essential for effective mitigation of cyber-attack propagation in WAMSs.

The contents of each chapter can be summarized as follows:

Chapter 2 unveiled a dynamical model for analyzing the propagation of cyber-attack in WAMS. Several rates such as β and β' as well as γ and θ, which the first two are the capabilities of an attacker and the second two are the network's operator abilities were defined and their impacts on attack propagation were analyzed.

Chapter 3 proposed an LBF to estimate an attacker's ability using the information received from IDS. Several input information analyzed and the minimum number of inputs were identified.

Similar to Chapter 3, Chapter 4 presented an LBF to determine the required defensive strategy against a propagable attack using NNs (i.e., recovering rate γ and hardening rate θ') based on the estimated capability for attackers (rates β and β') to stop attack propagation. The proposed LBF could efficiently stop the propagation of attack using the information of IDS.

Chapter 5 presented an LP optimization framework reconfigure the communication network of WAMSs such that the probability of infection for critical nodes is minimized. The proposed model could guarantee the observability of the power system as well.

6.2 Contributions

The research presented in this dissertation made the following main contributions:

- **Modeling and analyzing a mathematical model for attack propagation in WAMSs:** A state-transition diagram was proposed to model and analyze cyber-attack propagation in WAMS. Then, the impacts of the attacker's capability and the network operator's defensive ability on attack propagation were investigated in detail.

- **Estimating the capability of an attacker using the information received from the IDS:** An LBF was proposed to estimate the attacker's capability using NN. This LBF is trained offline using a comprehensive database obtained from the proposed attack propagation model and is supposed to be used online.

- **Estimating the required defensive strategy against a propagable attack:** An LBF was developed to determine the required defense strategy, including recovery and hardening rates, to stop the propagation of attacks in WAMSs.

- **Mitigating the attack propagation using communication network reconfiguration:** First, the probability of infection for critical nodes in WAMSs was obtained. Afterward, an LP problem was solved to minimize the probability of infection for critical nodes using communication network reconfiguration.

6.3 Directions for Future Work

Further research on the cyber-security of WAMSs in general, and attack propagation in particular, may include the topics listed below:

- Considering communication failure and propose a mitigation method under this condition.
- Fixing the vulnerabilities of the clean components after initiating a propagable attack
- Investigating the effect of attack propagation on various WAMPAC applications such as wide area control

References

[1] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid — the new and improved power grid: A survey. *IEEE Communications Surveys Tutorials*, 14(4):944–980, 2012.

[2] Febraury NIST. Nist framework and roadmap for smart grid interoperability standards, release 2.0. *NIST Special Publication 1108R2, NIST-National Institute of Standards and Technology. URL http://www. nist. gov/smartgrid/upload/NIST_Framework_ Release_2-0_corr. pdf*, 2012.

[3] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-art. *International Journal of Electrical Power & Energy Systems*, 99:45–56, 2018.

[4] Zakaria Elmrabet, Hamid Elghazi, Tayeb Sadiki, and Hassan Elghazi. A new secure network architecture to increase security among virtual machines in cloud computing.

In *Advances in Ubiquitous Networking: Proceedings of the UNet'15 1*, pages 105–116. Springer, 2016.

[5] Himanshu Khurana, Mark Hadley, Ning Lu, and Deborah A. Frincke. Smart-grid security issues. *IEEE Security Privacy*, 8(1):81–85, 2010.

[6] Raja Waseem Anwar, Anazida Zainal, Tariq Abdullah, and Saleem Iqbal. Security threats and challenges to iot and its applications: a review. In *2020 Fifth international conference on fog and mobile edge computing (FMEC)*, pages 301–305. IEEE, 2020.

[7] Maximilian Strobel, Norbert Wiedermann, and Claudia Eckert. Novel weaknesses in iec 62351 protected smart grid control systems. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 266–270. IEEE, 2016.

[8] Vasco Delgado-Gomes, Joao F Martins, Celson Lima, and Paul Nicolae Borza. Smart grid security issues. In *2015 9th International conference on compatibility and power electronics (CPE)*, pages 534–538. IEEE, 2015.

[9] Monika Skowron, Artur Janicki, and Wojciech Mazurczyk. Traffic fingerprinting attacks on internet of things using machine learning. *IEEE Access*, 8:20386–20400, 2020.

