Proactive defense strategies against net load redistribution attacks in cyber-physical smart grids

by

Hang Zhang

B.S., Kansas State University, 2017

M.S., Kansas State University, 2018

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wiegers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Abstract

Recent advances in the cyber-physical smart grid (CPSG) have enabled a broad range of new devices based on information and communication technology (ICT). An open network environment in CPSG provides frequent interaction between information and physical components. However, this interaction also exposes the ICT-enabled devices to a growing threat of cyberattacks. Such threats have been alerted by recent cybersecurity incidents, and the security issues have strongly restricted the development of CPSG. Among various CPS cybersecurity incidents, cyber data attacks invade the cyber layer to destroy data integrity. Through elaborately eavesdropping on the transferred measurement data, the attacks can mislead the state estimation (SE) while keeping stealthy to conventional bad data detection (BDD). Due to the SE being the critical function of CPSG control, the cyber data attacks may cause massive economic loss, power system instability, or even cascading failures. Therefore, this dissertation focuses on the detection of stealthy data integrity attacks.

This dissertation first performs a thorough review of the state-of-the-art cyber-physical security of the smart grid. By focusing on the physical layer of the CPSG, this work provides an abstracted and unified state-space model in which cyber-physical attack and defense models can be effectively generalized. The existing cyber-physical attacks are categorized in terms of their target components. In addition, this work discusses several operational and informational defense approaches that present the current state-of-the-art in the field, including moving target defense (MTD), watermarking, and data-driven strategies. The challenges and future opportunities associated with the smart grid cyber-physical security is also discussed. Further, a real-time digital simulator, namely Typhoon HIL, is utilized to visualize the random MTD against false data injection (FDI) attacks.

Given the review section as a background, a hidden, coordinated net load redistribution attack (NLRA) in an AC distribution system is proposed. The attacker's goal is to create

violations in nodal voltage magnitude estimation. An attacker can implement the NLRA strategy by using the local information of an attack region and power flow enhanced deep learning (PFEDL) state estimators. The NLRA is modeled as an attacker's modified AC optimal power flow problem to maximize the attack impact. Case study results indicate the PFEDL-based SE can provide the attacker with accurate system states in a low observable distribution system where conventional lease square-based SE cannot converge. The stealthiness of the hidden NLRA is validated in multiple attack cases. The influence of NLRA on the distribution system is assessed, and the impact of attack regions, attack timing, and attack area size are also revealed.

Next, this dissertation highlights that current MTD strategies myopically perturb the reactance of D-FACTS lines without considering the system voltage stability. Voltage instability induced by MTDs is illustrated in a three-bus system and two more complicated systems with real-world load profiles. Further, a novel MTD framework that explicitly considers system voltage stability using continuation power flow and voltage stability indices is proposed to avoid MTD-induced voltage instability. In addition, this dissertation mathematically derives the sensitivity matrix of voltage stability index to line impedance, on which an optimization problem for maximizing voltage stability index is formulated. This framework is tested on the IEEE 14-bus and the IEEE 118-bus transmission systems, in which sophisticated attackers launch NLRAs. The simulation results show the effectiveness of the proposed framework in circumventing voltage instability while maintaining the detection effectiveness of MTD. Case studies are conducted with and without the proposed framework under different MTD planning and operational methods. The impacts of the proposed two methods on attack detection effectiveness and system economic metrics are also revealed.

Finally, this dissertation proposes utilizing smart inverters to implement a novel meter encoding scheme in distribution systems. The proposed meter encoding scheme is a software-based active detection method, which neither requires additional hardware devices nor causes system instability, compared with MTD and watermarking. By elaborately constructing the encoding vector, the proposed smart-inverter-based meter encoding can mislead the

attacker's SE while being hidden from alert attackers. In addition, by utilizing the topology of radial distribution systems, the proposed encoding scheme encodes fewer meters than current schemes when protecting the same number of buses, which decreases the encoding cost. Simulation results from the IEEE 69-bus distribution system demonstrate that the proposed meter encoding scheme can mislead the attacker's state estimation on all the downstream buses of an encoded bus without arousing the attacker's suspicion. FDI attacks constructed based on the misled estimated states are highly possible to trigger the defender's BDD alarm.

Proactive defense strategies against net load redistribution attacks in cyber-physical smart grids

by

Hang Zhang

B.S., Kansas State University, 2017

M.S., Kansas State University, 2018

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wiegers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Dr. Hongyu Wu

Copyright

Abstract

Recent advances in the cyber-physical smart grid (CPSG) have enabled a broad range of new devices based on information and communication technology (ICT). An open network environment in CPSG provides frequent interaction between information and physical components. However, this interaction also exposes the ICT-enabled devices to a growing threat of cyberattacks. Such threats have been alerted by recent cybersecurity incidents, and the security issues have strongly restricted the development of CPSG. Among various CPS cybersecurity incidents, cyber data attacks invade the cyber layer to destroy data integrity. Through elaborately eavesdropping on the transferred measurement data, the attacks can mislead the state estimation (SE) while keeping stealthy to conventional bad data detection (BDD). Due to the SE being the critical function of CPSG control, the cyber data attacks may cause massive economic loss, power system instability, or even cascading failures. Therefore, this dissertation focuses on the detection of stealthy data integrity attacks.

This dissertation first performs a thorough review of the state-of-the-art cyber-physical security of the smart grid. By focusing on the physical layer of the CPSG, this work provides an abstracted and unified state-space model in which cyber-physical attack and defense models can be effectively generalized. The existing cyber-physical attacks are categorized in terms of their target components. In addition, this work discusses several operational and informational defense approaches that present the current state-of-the-art in the field, including moving target defense (MTD), watermarking, and data-driven strategies. The challenges and future opportunities associated with the smart grid cyber-physical security is also discussed. Further, a real-time digital simulator, namely Typhoon HIL, is utilized to visualize the random MTD against false data injection (FDI) attacks.

Given the review section as a background, a hidden, coordinated net load redistribution attack (NLRA) in an AC distribution system is proposed. The attacker's goal is to create

violations in nodal voltage magnitude estimation. An attacker can implement the NLRA strategy by using the local information of an attack region and power flow enhanced deep learning (PFEDL) state estimators. The NLRA is modeled as an attacker's modified AC optimal power flow problem to maximize the attack impact. Case study results indicate the PFEDL-based SE can provide the attacker with accurate system states in a low observable distribution system where conventional lease square-based SE cannot converge. The stealthiness of the hidden NLRA is validated in multiple attack cases. The influence of NLRA on the distribution system is assessed, and the impact of attack regions, attack timing, and attack area size are also revealed.

Next, this dissertation highlights that current MTD strategies myopically perturb the reactance of D-FACTS lines without considering the system voltage stability. Voltage instability induced by MTDs is illustrated in a three-bus system and two more complicated systems with real-world load profiles. Further, a novel MTD framework that explicitly considers system voltage stability using continuation power flow and voltage stability indices is proposed to avoid MTD-induced voltage instability. In addition, this dissertation mathematically derives the sensitivity matrix of voltage stability index to line impedance, on which an optimization problem for maximizing voltage stability index is formulated. This framework is tested on the IEEE 14-bus and the IEEE 118-bus transmission systems, in which sophisticated attackers launch NLRAs. The simulation results show the effectiveness of the proposed framework in circumventing voltage instability while maintaining the detection effectiveness of MTD. Case studies are conducted with and without the proposed framework under different MTD planning and operational methods. The impacts of the proposed two methods on attack detection effectiveness and system economic metrics are also revealed.

Finally, this dissertation proposes utilizing smart inverters to implement a novel meter encoding scheme in distribution systems. The proposed meter encoding scheme is a software-based active detection method, which neither requires additional hardware devices nor causes system instability, compared with MTD and watermarking. By elaborately constructing the encoding vector, the proposed smart-inverter-based meter encoding can mislead the

attacker's SE while being hidden from alert attackers. In addition, by utilizing the topology of radial distribution systems, the proposed encoding scheme encodes fewer meters than current schemes when protecting the same number of buses, which decreases the encoding cost. Simulation results from the IEEE 69-bus distribution system demonstrate that the proposed meter encoding scheme can mislead the attacker's state estimation on all the downstream buses of an encoded bus without arousing the attacker's suspicion. FDI attacks constructed based on the misled estimated states are highly possible to trigger the defender's BDD alarm.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my gratitude and appreciation for Dr. Hongyu Wu. Your guidance, support, and encouragement have been invaluable throughout this study. It is a great honor to work with you.

I would like to extend my appreciation to my committee members, Dr. Anil Pahwa, Dr. Behrooz Mirafzal, and Dr. Ashesh Sinha, for their time and comments that improved this work.

Finally, I would like to acknowledge every Smart Energy Systems (SES) Group member, Dr. Bo Liu, Dr. Lawryn Kiboma, Xuebo Liu, Li Wang, and Dr. Ananth Palani. Many thanks for your support and the happy time we had.

This work was supported in part by the U.S. National Science Foundation under Grant No. 1929147 and No. 2146156, and in part by the U.S. Department of Energy under Award No. DE-EE0008767.

Dedication

This dissertation is dedicated to my beloved wife Shimeng Wang, who has been a constant source of love and encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life.

To my parents, Xiangduan Zhang and Shufang Zhang, who have always loved me unconditionally. Your good examples have taught me to work hard for the things that I aspire to achieve.

To my parents-in-law, Guochao Wang, Ming Ma. I will forever be grateful for your support and suggestions.

Chapter 1

Introduction

This chapter introduces the concepts studied in this dissertation. It also emphasizes the importance and necessity of this research as well as provides detailed research questions and contributions of this dissertation.