[10] Beibei Li, Rongxing Lu, Gaoxi Xiao, Haiyong Bao, and Ali A Ghorbani. Towards insider threats detection in smart grid communication systems. *IET Communications*, 13(12):1728–1736, 2019.

[11] Florian Skopik, Ivo Friedberg, and Roman Fiedler. Dealing with advanced persistent threats in smart grid ict networks. In *ISGT 2014*, pages 1–5. IEEE, 2014.

[12] Yang Wang, Wenyuan Li, and Jiping Lu. Reliability analysis of wide-area measurement system. *IEEE Transactions on Power Delivery*, 25(3):1483–1491, 2010.

[13] Yuri V Makarov, Shuai Lu, Xinxin Guo, James Gronquist, Pengwei Du, Tony B Nguyen, and JW Burns. Wide area security region final report. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2010.

[14] Kateb Reem. *On The Security of Wide Area Measurement System and Phasor Data Collection*. PhD thesis, Concordia University, 2019.

[15] US DOE Electricity Delivery. Synchrophasor technologies and their deployment in the recovery act smart grid programs, 2013.

[16] Mohammad Shahraeini, Mohammad Hossein Javidi, and Mohammad Sadegh Ghazizadeh. Comparison between communication infrastructures of centralized and decentralized wide area measurement systems. *IEEE Transactions on Smart Grid*, 2(1):206–211, 2010.

[17] Kun Zhu, Ahmad T Al-Hammouri, and Lars Nordström. To concentrate or not to concentrate: Performance analysis of ict system with data concentrations for wide-area monitoring and control systems. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–7. IEEE, 2012.

[18] Sarasij Das and Tarlochan Singh Sidhu. Application of compressive sampling in synchrophasor data communication in wams. *IEEE Transactions on Industrial Informatics*, 10(1):450–460, 2013.

[19] Mohd Rihan, Mukhtar Ahmad, M Salim Beg, et al. Vulnerability analysis of wide area measurement system in the smart grid. *Smart Grid and Renewable Energy*, 4(06):1, 2013.

[20] Xing Liu, Cheng Qian, William Grant Hatcher, Hansong Xu, Weixian Liao, and Wei Yu. Secure internet of things (iot)-based smart-world critical infrastructures: Survey, case study and research opportunities. *IEEE Access*, 7:79523–79544, 2019.

[21] Mohamed Abomhara and Geir M Køien. Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. *Journal of Cyber Security and Mobility*, 4:65–88, 2015.

[22] Nataliia Neshenko, Elias Bou-Harb, Jorge Crichigno, Georges Kaddoum, and Nasir Ghani. Demystifying iot security: An exhaustive survey on iot vulnerabilities and a first empirical look on internet-scale iot exploitations. *IEEE Communications Surveys & Tutorials*, 21(3):2702–2733, 2019.

[23] B. Parno, A. Perrig, and V. Gligor. Distributed detection of node replication attacks in sensor networks. In *2005 IEEE Symposium on Security and Privacy (S P'05)*, pages 49–63, 2005.

[24] John Paul Walters, Zhengqiang Liang, Weisong Shi, and Vishwamitra Chaudhary. [wireless sensor network security: A survey. *Security in Distributed, Grid, and Pervasive Computing*, 01 2006.

[25] Elisa Bertino and Nayeem Islam. Botnets and internet of things security. *Computer*, 50(2):76–79, 2017.

[26] M Umar Farooq, Muhammad Waseem, Sadia Mazhar, Anjum Khairi, and Talha Kamal. A review on internet of things (iot). *International journal of computer applications*, 113(1):1–7, 2015.

[27] Xinyu Yang, Xialei Zhang, Jie Lin, Wei Yu, Xinwen Fu, and Wei Zhao. Data integrity attacks against the distributed real-time pricing in the smart grid. In *2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC)*, pages 1–8, 2016.

[28] Luca Schenato, Bruno Sinopoli, Massimo Franceschetti, Kameshwar Poolla, and S Shankar Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163–187, 2007.

[29] Kevin Hamlen, Murat Kantarcioglu, Latifur Khan, and Bhavani Thuraisingham. Security issues for cloud computing. *International Journal of Information Security and Privacy (IJISP)*, 4(2):36–48, 2010.