1.1 Background

Cyber-physical systems (CPSs) are smart systems that include engineered interacting networks of physical and computational components[1]. The comprehensively interconnected and integrated systems contribute new functionalities to enable technological development in critical infrastructures, such as electric power systems, water networks, transportation, home automation, and health care. A CPS encompasses complex systems of control, awareness, computing, and communication. The complexity and heterogeneity have indicated potential challenges to the security and resilience of CPSs. The interconnection of bulk physical layer components is challenging the protection against inherent physical vulnerabilities therein. On the other hand, cyber-integration, which relies on network communication and the internet of things (IoT) based devices, requires extraordinary investments in security designs and upgrades against unanticipated threats from cyberspace[2]. A cyber-physical attack is defined as a security breach in cyberspace that adversely affects the physical space of a

CPS[3]. Cyber-physical attacks compromise the confidentiality, integrity, and availability of information by coupling cyber and physical spaces in a CPS. In the past decades, several noteworthy cyber-physical attacks have been reported in the industry, facilitating synergistic efforts from industry practitioners and research communities towards a new CPS security era[4]. The first proclaimed cyber-physical attack dated back to 1982 in the Siberian wilderness, where attackers manipulated the pipeline control software, which led the valves' control to misbehave, resulting in the severe crossing of pressure limits and eventually a massive explosion[5]. In 2003, the Slammer worm invaded the control system of the David-Besse nuclear plant in Ohio through a contractor's network, which disabled the supervisory system for 5 hours[6]. In June 2010, a cyber worm dubbed "Stuxnet" struck the Iranian nuclear fuel enrichment plant by utilizing four zero-day vulnerabilities and digitally signed certificates to bypass intrusion detection. The targets were the programmable logic controllers in the supervisory control and data acquisition (SCADA) system[7]. The Stuxnet maliciously alternated the frequency of electrical current powering the centrifuges and then switched them between high and low speeds at intervals for which the machines were not designed[8]. In December 2015, a coordinated cyberattack compromised three Ukrainian electric power distribution companies. Thirty substations suffered blackout for about three hours, resulting in wide-area power outages affecting approximately 225,000 customers. BlackEnergy3 malware was used to steal the authorized users' virtual private network credentials, and a telephonic denial-of-service (DoS) attack was executed to frustrate reports of outages[9].

The smart grid landscape, arguably one of the most complex CPSs in history, is undergoing a radical transformation. Particularly, increased renewable energy resources, demand diversification, and integration of information and communication technologies (ICTs)[10]. The cyber-physical smart grid (CPSG) that has organized a universal cyberinfrastructure interwoven with the bulk physical systems is susceptible to cyber-physical attacks. A wide variety of motivations exist for launching such an attack in smart grids, ranging from economic reasons, to terrorism, to a grudge (a disgruntled employee[11]). A large body of recent work has been dedicated to addressing the cyber-physical security of smart grids, with many warnings becomes prominent[12–15] and new vulnerabilities are continuously unveiled[16]. Regarding cyber and physical security, neither of them alone can provide broad solutions without incorporating the other. In this regard, the investigations of the cyber-physical attacks and the development of effective defense strategies are still incomprehensive. Thereby, it has become paramount to keep up with the latest progress along the research frontier of smart grid security.

This dissertation focuses on enhancing the cybersecurity of power systems against data integrity attacks. The state-of-the-art research on cyber-physical attacks and defense methods is reviewed. Then, a realistic data integrity attack, namely net load redistribution attack (NLRA), is introduced in AC distribution systems. Two emerging defense mechanisms, i.e., moving target defense (MTD) and smart-inverter-based meter encoding, are proposed to detect the NLRA. The remainder of this chapter introduces the control-measurement loop of the cyber-physical smart grid, the cyber-physical attacks in smart grids, the MTD methods, and the meter encoding methods.

1.2 Introduction to Cyber-physical Attacks in Smart Grid

This section proposes a discrete-time nonlinear time-invariant system to represent a CPSG using the state-space representation. Such a high-level abstraction is a useful strategy to form the foundation and generalize a defense analysis across all attack types.

A CPSG is a monolithic system with electricity generation, transmission, and distribution sectors[17]. The physical systems are interconnected through transmission lines and substations deployed in the field. The integration and coordination of heterogeneous components require reliable capabilities in information, computation, and communication. These requirements rely on a ubiquitous cyber-infrastructure interwoven with the physical systems. Measurements and commands are constantly generated and transmitted through communi-

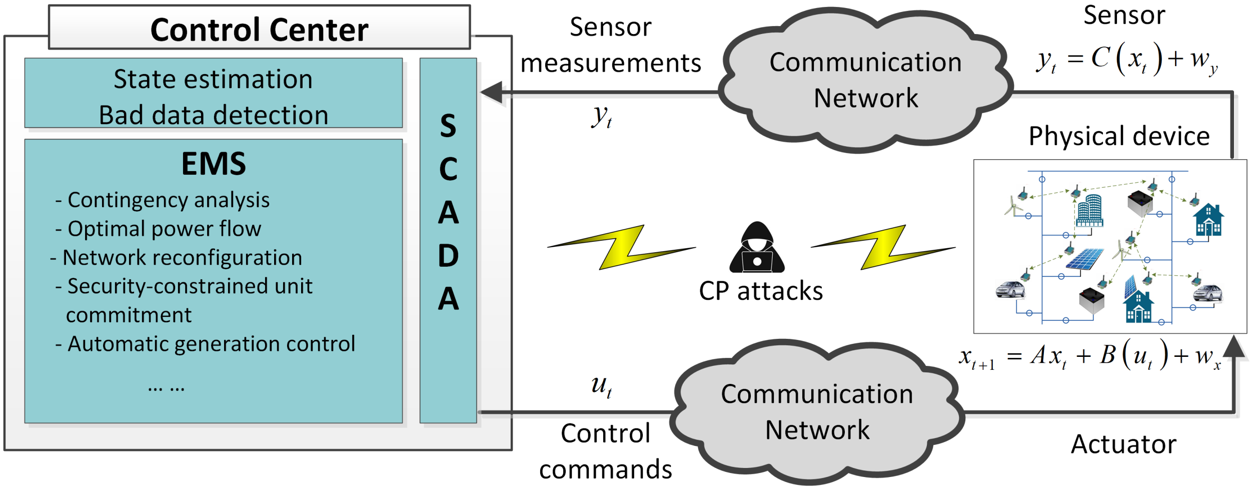


Figure 1.1: Illustration of cyber-physical attacks on smart grid. This Dissertation focuses on reviewing attacks that target either the EMS within the control center or physical devices in the field. Defense mechanisms against those attacks are also discussed.

cation channels. A CPSG consists of physical devices, actuators, sensors, communication channels, a centralized control center equipped with a state estimator, a bad data detector, and an energy management system (EMS), as shown in Fig. 1.1.

This dissertation describes the CPSG as a discrete-time nonlinear time-invariant system by using a state-space representation as follows:

$$x_{t+1} = A(x_t) + B\left(u_t\right) + w_t \tag{1.1}$$

$$y_t = C\left(x_t\right) + v_t \tag{1.2}$$

where $x_t \in \mathbb{R}^n$ and $y_t \in \mathbb{R}^m$ are system state and measurements at time interval t, respectively; m is the number of measurements; n is the number of system states (usually $m \geq n$). Typically, system measurements include nodal net injections, line power flows, line current phasors, and bus voltage phases from the emerging phasor measurement units (PMUs). System states include bus voltage magnitudes and angles. $A\left(\bullet\right)$ denotes a system state function; $B\left(\bullet\right)$ is a control function; $C\left(\bullet\right)$ is a nonlinear measurement function; $w_x \in \mathbb{R}^n$ and $w_y \in \mathbb{R}^m$ are system operating noise and measurement noise, respectively. The measurement function is reliant on the specific measurement type and involves the power system network topology and parameters, such as line impedance and transformer tap ratios. The noise is generally assumed to be Gaussian distributed with a covariance matrix $R \in \mathbb{R}^{m \times m}$. The received sensor measurement data, which are called raw data, cannot be utilized directly by the EMS and must be processed by state estimation (SE) and bad data detection (BDD).

The wide-area field sensors and communication channels are exposed to an increased level of cyber threats. As shown in Fig. 1.1, the communication networks are vulnerable to adversaries who can manipulate the control and measurement signals. For countermeasures, the National Electric Sector Cybersecurity Organization Resource (NESCOR) has conducted impact analyses and assessment of data integrity attacks against the wide-area monitoring, protection, and control (WAMPAC) systems[18], in which a dozen attack scenarios are discussed with the corresponding failure scenarios, including line trip, improper synchronous closing, and control actions that create undesirable states. For instance, the WAMPAC.2 scenario indicated that the network equipment could be leveraged to spoof WAMPAC messages[18]. A threat agent may perform a spoofing attack and inject messages into WAMPAC network equipment (router, switch, etc.). The altered messages involve measurement that goes into the WAMPAC algorithms or control commands to PMUs or phasor data concentrators (PDCs). The WAMPAC.4 scenario leverages the compromised PDC authentication to manipulate the measurement data. Such compromise may be due to a backdoor or network sniffing, which allows the malicious introduction of false measurement data. The altered data can trigger actions when none are necessary or fail to take action when needed. Meanwhile, The WAMPAC.8 scenario shows an attacker can insert malware in PMU/PDC firmware to alter measurements. When the altering action is triggered, significant effort or cost is invested in troubleshooting the systems given the lack of measurement consistency, followed by equipment replacement[18].

Analyzing the vulnerabilities of a CPSG has attracted increasing attention in the last few years. The general approach is to study specific attacks against a particular system component. A CPSG consists of information technologies (IT) and operational technologies

(OT). IT refers to the application of networks that deal with the data and the flow of digital information. In contrast, OT refers to technology that monitors and controls specific devices, such as the SCADA system. IT and OT are merging, known as IT-OT convergence, and the boundary between them has become blurry. This dissertation primarily focuses on OT attacks and defense approaches in smart grids. Figure 1.2 shows the attacks and their targets surveyed in this dissertation.