[30] YongBin Zhou and DengGuo Feng. Side-channel attacks: Ten years after its publication and the impacts on cryptographic module security testing. *IACR Cryptol. ePrint Arch.*, 2005:388, 2005.

[31] Mehrnaz Sharifian Esfahani. *Security analysis of phasor measurement units in smart grid communication infrastructures*. PhD thesis, University of Nebraska-Lincoln, 2014.

[32] Waseem Iqbal, Haider Abbas, Mahmoud Daneshmand, Bilal Rauf, and Yawar Abbas Bangash. An in-depth analysis of iot security requirements, challenges, and their countermeasures via software-defined security. *IEEE Internet of Things Journal*, 7(10):10250–10276, 2020.

[33] Mark Stanislav and Tod Beardsley. Hacking iot: A case study on baby monitor exposures and vulnerabilities. *Rapid7 Report*, 2015.

[34] Prajoy Podder, M Mondal, Subrato Bharati, and Pinto Kumar Paul. Review on the security threats of internet of things. *arXiv preprint arXiv:2101.05614*, 2021.

[35] Stephan Berger, Olga Bürger, and Maximilian Röglinger. Attacks on the industrial internet of things–development of a multi-layer taxonomy. *Computers & Security*, 93:101790, 2020.

[36] Farouq Aliyu, Tarek Sheltami, and Elhadi M Shakshuki. A detection and prevention technique for man in the middle attack in fog computing. *Procedia Computer Science*, 141:24–31, 2018.

[37] Ayyoob Hamza, Hassan Habibi Gharakheili, and Vijay Sivaraman. Iot network security: Requirements, threats, and countermeasures. *arXiv preprint arXiv:2008.09339*, 2020.

[38] Seyedbehzad Nabavi, Jianhua Zhang, and Aranya Chakrabortty. Distributed optimization algorithms for wide-area oscillation monitoring in power systems using interregional pmu-pdc architectures. *IEEE Transactions on Smart Grid*, 6(5):2529–2538, 2015.

[39] Amir Ameli, Abdelrahman Ayad, Ehab F. El-Saadany, Magdy M. A. Salama, and Amr Youssef. A learning-based framework for detecting cyber-attacks against line current differential relays. *IEEE Transactions on Power Delivery*, 36(4):2274–2286, 2021.

[40] Aditya Sundararajan, Tanwir Khan, Amir Moghadasi, and Arif I Sarwat. Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies. *Journal of Modern Power Systems and Clean Energy*, 7(3):449–467, 2019.

[41] Amir Ameli and Hamed Sarjan. False data injection attacks in power systems. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2022, Accepted, Unpublished.

[42] Amir Ameli, Mohsen Ghafouri, Magdy M. A. Salama, and Ehab F. El-Saadany. An auxiliary framework to mitigate measurement inaccuracies caused by capacitive volt-

age transformers. *IEEE Transactions on Instrumentation and Measurement*, 71:1–11, 2022.

[43] Daniel P Shepard, Todd E Humphreys, and Aaron A Fansler. Evaluation of the vulnerability of phasor measurement units to gps spoofing attacks. *International Journal of Critical Infrastructure Protection*, 5(3-4):146–153, 2012.

[44] Maëlle Kabir-Querrec, Stéphane Mocanu, Jean-Marc Thiriet, and Eric Savary. A test bed dedicated to the study of vulnerabilities in iec 61850 power utility automation networks. In *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4. IEEE, 2016.

[45] Chai Jiwen and Liu Shanmei. Cyber security vulnerability assessment for smart substations. In *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pages 1368–1373. IEEE, 2016.

[46] David E Whitehead, Kevin Owens, Dennis Gammel, and Jess Smith. Ukraine cyber-induced power outage: Analysis and practical mitigation strategies. In *2017 70th Annual conference for protective relay engineers (CPRE)*, pages 1–8. IEEE, 2017.

[47] Jason Allnutt, Dhananjay Anand, Douglas Arnold, Allen Goldstein, Ya-Shian Li-Baboud, Aaron Martin, Cuong Nguyen, Robert Noseworthy, Ravi Subramaniam, and Marc Weiss. Timing challenges in the smart grid. *NIST Special Publication*, 1500:08, 2017.