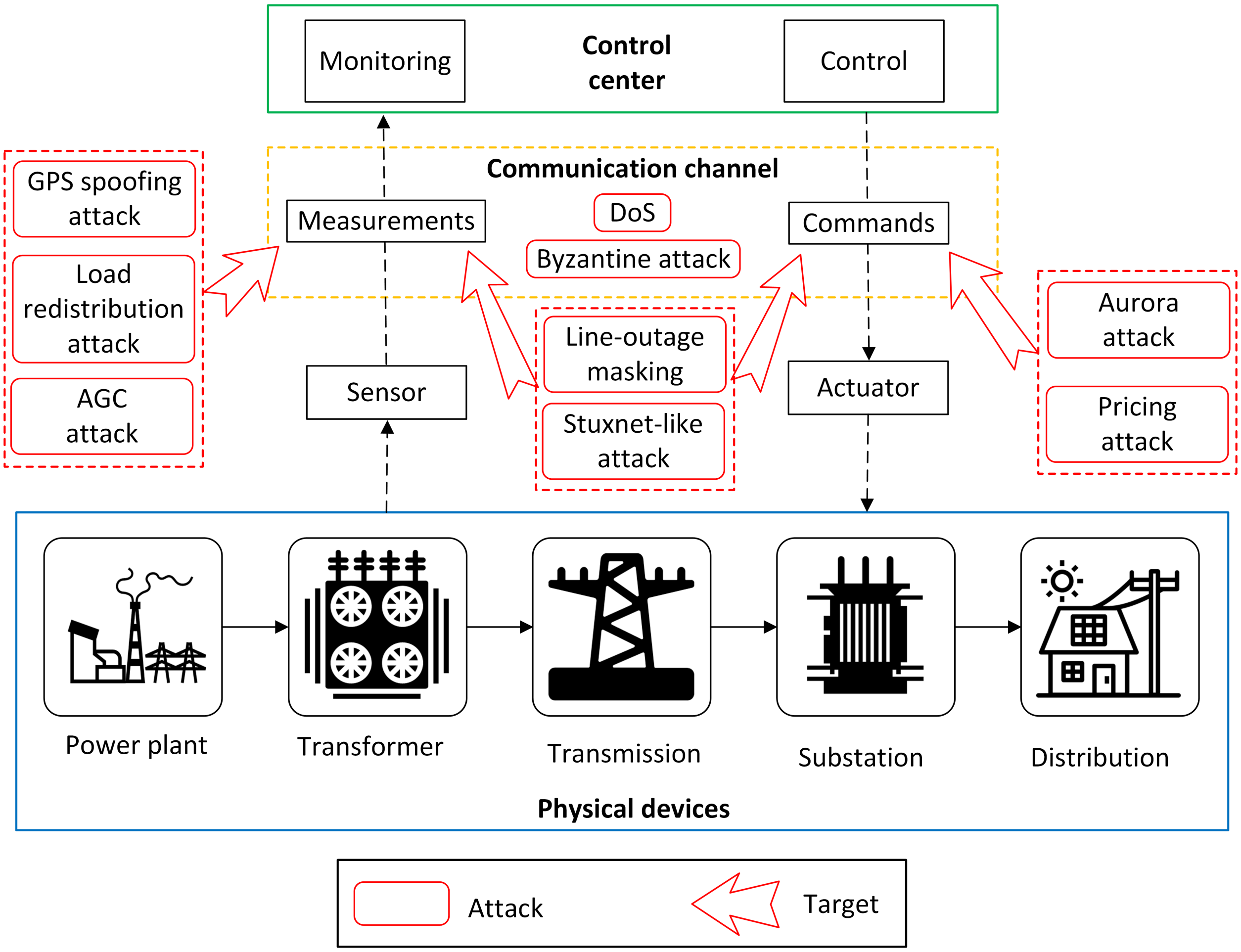


Figure 1.2: Cyber-physical attacks and their targets reviewed in this dissertation.

1.3 False Data Injection Attack

FDI attacks have become a significant threat to smart grids. As the communication channels between meters and control centers are vulnerable to cyberattacks, attackers can launch man-in-the-middle attacks to eavesdrop and manipulate the measurement signals. FDI attack, as one of the most infamous cyberattacks, can stealthily mislead the power system state estimation by injecting false data into legitimate measurements. To bypass the defender's bad data detector, the attacker needs to elaborately construct an AC-FDI attack vector a by following the equation:

$$\mathbf{a} \triangleq h(\hat{x} + c) - h(\hat{x}) \tag{1.3}$$

where $h(\bullet)$ are the measurement functions, $c \in \mathbb{R}^n$ denotes the bias vector that the attacker intends to mislead the state estimation, n is the number of system states, and $\hat{x}$ is the estimated state. When the legitimate measurement $\mathbf{M}$ is altered by the manipulated measurement $\mathbf{M}_a = \mathbf{M} + \mathbf{a}$, the BDD residual after the attack will be less or equal to the residual before the attack, i.e., the attack is stealthy.

FDI attack vectors can also be implemented in DC power systems. In a DC model, attackers need to specify a state increment, i.e., $\Delta\theta$, to construct and launch a successful DC-FDI attack. An FDI attacker can compromise estimated states without being detected by BDD, if the attack vector $\mathbf{a}$ is calculated by $\mathbf{a} = \mathbf{H} \cdot \Delta\theta$[19].

The illustration of FDI attacks is shown in Figure 1.3. FDI attackers inject adversary data $\mathbf{a}$ into the legitimate measurements $\mathbf{M}[t]$ based on the DC- or AC-FDI attack model. The manipulated measurements $\mathbf{M}_a[t]$ received by the SCADA system can bypass bad data detection without alerts and induce a bias in the estimated states. The estimated states will be used in the applications of EMS, including the contingency analysis, automatic generation control, and optimal power flow (OPF) model. S_d and S_G are the apparent load power and generation, respectively.

Yuan et al.[20], for the first time, proposed a special case of FDI attacks, i.e., load redistribution (LR) attack. With the increasing penetration of renewable-based distributed energy resources (DERs), the malicious manipulation of net load measurements (load minus DER generation) at DER buses can be disguised as renewable generation uncertainty. Therefore,

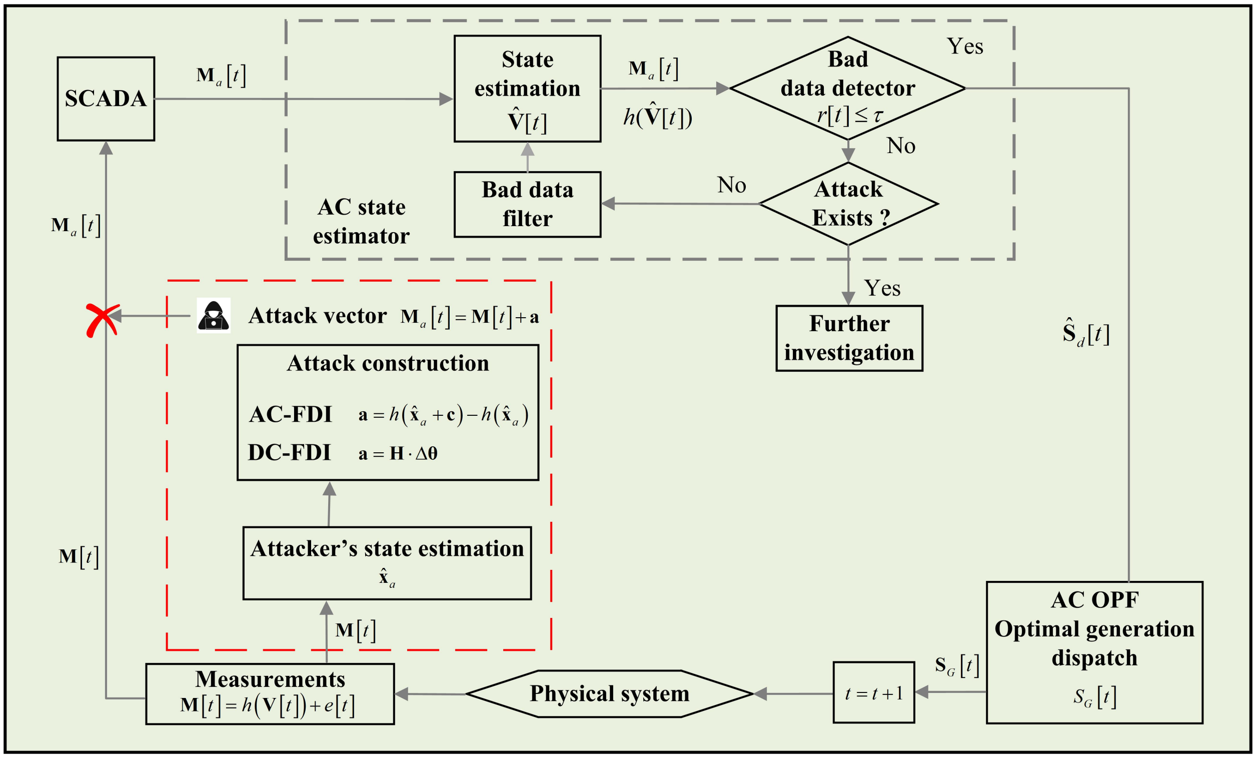

SCADA
$\mathbf{M}_a[t]$
State estimation $\hat{\mathbf{V}}[t]$
$\mathbf{M}_a[t]$
$h(\hat{\mathbf{V}}[t])$
Bad data detector $r[t] \le \tau$
Yes
No
Bad data filter
No
Attack Exists ?
AC state estimator
$\mathbf{M}_a[t]$
Attack vector $\mathbf{M}_a[t] = \mathbf{M}[t] + \mathbf{a}$
Yes
Further investigation
$\hat{\mathbf{S}}_d[t]$
Attack construction
AC-FDI $\mathbf{a} = h(\hat{\mathbf{x}}_a + \mathbf{c}) - h(\hat{\mathbf{x}}_a)$
DC-FDI $\mathbf{a} = \mathbf{H} \cdot \Delta\boldsymbol{\theta}$
Attacker's state estimation $\hat{\mathbf{x}}_a$
$\mathbf{M}[t]$
$\mathbf{M}[t]$
AC OPF Optimal generation dispatch $S_G[t]$
$\mathbf{S}_G[t]$
$t = t+1$
Physical system
Measurements $\mathbf{M}[t] = h(\mathbf{V}[t]) + e[t]$

Figure 1.3: Illustration of FDI attack model in the smart grid.

considering the attacker's practical capability of manipulating the net load measurements, this dissertation introduces an improved LR attack strategy, namely net load redistribution attack[21]. The goal of the net load redistribution attack is to mislead the AC state estimation with an illusory over- or under-voltage issue by injecting highly-structured attack vectors into the measurements. To bypass the BDD, the net load redistribution attack stealthiness constraints pertaining to boundary conditions between the attack and non-attack areas were proposed. Those constraints included restrictions on voltage magnitude measurements on the boundary buses and power flow measurements on the tie lines. With the required local information within the attack region and the stealthiness constraints, the net load redistribution attack is modeled as an ACOPF problem for attackers, in which the prevailing ACOPF constraints hold.

1.4 State Estimation and Bad Data Detection

SE is essential in power systems, providing the estimated voltage states to EMS applications. Given the measurements $\mathbf{M}$, the system states can be estimated by solving the following weighted least square (WLS) optimization[22]:

$$\hat{\mathbf{x}} = \min_{x} \left[\mathbf{M} - h\left(\mathbf{x}\right)\right]^T \mathbf{W}^{-1} \left[\mathbf{M} - h\left(\mathbf{x}\right)\right]$$

where $h\left(\textbf{.}\right) : \mathbb{R}^{2n-1} \rightarrow \mathbb{R}^{m}$ is a vector of nonlinear functions that reveal the relationship between the measurements vector $\mathbf{M} \in \mathbb{R}^{m}$ and the state vector $\mathbf{x} \in \mathbb{R}^{2n-1}$.