[48] John C Eidson, Mike Fischer, and Joe White. IEEE-1588 standard for a precision clock synchronization protocol for networked measurement and control systems. In *Proceedings of teh 34th Annual Precise Time and Time Interval Systems and Applications Meeting*, pages 243–254, 2002.

[49] force NASPI Time Synchronization Task. Time synchronization in the electric power system. *NASPI, Richland, WA, USA, Rep. NASPI-2017-TR-001*, 2017.

[50] Mouhammd Al-Kasassbeh, Mohammad A Abbadi, and Ahmed M Al-Bustanji. Lightgbm algorithm for malware detection. In *Intelligent Computing: Proceedings of the 2020 Computing Conference, Volume 3*, pages 391–403. Springer, 2020.

[51] Shui Yu, Guofei Gu, Ahmed Barnawi, Song Guo, and Ivan Stojmenovic. Malware propagation in large-scale networks. *IEEE Transactions on Knowledge and Data Engineering*, 27(1):170–179, 2015.

[52] Zhenhua Yu, Hongxia Gao, Dan Wang, Abeer Ali Alnuaim, Muhammad Firdausi, and Almetwally M Mostafa. Sei2rs malware propagation model considering two infection rates in cyber–physical systems. *Physica A: Statistical Mechanics and its Applications*, 597:127207, 2022.

[53] Sheng Xu, Yongxiang Xia, and Hui-Liang Shen. Analysis of malware-induced cyber attacks in cyber-physical power systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(12):3482–3486, 2020.

[54] Tafseer Akhtar, B. B. Gupta, and Shingo Yamaguchi. Malware propagation effects on scada system and smart power grid. In *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–6, 2018.

[55] Yonghe Guo, Chee-Wooi Ten, Shiyan Hu, and Wayne W. Weaver. Preventive maintenance for advanced metering infrastructure against malware propagation. *IEEE Transactions on Smart Grid*, 7(3):1314–1328, 2016.

[56] Senthilkumar Muthukrishnan, Sumathi Muthukumar, and Veeramani Chinnadurai. Optimal control of malware spreading model with tracing and patching in wireless sensor networks. *Wireless Personal Communications*, 117(3):2061–2083, 2021.

[57] Shigen Shen, Haiping Zhou, Sheng Feng, Jianhua Liu, and Qiying Cao. Snird: Disclosing rules of malware spread in heterogeneous wireless sensor networks. *IEEE Access*, 7:92881–92892, 2019.

[58] Mine Altunay, Sven Leyffer, Jeffrey T Linderoth, and Zhen Xie. Optimal response to attacks on the open science grid. *Computer Networks*, 55(1):61–73, 2011.

[59] Seyedamirabbas Mousavian, Jorge Valenzuela, and Jianhui Wang. A probabilistic risk mitigation model for cyber-attacks to pmu networks. *IEEE Transactions on Power Systems*, 30(1):156–165, 2014.

[60] Reem Kateb, Mosaddek Hossain Kamal Tushar, Chadi Assi, and Mourad Debbabi. Optimal tree construction model for cyber-attacks to wide area measurement systems. *IEEE Transactions on Smart Grid*, 9(1):25–34, 2016.

[61] María Teresa Signes-Pont, Antonio Cortés-Castillo, Higinio Mora-Mora, and Julian Szymanski. Modelling the malware propagation in mobile computer devices. *Computers & Security*, 79:80–93, 2018.

[62] Dilara Acarali, Muttukrishnan Rajarajan, Nikos Komninos, and Bruno Bogaz Zarpelão. Modelling the spread of botnet malware in iot-based wireless sensor networks. *Security and Communication Networks*, 2019, 2019.

[63] William O Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics—ii. the problem of endemicity. *Bulletin of mathematical biology*, 53(1-2):57–87, 1991.

[64] Nicholas Metropolis and Stanislaw Ulam. The monte carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949.

[65] Tom M Mitchell and Tom M Mitchell. *Machine learning*, volume 1. McGraw-hill New York, 1997.

[66] Daniel Svozil, Vladimir Kvasnicka, and Jiri Pospichal. Introduction to multi-layer feed-forward neural networks. *Chemometrics and intelligent laboratory systems*, 39(1):43–62, 1997.

[67] Tom Michael Mitchell et al. *Machine learning*, volume 1. McGraw-hill New York, 2007.