In power systems, it is customary to use the Gauss-Newton iterative algorithm to solve the above optimization problem. The iterative process converges when the difference between the system states in two iterative is smaller than a pre-determined threshold.

$$\hat{\mathbf{x}}_{k+1}{=}\hat{\mathbf{x}}_k + \Big(\ \mathrm{H}\left(\hat{\mathbf{x}}_k\right)^T \mathbf{W}^{-1} \mathrm{H}\left(\hat{\mathbf{x}}_k\right)\Big)^{-1} \mathrm{H}\left(\hat{\mathbf{x}}_k\right) \mathbf{W}^{-1} \left(\mathbf{M} - h\left(\hat{\mathbf{x}}_k\right)\right)$$

where $\mathrm{H}\left(\mathbf{x}\right)^T = \partial h\left(\mathbf{x}\right) / \partial \mathbf{x}$ is the Jacobian matrix.

Bad Data Detection (BDD) in AC state estimation is based on a residual analysis of $\mathbf{r} = \mathbf{M} - h(\hat{\mathbf{x}})$, where $\mathbf{r}$ is the residual, $\hat{\mathbf{x}}$ is the estimated system states. The residual can be attributed to noise and inaccuracy of the measurements as well as injected false data. The defender, usually a system operator, runs the BDD tests by comparing the residual with a pre-defined threshold τ calculated at a certain confidence interval.

$$\|\mathbf{r}\|_2 = \|\mathbf{M} - h(\hat{\mathbf{x}})\|_2 < \tau$$

When there is at least one faulty measurement, the l_2 norm of residual exceeds the threshold.

1.5 Introduction to Moving Target Defense

The United States Department of Homeland Security defines MTD as the idea of controlling perturbations across multiple system dimensions to increase uncertainty and apparent complexity for adversaries, downsizing their window of opportunity, and increasing the costs of their probing and attack efforts. Researchers have proved that MTD is a promising defense method in IT systems[23]. IT systems usually operate in static configurations, and attackers have ample reconnaissance time. Static configurations are vulnerable to various attacks, including replay attacks, denial-of-service attacks, and man-in-the-middle attacks[24]. When MTD is implemented, it provides the systems with improved security by constantly altering the system configurations. Researchers have proposed various MTD applications, including network address shuffling[25], address space layout randomization[26], moving target internet protocol version 6 defense[24], MTD platform for cloud-based IT systems[23], and instruction-set randomization[27].

MTD in power systems provides proactive defense in contrast to the traditional remedial defense approaches. Unlike the MTD in information technology systems which highlights the changes in the network layer, MTD in power systems require physical devices and extra control. The essence of MTD is that it actively perturbs the transmission line impedance to invalidate attackers' knowledge about the power system configurations. Since MTDs can utilize many devices to perturb the line impedance equivalently, this dissertation focuses on MTDs based on distributed flexible AC transmission system (D-FACTS) devices. D-FACTS devices, including Static Var Compensators (SVC), Thyristor Controlled Series Capacitors (TCSC), and Static Synchronous Series Compensators (SSSC), are utilized initially to control power flows. They can also manage power congestion and minimize system losses by altering the impedance of power lines[28]. With the increase of D-FACTS devices[29], their add-on cyber-physical security benefits via MTD have attracted increasing attention in the cyber-physical security research community.

The block diagram in Figure 1.4 shows an MTD-enabled power system measurement-control-loop in wide area monitoring, protection and control. Attackers can eavesdrop on the power system measurement data and inject the manipulated measurement back into the system. If the attackers have knowledge of the system configuration, they can construct and inject stealthy FDI attack vector M_a into the SCADA system. M_a can bypass an AC state estimation based BDD[30] if there is no MTD activated. When an MTD is activated, the attacker's knowledge about the system configuration $h(\bullet)$ will be outdated, and the injected attack vector that is constructed based on the outdated $h(\bullet)$ can be detected by BDD[31]. In this case, further investigation can be conducted to identify the attack vector under some conditions[32].

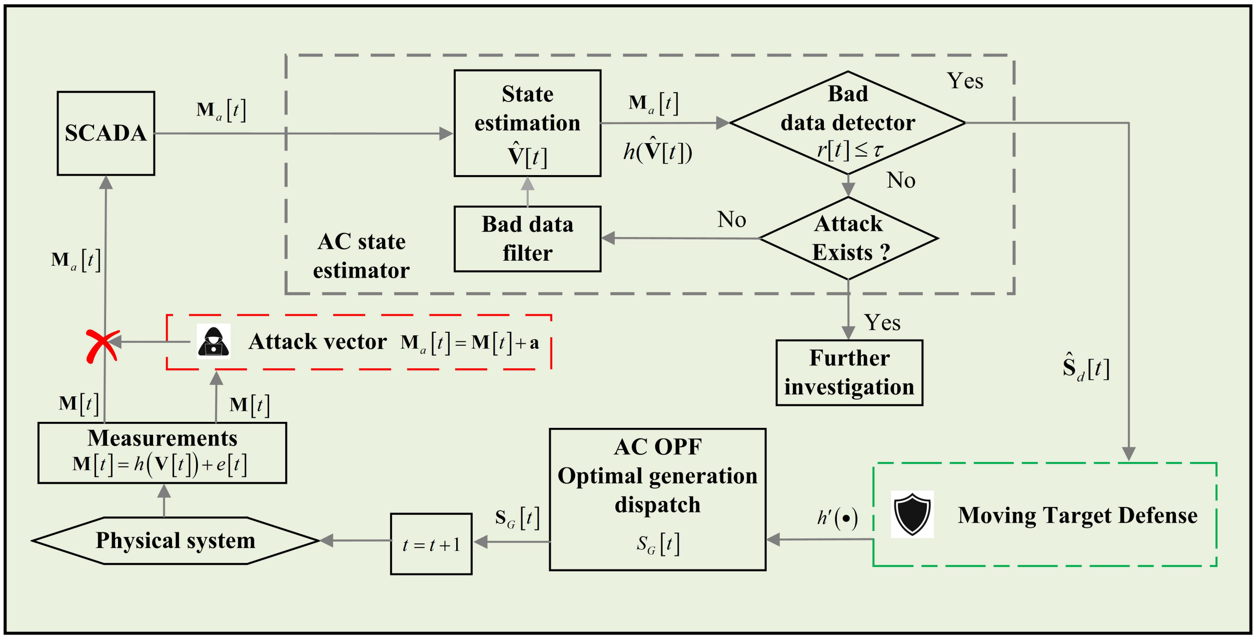


Figure 1.4: Illustration of MTD model in the smart grid.

This dissertation constructs a dynamic simulation on the Typhoon real-time digital simulator to visualize the aforementioned MTD against FDI process. The simulation platform consists of a digital circuit and a SCADA system, as demonstrated in Fig. 1.5 and 1.6. The meters deployed in the circuit collect measurement data and transfer them to the SCADA system. The SE and BDD then process the received data to evaluate the measurement. When the residual value is lower than the threshold value, the BDD alarm light is off, and the system is considered free of bad data or attack. When a stealthy FDI attack is injected, the estimated voltage angle at bus 10 is incorrect, as shown in Fig. 1.7. Meanwhile, the BDD residual under attack is still lower than the threshold, meaning the attack is stealthy. However, when the MTD is activated, the system line impedance changes, and the attacker cannot obtain the updated line impedance in time. In this case, the injected FDI attack vector will be detected by the BDD, as shown in Fig. 1.8.

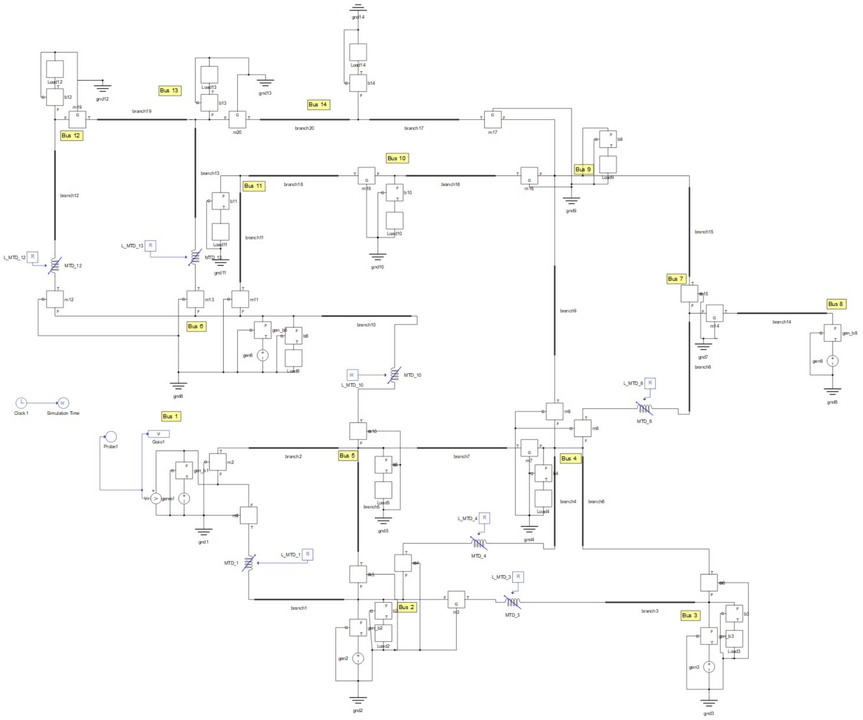
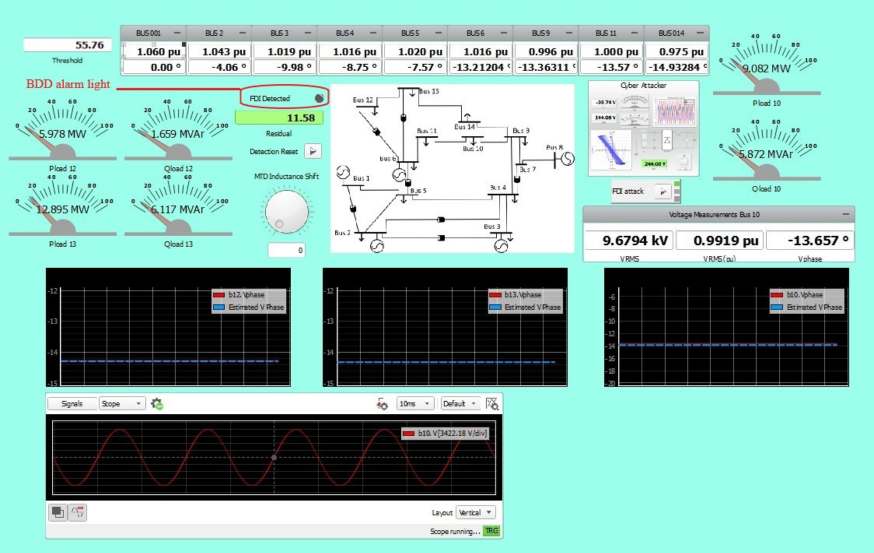


Figure 1.5: 14-bus circuit in Typhoon real-time simulator.

1.6 Introduction to Meter Encoding

Besides the research on MTD, researchers have proposed other methods for detecting data integrity attacks in cyber-physical systems. One such technique, known as meter encoding, involves changing the values of sensor readings to spoil the capability of the attacker to design a stealthy attack sequence in measurements. Liu[33] created a meter encoding strategy to increase the residuals obtained by the Kalman Filter applied to the undetectable attack sequence capable of introducing estimation errors. Existing methods, however, are tailored for IT systems, and their applicability to power systems is limited.



BUS 001	BUS 2	BUS 3	BUS 4	BUS 5	BUS 6	BUS 9	BUS 11	BUS 014
1.060 pu	1.043 pu	1.019 pu	1.016 pu	1.020 pu	1.016 pu	0.996 pu	1.000 pu	0.975 pu
0.00 °	-4.06 °	-9.98 °	-8.75 °	-7.57 °	-13.21204 °	-13.36311 °	-13.57 °	-14.93284 °

Threshold: 55.76

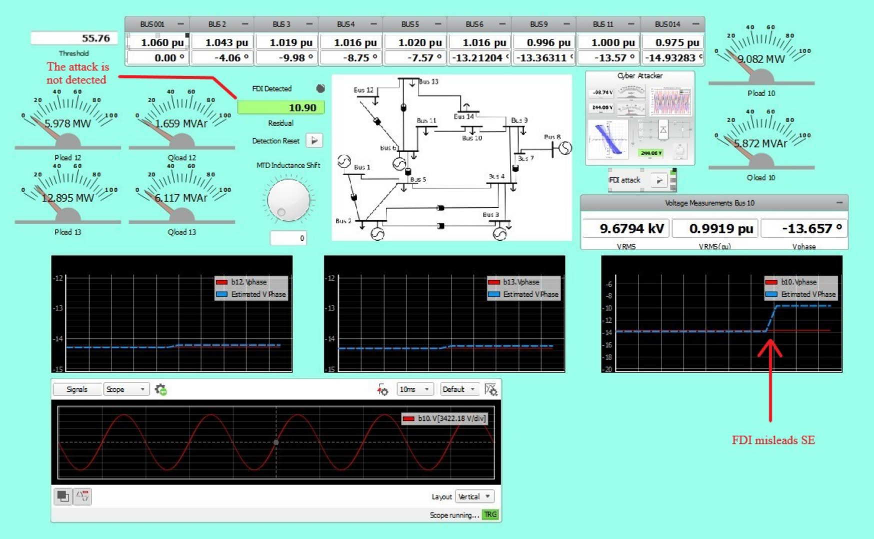
Residual: 11.58

MTD Inductance Shift: 0

Voltage Measurements Bus 10

VRMS	VRMS(pu)	Vphase
9.6794 kV	0.9919 pu	-13.657 °

Figure 1.6: 14-bus SCADA in Typhoon real-time simulator.



BUS 001	BUS 2	BUS 3	BUS 4	BUS 5	BUS 6	BUS 9	BUS 11	BUS 014
1.060 pu	1.043 pu	1.019 pu	1.016 pu	1.020 pu	1.016 pu	0.996 pu	1.000 pu	0.975 pu
0.00 °	-4.06 °	-9.98 °	-8.75 °	-7.57 °	-13.21204 °	-13.36311 °	-13.57 °	-14.93283 °

Threshold: 55.76

Residual: 10.90

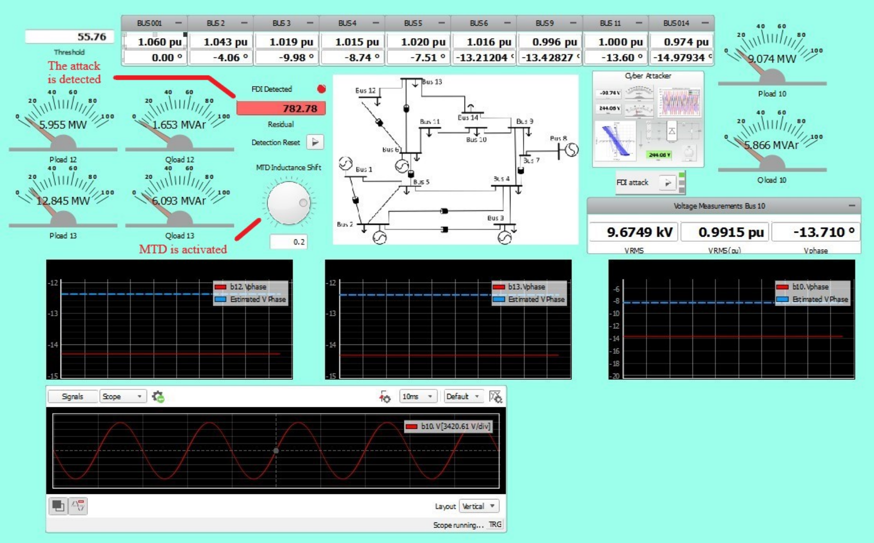
MTD Inductance Shift: 0

Voltage Measurements Bus 10

VRMS	VRMS(pu)	Vphase
9.6794 kV	0.9919 pu	-13.657 °

Figure 1.7: FDI attack impact in Typhoon real-time simulator.

Various defending strategies have been proposed to detect stealthy FDI attacks recently in smart grids. For instance, protection-based defenses[34;35] are proposed to deter attacks by protecting a set of critical measurements. One disadvantage of the protection-based strategies is that advanced encryption technologies are required, which can cause unacceptable latency[36;37] in power systems. Therefore, IEC 62351 standard stipulates that encryption algorithms are not recommended[38]. Another type of emerging defense strategy is proactive detection, which exploits the fact that the stealthiness of FDI attacks closely depends on the



BUS001	BUS 2	BUS 3	BUS4	BUS5	BUS6	BUS9	BUS 11	BUS014
1.060 pu	1.043 pu	1.019 pu	1.015 pu	1.020 pu	1.016 pu	0.996 pu	1.000 pu	0.974 pu
0.00 °	-4.06 °	-9.98 °	-8.74 °	-7.51 °	-13.21204 °	-13.42827 °	-13.60 °	-14.97934 °

Threshold: 55.76

The attack is detected

FDI Detected

Residual: 782.78

Detection Reset

MTD Inductance Shift: 0.2

MTD is activated

Voltage Measurements Bus 10

VRMS	VRMS(pu)	Vphase
9.6749 kV	0.9915 pu	-13.710 °

Figure 1.8: MTD detection in Typhoon real-time simulator.

attacker's prior knowledge of the power system, e.g., line impedance and real-time measurements. Proactive detection methods, such as MTD[31;39–43], dynamic watermarking[44;45], and meter encoding[46–49], have been studied to detect FDI attacks by actively perturbing transmission line reactance, control inputs, or online measurements, respectively. Meter encoding is a side-effect-free technique to detect FDI attacks compared with MTD and watermarking, which may unavoidably increase power losses or decrease system stability. Meter encoding strategies can detect FDI attacks by encoding the meter outputs and decoding them at control centers without significantly affecting the physical operation of power systems.

The flowchart of the meter encoding against stealthy FDI attacks is shown in Figure 1.9. The measurement vector $\mathbf{M}$ from the physical power system is processed by the encoding scheme before being transmitted by the communication network. Given a secret encoding vector μ and the encoding function $f(\mathbf{M}, \mu)$, the output $\mathbf{M}^{\mu}$ of an encoder consists of the encoded measurements from the encoded meters and legitimate measurements from the conventional meters. Both encoder and decoder have access to the secret encoding vector μ, which is the bias between the encoded and original measurements. The encoding function f in the encoder adds the encoding vector to the original measurements from the encoded meters. Correspondingly, a decoding function g substracts the encoding vector from the encoded measurements. After receiving $\mathbf{M}^\mu$, the control center will decode the received data by using the decoding function $g(\mathbf{M}^\mu, \mu)$. The decoded measurements $\mathbf{M}^d$ will be tested by the state estimation-based BDD first to check if the measurements contain bad or manipulated data.