[68] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning internal representations by error propagation. Technical report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[69] Frank Burden and Dave Winkler. Bayesian regularization of neural networks. *Artificial neural networks: methods and applications*, pages 23–42, 2009.

[70] Irwan Bello, Barret Zoph, Vijay Vasudevan, and Quoc V Le. Neural optimizer search with reinforcement learning. In *International Conference on Machine Learning*, pages 459–468. PMLR, 2017.

[71] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Computer Science*, 7:e623, 2021.

[72] R Christie. Power systems test case archive, university of washington. *Electrical Engineering. https://www2. ee. washington. edu/research/pstca*, 2000.

[73] Xiaoping Xiong, Jiancheng Tan, and Xiangning Lin. Study on communication architecture design of wide-area measurement system. *IEEE Transactions on Power Delivery*, 28(3):1542–1547, 2013.

[74] Hamed Sarjan, Amir Ameli, and Mohsen Ghafouri. On propagation of cyber-attacks in wide-area measurement systems. In *2022 IEEE Electrical Power and Energy Conference (EPEC)*, pages 67–72. IEEE, 2022.

[75] Hui Lin, Chen Chen, Jianhui Wang, Junjian Qi, Dong Jin, Zbigniew T Kalbarczyk, and Ravishankar K Iyer. Self-healing attack-resilient pmu network for power system operation. *IEEE Transactions on Smart Grid*, 9(3):1551–1565, 2016.

[76] Ali Abur and Antonio Gomez Exposito. *Power system state estimation: theory and implementation*. CRC press, 2004.

[77] Mohammad Asghari, Amir M Fathollahi-Fard, SMJ Mirzapour Al-e hashem, and Maxim A Dulebenets. Transformation and linearization techniques in optimization: A state-of-the-art survey. *Mathematics*, 10(2):283, 2022.

[78] Bora A Akyol. Cyber security challenges in using cloud computing in the electric utility industry. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2012.

[79] Shui Yu, Guofei Gu, Ahmed Barnawi, Song Guo, and Ivan Stojmenovic. Malware propagation in large-scale networks. *IEEE Transactions on Knowledge and data engineering*, 27(1):170–179, 2014.

[80] Seungwon Shin, Guofei Gu, Narasimha Reddy, and Christopher P Lee. A large-scale empirical study of conficker. *IEEE Transactions on Information Forensics and Security*, 7(2):676–690, 2011.

[81] Hamed Sarjan, Amir Ameli, and Mohsen Ghafouri. Cyber-security of industrial internet of things in electric power systems. *IEEE Access*, 10:92390–92409, 2022.

[82] Saghar Vahidi, Mohsen Ghafouri, Minh Au, Marthe Kassouf, Arash Mohammadi, and Mourad Debbabi. Security of wide-area monitoring, protection, and control (wampac) systems of the smart grid: A survey on challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 2023.

APPENDICES

Appendix A

List of Publications

The following is a list of publications by the author during master's studies.

A.1 Peer-Reviewed Journal Articles

[1] **H. Sarjan**, A. Ameli and M. Ghafouri, "Cyber-Security of Industrial Internet of Things in Electric Power Systems," *IEEE Access*, vol. 10, pp. 92390-92409, 2022, doi: 10.1109/ACCESS.2022.3202914.

A.2 Book Chapters

[1] A. Ameli, **H. Sarjan**, "False Data Injection Attacks in Power Systems," *Wiley Encyclopedia of Electrical and Electronics Engineering*, vol. 2, pp 1-15, 2023, doi: 10.1002/047134608X.W8446.

A.3 Submitted Journal Articles

[1] **H. Sarjan**, A. Ameli and M. Ghafouri, “Mitigating Propagation of Cyber-Attacks in Wide-Area Measurement Systems”, submitted to *IEEE Transactions on Smart Grid*, Apr. 2023.

A.4 Conference Proceedings

[1] **H. Sarjan**, A. Ameli and M. Ghafouri, ”On Propagation of Cyber-Attacks in Wide-Area Measurement Systems,” *2022 IEEE Electrical Power and Energy Conference (EPEC)*, Victoria, BC, Canada, 2022, pp. 67-72, doi: 10.1109/EPEC56903.2022.10000207.