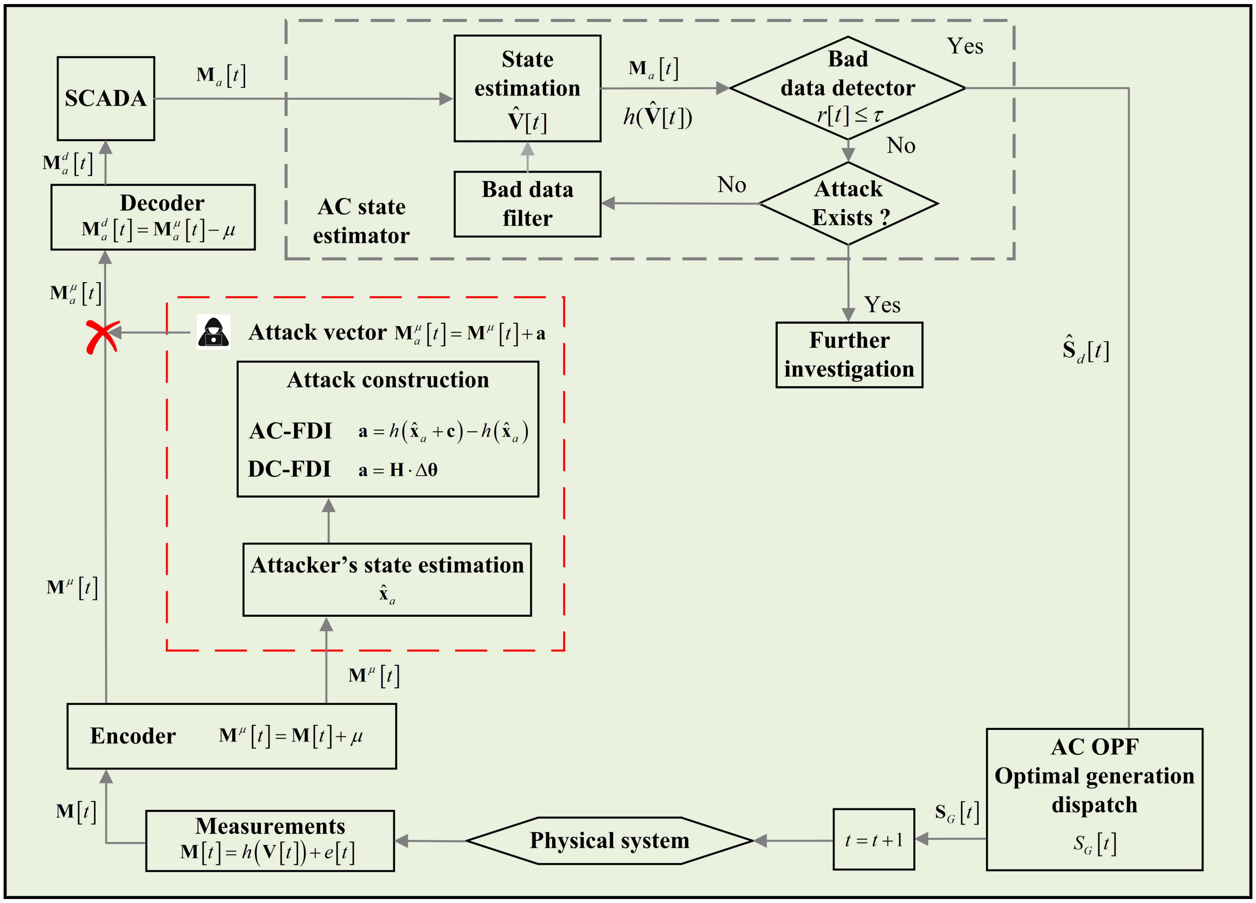

SCADA
$\mathbf{M}_a[t]$
State estimation $\hat{\mathbf{V}}[t]$
$\mathbf{M}_a[t]$
$h(\hat{\mathbf{V}}[t])$
Bad data detector $r[t] \le \tau$
Yes
No
No
Attack Exists ?
Bad data filter
AC state estimator
$\mathbf{M}_a^d[t]$
Decoder $\mathbf{M}_a^d[t] = \mathbf{M}_a^\mu[t] - \mu$
$\mathbf{M}_a^\mu[t]$
Attack vector $\mathbf{M}_a^\mu[t] = \mathbf{M}^\mu[t] + \mathbf{a}$
Attack construction
AC-FDI $\mathbf{a} = h(\hat{\mathbf{x}}_a + \mathbf{c}) - h(\hat{\mathbf{x}}_a)$
DC-FDI $\mathbf{a} = \mathbf{H} \cdot \Delta\boldsymbol{\theta}$
Attacker's state estimation $\hat{\mathbf{x}}_a$
Yes
Further investigation
$\hat{\mathbf{S}}_d[t]$
$\mathbf{M}^\mu[t]$
$\mathbf{M}^\mu[t]$
Encoder $\mathbf{M}^\mu[t] = \mathbf{M}[t] + \mu$
$\mathbf{M}[t]$
Measurements $\mathbf{M}[t] = h(\mathbf{V}[t]) + e[t]$
Physical system
$t = t + 1$
$\mathbf{S}_G[t]$
AC OPF Optimal generation dispatch $S_G[t]$

Figure 1.9: Illustration of meter encoding in the smart grid.

1.7 Research Motivations

Cyber-physical incidents introduced in Chapter 1.1 on page 1 pose a major threat to the cyber-physical security of smart grids. The attacks listed in Fig. 1.2 can mislead the smart grid control system, cause physical damage to devices, induce blackout, and even trigger cascading failures. Since smart grids, as critical infrastructures, play a crucial role in our everyday life, their security needs to be considered as one of the most important challenges in this modern era. Nations worldwide have recognized the threats of cyber-physical attacks against smart grids. The United States has invested $210 million in smart grid cybersecurity research since 2010[50]. Canada has invested $40 billion in enhancing power system infrastructures to achieve a secure and reliable power system[50]. It is urgent to study the cybersecurity of smart grids from both an attacker's and a defender's perspectives to improve the reliability of smart grids. This dissertation can help to achieve this goal in four ways, including reviewing the state-of-the-art cybersecurity in smart grids, studying a realistic FDI attack, improving the existing MTD approaches, and proposing the smart inverter-based meter encoding.

From an attacker's perspective, existing FDI constructions are costly and difficult to achieve. There are two strong assumptions in current FDI literature, i.e., attackers know the complete topology of a power system and the accurate system state to construct an FDI attack vector. How to release these two strong assumptions is the crucial question for an attacker to launch realistic attacks on power systems. Meanwhile, the majority of FDI attacks are studied in transmission systems. The impact of FDI attacks on distribution systems considering the DER is still unclear. Unlike a transmission system, the distribution system is characterized by a radial network typology and a low X/R ratio. Therefore, the current approximation-based FDI construction method[51] does not apply to the distribution system as significant approximation errors occur in the estimation of voltage angle differences in the distribution system. Furthermore, there has been little effort to model a stealthy FDI attack with a tangible attack goal in the distribution system. Without such a goal, it is impossible for a defender to analyze the real consequence of an FDI attack. The author of this dissertation wants to emphasize that the intention of this dissertation is not to educate the attackers on how to perform FDI attacks but to provide power grid operators with a better understanding of attack consequences, which in turn can assist in devising effective defense approaches.

Existing MTD operations may deviate the steady-state operating point of a power system from its optimal one, causing massive economic and stability impacts[52]. In[53], voltage stability is defined as the ability of a power system to maintain steady voltages at all buses after being subjected to a disturbance. One of the most common disturbances is the load increases that occur due to the peak load period. To maintain stability after such disturbance, the system needs the preserved capabilities of the transmission network for power transfer. The action of MTD perturbation which changes the transmission line impedance, may degrade the power transfer capability and cause voltage instability during the peak load period. To the best of the author's knowledge, there is no research on MTD operations to detect FDI attacks while guaranteeing system stability. Furthermore, even if existing MTD operational approaches[54–56] are proposed to follow some security constraints such as power flow limits and safe voltage boundaries, all those approaches consider a single-hour system load without taking into account forecasted load variations in look-ahead time periods. This might be plausible for AC-OPF since it is frequently implemented, e.g., on an hourly basis. However, the frequency of the MTD can be several hours to a few days depending on the attacker's capabilities as well as how a system operator executes it (e.g., an event-based MTD strategy[57]). The lack of such look-ahead capabilities in existing MTD methods may cause voltage instability or even voltage collapse due to the reduction of load margin or voltage stability degradation between two consecutive MTD executions. Therefore, addressing the impact of MTD on the system voltage stability is an urgent task.

As D-FACTS devices are usually unavailable in distribution systems, meter encoding strategies are proposed to replace MTD in detecting FDI attacks. There are two metrics to evaluate the meter encoding methods, i.e., encoding cost and encoding hiddenness. Most current meter encoding focuses on implementing their proposed schemes on conventional meters. Additional ZigBee modules and cellular network communication devices must be deployed for the conventional meters within different substations to support data transmission between meters and encoders. Meanwhile, defenders need microprocessors to implement meter encoding because conventional meters are not programmable and can only report correct measurements. These additional devices are expensive and dramatically increase the system's operational cost. As for hiddenness, existing meter encoding methods can be detected by alert attackers using BDD before launching attacks because the encoded measurements are inconsistent with physical laws like Kirchhoff's circuit laws or power flow laws. When attackers detect the unhidden meter encoding, they will not launch attacks until they crack the coding matrices. To fill this gap, proposing a cheap and hidden meter encoding scheme to detect FDI attacks is worthwhile.

This dissertation seeks to address the following unsolved research questions plaguing smart grid cyber-physical security. These questions are from either an attacker's or a defender's perspective.

Question 1: What is the state-of-the-art cyber-physical security in smart grids?

Question 2: How can attackers construct stealthy FDI attacks without the knowledge of complete system topology and the capability of WLS-based SE?

Question 3: When system operators utilize MTD to detect stealthy FDI attacks, what is the impact of MTD on the system voltage stability? Is it possible for an MTD to induce voltage instability? If yes, how to re-dispatch an MTD to avoid voltage instability while keeping the performance of the original MTD strategy?

Question 4: Since D-FACTS devices are usually not available in distribution systems, how can system operators detect stealthy FDI attacks by utilizing programmable smart inverters?

1.8 Research Contributions

This work is unique in that it considers the cyber-physical security of smart grids from both an attacker's and a system operator's perspectives. This dissertation aims to study the construction of stealthy FDI attacks using limited resources and enhance the detection of stealthy FDI attacks with proactive detection methods. Interconnections between these studies also provide a comprehensive attack and detection framework, which can further enhance the cybersecurity of smart grids against data integrity attacks. The objectives, contributions, and research outcomes of this dissertation are outlined in response to the questions raised in Section 1.7.

Question 1: The review chapter comprehensively reviews cyber-physical threat models and defense mechanisms. Over the last five years, several survey and review papers on the cyber-physical security of the smart grid have been published. Table 1.1 lists a comparison between this work and other works regarding the publication year, smart grid models, attack taxonomy, technological focus, challenges and opportunities, and the review scope. The contributions of this review chapter, as illustrated in Table 1.1, are four-fold.

- A discrete-time nonlinear time-invariant system is proposed to represent a CPSG by using the state-space representation. Such a high-level abstraction is a useful strategy to form the foundation and generalize a defense analysis across all attack types.
- The state-of-the-art cyber-physical attack models are summarized based on the proposed abstraction and categorized according to the control-feedback loop segment each attack involves. This new taxonomy provides the grid operator with intuitive situational awareness of enhancing the system's cyber-physical security.
- In order to provide a timely review, this review chapter surveys the most recent publications, including 78 in the last five years (i.e., 2016-2020), 49 of which were published in the past three years (i.e., 2018-2020). A thorough review of cutting-edge defense approaches such as data-driven machine learning, moving target defense, and watermarking is provided.
- The challenges and opportunities of future CPSGs are discussed, which may shed light on cyber-physical security issues that the next-generation smart grid needs to tackle.

These contributions are discussed in Chapter 2 and in the following article:

H. Zhang, B. Liu and H. Wu, "Smart Grid Cyber-Physical Attack and Defense: A Review," in IEEE Access, vol. 9, pp. 29641-29659, 2021, doi: 10.1109/ACCESS.2021.3058628.[58]

Question 2: A novel net load redistribution attack is proposed to falsify nodal voltage magnitude estimation in the AC distribution system. Specifically, the attackers intend to create illusory voltage violations, such as under-voltage violations, to the system operator.

- The NLRA is constructed as an attacker's AC-OPF problem that only requires local line impedance information.
- The machine learning-based SE with two deep neural networks is proposed to construct NLRA. This SE can address the issue that attackers may not have enough redundant measurements to implement the conventional WLS-based SE in distribution systems. The performance of the attacker's machine learning-based SE is evaluated under low-observable conditions.
- The NLRA model is solved by using an interior-point algorithm, and simulations are conducted on a modified PG&E 69-node distribution system.

These contributions are discussed in Chapter 3 and in the following article:

H. Zhang, B. Liu and H. Wu, "Net Load Redistribution Attacks on Nodal Voltage Magnitude Estimation in AC Distribution Networks," 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 2020, pp. 46-50, doi: 10.1109/ISGT-Europe47291.2020.9248915.[21]

Question 3: Chapter 4 proposes a novel voltage-stability-constrained MTD framework against highly structured FDI attacks, especially in the presence of stressful system conditions. The contributions of Chapter 4 are described as follows:

- It is revealed through a 3-bus system and two more complex systems that a system with the existing MTD operation methods can suffer voltage instability or even experience voltage collapse at the peak load.
- A voltage stability (t-index) optimization method is proposed to enhance the original MTD strategies. Specifically, the sensitivity matrix of the voltage stability index with respect to line impedance is mathematically derived. The proposed optimization method maximizes the lowest index value among all the load buses with the minimum impedance adjustment; therefore, the system voltage stability is considered while the impact on the original MTD strategy is minimized.

- A load margin optimization method is developed based on Continuation Power Flow[69] (CPF) to ensure a sufficient load margin for system voltage stability at the most stressful time period. The power injection to impedance sensitivity is utilized to calculate safe MTD setpoints adjustment with ample load margins.

- A new MTD framework is presented, which seamlessly integrates the above two voltage stability constrained methods into the original MTD operational methods. Case studies on IEEE 14-bus and 118-bus systems are conducted to test the proposed MTD framework against one of the most sophisticated FDI attacks, i.e., NLRA.

These contributions are discussed in Chapter 4 and in the following articles:

H. Zhang, B. Liu, X. Liu, A. Pahwa and H. Wu, "Voltage Stability Constrained Moving Target Defense Against Net Load Redistribution Attacks," in IEEE Transactions on Smart Grid, vol. 13, no. 5, pp. 3748-3759, Sept. 2022, doi: 10.1109/TSG.2022.3170839.[42]

H. Zhang, N. Fulk, B. Liu, L. Edmonds, X. Liu and H. Wu, "Load Margin Constrained Moving Target Defense against False Data Injection Attacks," 2022 IEEE Green Technologies Conference (GreenTech), 2022, pp. 51-56, doi: 10.1109/GreenTech52845.2022.9772024.[70]

Question 4: Chapter 5 aims to detect FDI attacks in distribution systems by designing a smart-inverter-based meter encoding scheme. The main contributions of this chapter are outlined as follows:

- A smart-inverter-based meter encoding scheme is proposed to detect FDI attacks in distribution systems. The proposed meter encoding can mislead the attacker's SE while not being detected by alert attackers.

- It is proved that if an inverter bus is encoded, all the downstream buses on that lateral will be protected by the proposed smart-inverter-based meter encoding. FDI attacks that target these buses will be detected.

- A comprehensive evaluation is constructed to test the detection effectiveness of the proposed smart-inverter-based meter encoding against strong attackers. In this dissertation, FDI attackers can obtain the necessary system state using either the WLS-based state estimator or the power flow enhanced deep learning (PFEDL) state estimator.

These contributions are discussed in Chapter 5 and in the following article:

H. Zhang, B. Liu and H. Wu, "Smart Inverter Enabled Coding Scheme for Detecting False Data Injection Attacks in Distribution System State Estimation" in IEEE Open Access Journal of Power and Energy, 2022, under review

In conclusion, this work proposes the utilization of a realistic FDI attack and two proactive defense methods to enhance the cybersecurity of smart grids. First, the NLRA model is proposed to release the strong assumptions in the existing FDI attack construction. Next, the impact of MTD on long-term voltage stability is explored. Two optimization models are proposed to avoid MTD-induced voltage instability while maintaining the detection effectiveness against NLRA. Finally, smart-inverter-based meter encoding is proposed to detect FDI attacks in distribution systems. Compared with MTD, the proposed meter encoding does not require installing additional hardware devices, such as D-FACTS.

1.9 Organization of This Dissertation

The organization of this dissertation is shown in Fig. 1.10. In each block in this figure, the research questions and the main contributions are summarized. Chapter 1 introduced the concepts in this dissertation, including the background, research motivations, and contributions of this dissertation. In Chapter 2, a comprehensive review is presented to study the state-of-the-art cyber-physical security in smart grids. In Chapter 3, a realistic FDI attack, namely NLRA, is proposed to mislead the operator's situational awareness by manipulating the measurements. The attackers only need lock line impedance information and much fewer measurements than traditional FDI strategies to construct the proposed attack vector in distribution systems. Chapter 4 first evaluates the impact of existing MTD on long-term voltage stability, then proposes two optimization methods to ensure voltage stability while maintaining the performance of the original MTD strategies. In Chapter 5, a software-based detection method, namely smart-inverter-based meter encoding, is proposed to detect FDI attacks in distribution systems. The proposed meter encoding does not require D-FACTS devices, thus, addressing the issue that it is unrealistic to implement MTD in distribution systems. Finally, Chapter 6 presents key conclusions of this work and discusses future research directions.

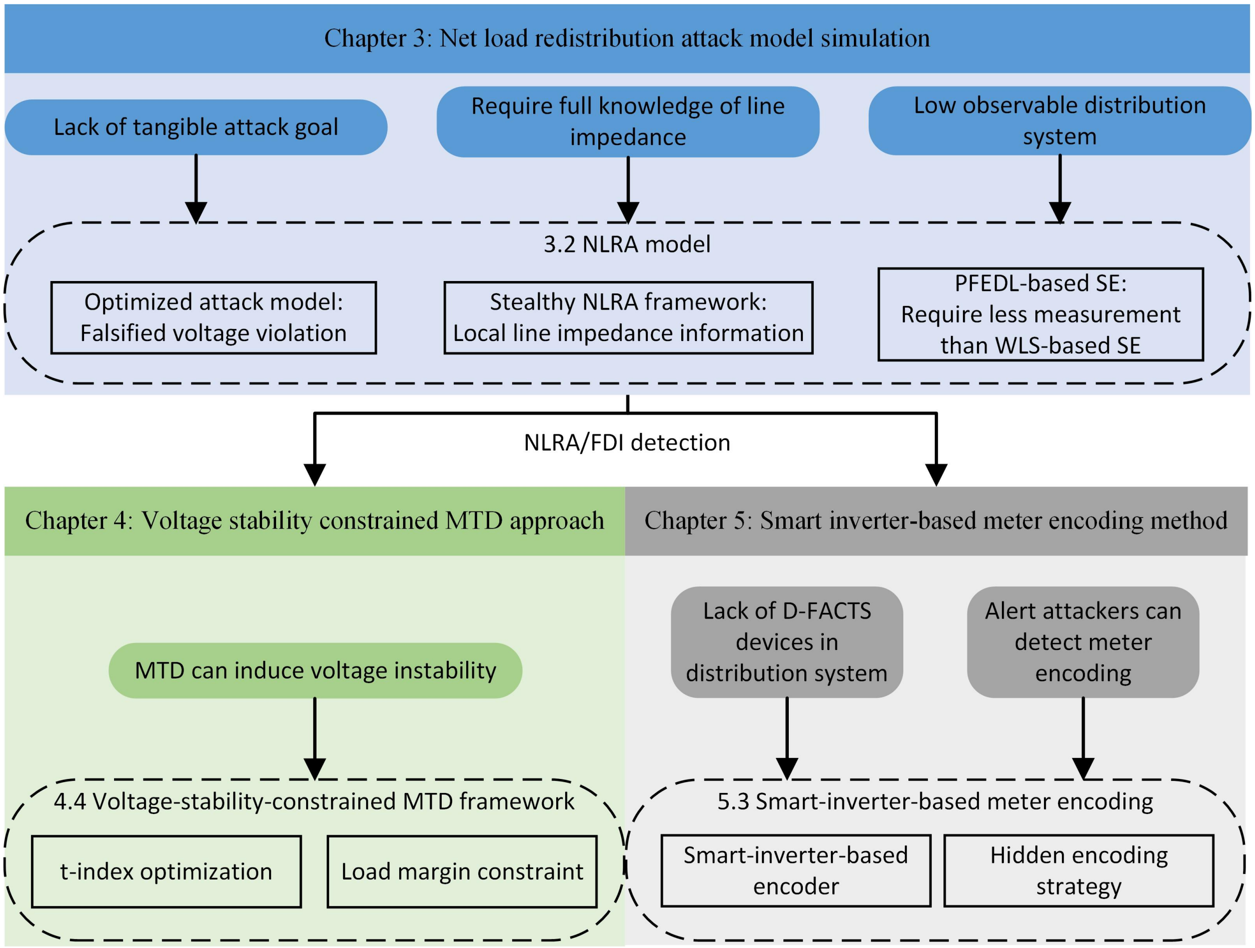

Chapter 3: Net load redistribution attack model simulation
Lack of tangible attack goal
Require full knowledge of line impedance
Low observable distribution system
3.2 NLRA model
Optimized attack model: Falsified voltage violation
Stealthy NLRA framework: Local line impedance information
PFEDL-based SE: Require less measurement than WLS-based SE
NLRA/FDI detection
Chapter 4: Voltage stability constrained MTD approach
Chapter 5: Smart inverter-based meter encoding method
MTD can induce voltage instability
Lack of D-FACTS devices in distribution system
Alert attackers can detect meter encoding
4.4 Voltage-stability-constrained MTD framework
t-index optimization
Load margin constraint
5.3 Smart-inverter-based meter encoding
Smart-inverter-based encoder
Hidden encoding strategy

Figure 1.10: Structure diagram of this dissertation.

Table 1.1: A comparison of related literature

Ref.-Yr	**CPSG model**	**Taxonomy**	**Attack types**	**Challenge and opportunity**	**Scope (attack or defense)**	**Technological focus**
[59]-2015	Conceptual model	N/A	DoS	Yes	Both attack and defense	Informational
[60]-2015	Conceptual model	Layers (physical, MAC, network, application)	CIA triad attacks	Yes	Attack only	Both informational and operational
[61]-2016	N/A	CPSG components	CIA triad attacks	No	Attack only	Informational
[62]-2016	Abstract model	Security objectives (Confidentiality, integrity, availability)	CIA triad attacks	No	Both attack and defense	Both informational and operational
[63]-2016	N/A	Layers (communication, measurement, control)	DoS, wrapping, phishing attacks	No	Attack only	Both informational and operational
[2]-2016	Abstract model	Layers (generation, transmission, distribution)	Control, measurement attacks	Yes	Both attack and defense	Both informational and operational
[64]-2017	N/A	N/A	Malware	Yes	Attack only	Informational
[65]-2018	Abstract model	Source of threats	Technical and non-technical attacks	No	Attack only	Both informational and operational
[66]-2018	N/A	Attack behavior	Interruption, interception, modification, fabrication attacks	No	Both attack and defense	Both informational and operational
[67]-2019	Mathematical model (linear time invariant)	Spatial–temporal hiddenness	FDI, topology attack, DoS, replay attack, Stuxnet, dynamic attack	No	Both attack and defense	Both informational and operational
[68]-2020	Conceptual model	Security objectives (Confidentiality, integrity, availability)	CIA triad attacks	Yes	Both attack and defense	Informational
This chapter-2021	Mathematical model (non-linear time invariant)	Control and feedback loop (control, measurement, control-measurement)	Aurora, pricing attacks, AGC attacks, FDI, Topology attacks, GPS-spoofing, Line-outage masking, Stuxnet-like attacks	Yes	Both attack and defense	Both informational and operational

Chapter 2

Smart Grid Cyber-Physical Attack and Defense: A Review

Recent advances in the CPSG have enabled a broad range of new devices based on information and communication technology. However, these ICT-enabled devices are susceptible to a growing threat of cyber-physical attacks. This chapter thoroughly reviews the state-of-the-art cyber-physical security of the smart grid. By focusing on the measurements and control loop introduced in Figure 1.1, this chapter first categorizes the existing cyber-physical attacks in terms of their target components. This chapter then discusses several operational and informational defense approaches that present the current state-of-the-art in the field, including MTD, watermarking, and data-driven strategies. Finally, this chapter discusses the challenges and future opportunities associated with the cyber-physical security of smart grid.

2.1 Smart Grid Cyber-Physical Attacks Categorization

Figure 1.2 illustrates the cyber-physical attacks and their corresponding targets. Following the WAMPAC scenarios prescribed in[18], this chapter summarizes the cyber-physical attacks in CPSGs in the following three categories:

1. Control signal attacks: By relying on the ability to bypass the data authentication and integration examinations, control signal attacks aim to acquire the physical device authority and then operate it at the attacker's will. This type of attack is usually designed to target mission-critical devices in power systems, such as automatic generation control (AGC), relays, smart inverters, flexible AC transmission system (FACTS) devices, and circuit breakers. To achieve the adversaries' malicious goals effectively, adversaries likely have knowledge of the target device (e.g., inverter P-Q setpoints, generator ramping limits, line flow limits). Despite the study of N-1 contingency for loss of a generator or transmission line, researchers show that by exploiting the clustering-based vulnerability, simultaneous attacks against the elaborately identified, most vulnerable devices may cause cascading failures[71]. Control attacks can achieve significant consequences in a short period. However, the lack of coordinated masks in the feedback measurement makes the attacks unhidden to detection methods.

2. Measurement attacks: These attacks focus on manipulating the sensor measurement data transferred through the communication channels or hacking the remote terminal units (RTU) in the field. Physical communication links are usually compromised to deliver falsified messages (e.g., false data injection attacks, GPS spoofing, and replay attacks). Depending on the attackers' capabilities, they may change the firmware of devices, eavesdrop on measurements for reconnaissance, and control sensors for reporting tampered measurements. For example, an attacker may change the Domain Name Systems (DNS) server of the device gateway to an attacker-controlled DNS server[72]. By doing this, DNS hijacking attacks can be implemented to control the device-remote server interactions. Once an attacker controls the communication between the gateway and the remote server, all the measurement reports are going to be sent to the malicious server instead of the legitimate server. In addition, traditional DoS or a Black Hole can block the packets in the network, decreasing the system's situational awareness. This type of attack may disable the system operator's situational awareness to cover

intrusions or induce inappropriate operations according to the falsified system state based on manipulated measurements.

3. Control-signal-measurement attacks: This type of attack is also called control-measurement-loop attacks, in which adversaries launch coordinated attacks on both the control signals and measurements. The control signal attack may cause immediate physical layer consequences, while the measurement attacks, such as replay attacks, can disguise the ongoing control signal attack. The manipulated measurements can pass existing anomaly detection mechanisms in the system. Existing research revealed that attackers might utilize the control-measurement-loop attack[73] (e.g., line outage masking attacks, Stuxnet attacks) to enhance the stealthiness of control signal attacks. The enhancement is achieved by masking the attack consequences and deceiving the attack detection and mitigation mechanism. For instance, the notorious Stuxnet attack[74;75] targeted the SCADA systems and caused substantial damage to the centrifuge of a nuclear plant. A Stuxnet attack can compromise the programmable logic controllers and give unexpected commands while returning normal operation system measurements to the SCADA.

2.2 Cyber-Physical Attacks

Given the categorization of IT and OT attacks, several attack behaviors against the IT systems of a CPSG are briefly reviewed first. Then, the OT attacks are discussed in greater detail in the rest of this section.

2.2.1 Data Availability Attacks

Since wireless communication is commonly used in a CPSG, adversaries can launch attack schemes against the communication channel. This dissertation classifies the attacks that impede data availability as IT attacks. For example, Byzantine attacks against communication networks such as cognitive radio networks and mobile Adhoc networks were discussed in[76;77]. These attacks are launched by compromised insider nodes to affect the trusted routing, which in turn reduces the overall network performance. After intrusion, a selfish sensing node can report falsified channel sensing results and increase its own gains at the cost of performance degradation of other honest nodes. Typically, attackers intentionally launch Byzantine attacks for two attack objectives. The first objective is vandalism, where attackers report channel vacancy when the sensing results indicate that the channel is busy. The second objective is exploitation, where attackers can access the idle channel exclusively by sending channel busy information when their sensing results indicate that the channel is idle. Attackers can pursue attack utility maximization of the above objectives[78].

Compared with Byzantine attacks that hinder data availability by degrading the communication channel, DoS is another notorious attack that blocks normal data transfer by occupying the communication channel with junk data. In a CPSG, the objective of a DoS attack is to disrupt the communication between a control center and sensors or actuators in the field. DoS attackers are not required to have knowledge of the CPSG configuration or the ability to manipulate the control or measurement data in the communication channel. The attack consequence is that system operators can easily notice the attack due to the loss of measurement data. However, the operators cannot mitigate the attack since they cannot send control signals to the actuators. An example of the DoS attacks is the incident of the Ukrainian electric power companies discussed earlier. In[79], Qin et al. considered how to damage the system performance most severely when launching a DoS attack against the state estimation over the packet-dropping network environment. They presented an optimal attack schedule that maximizes the trace of the average expected estimation error covariance. In[80], Zhang et al. proposed a scenario that a DoS attacker with attack cost constraint jams the sensor-to-estimator communication channel. The authors formulated an optimization problem that balances the destruction on the cost of system control and the cost of attack in an infinite time horizon concurrently.

2.2.2 Control Signal Attacks

Aurora Attacks

The aurora generator vulnerability was originally tested by the Idaho National Laboratory in 2007, where a hypothetical attacker maliciously opened and re-closed the circuit breaker of a generator by injecting a series of compromised control commands[81]. When disconnected from the power grid, the generator becomes desynchronized. The aurora attack is designed to re-close the breaker when the system and generator slip out of synchronism before the protection system responds to the attack. Since generator protection elements are intentionally delayed to prevent unnecessary tripping, attackers typically get a 15-cycle window to re-close the breaker before any protection device kicks in[82]. The physical damage to the generator is caused by the variation of electrical power output from the generator and the incremental generator rotating speed during the aurora attack. Each time the breakers are re-closed, the difference in frequency and phase angle between the main grid and the generator may result in high torque and currents, which can ultimately damage the generator[83].

A scoring methodology with vulnerability ranking criteria to find the most vulnerable breakers for an aurora attack has been presented in[84]. In[85], modeling and impact analysis of aurora attack targeting microgrid point of common coupling (PCC) and synchronous generator breakers are examined. The classic sync-check relays for coping with aurora attacks can lead to unintentional islanding in a microgrid, which is forbidden by the IEEE 1547 Standard[86]. The authors demonstrated that an attacker could successfully damage the microgrid synchronous generator by attacking the PCC breaker of a microgrid connected to the main grid.

Pricing Attacks

Demand-response programs have been drawing more attention from retail markets to increase the efficiency of the power grid. In a basic form, demand-response is a control

Table 2.1: Control signal attack

Control signal attack	**Target**	**Objective**	**Means**	**Consequence**	**Mathematical expression**
Aurora attack[83;86;87]	Generators in power plants, microgrid synchronous generators	Cause damage to generators, motors, and transformers	Intentionally open and close a breaker or PCC breaker	Electromagnetic torque and current fluctuations	Control command injection u_t^a
Pricing attack[88–90]	Price signal, transactive energy systems bid signal	Mismatch between the generated and the consumed power, profitability	Manipulate the price signal, bid prices and bid quantities	System emergencies (e.g., line and transformer overload), economic losses	Price signal manipulation λ_t^a, bid price manipulation b_i^a, bid quantity manipulation q_i^a

mechanism where the control signals are the incentives. In 2013, Tan et al.[87] introduced a pricing attack by performing scaling (sending the scaled value of the true price) and delay (sending old prices) attacks on the price signals. In 2017, Giraldo et al.[91] further improved the attack by modeling an attacker who aims to increase the mismatch between the generated and the consumed power by compromising the communication channel and deploying an attack time series to manipulate the price signal. In contrast to one-snapshot attacks, where the attackers inject malicious data only once, Maharjan et al.[89] consider attacks capable of injecting false pricing information at any moment and repeatedly over a long-time duration. The power mismatch caused by long-term attacks can lead to over-generation, economic losses, and poor power quality. To quantify the impact of the repeated attacks, the authors propose a sensitivity analysis method. In their analysis, the authors utilize a z-transform sensitivity function to model the dynamics of the system.

Zhang et al.[88] analyzed the vulnerabilities in transactive energy systems in 2020. In such a system, the home controllers at the end-user react to the price signal sent by the transactive market and return bid information automatically. Data exchanged between the prosumer and the market agent can be manipulated by attackers. The authors extended the pricing attack using malware to inject both malicious bidding prices and quantities from prosumers. Under these attacks, the market-clearing price was manipulated, and the energy consumption of each individual prosumer was affected, which in turn adversely influenced the overall demand on the distribution feeders. Two attack scenarios were studied in[88], where the first scenario aims at compromising the reliability of the system by manipulating the bid price to some extreme values, while the second scenario aims at making profits over time by manipulating the bid price within limits to avoid being detected. Note that prosumers know these bid limits from the service agreement. If the attacker manipulates the signals such that they are out of the limits, the manipulation will become obvious[89]. In contrast to the first scenario, the attack in the second scenario has a small impact on the total load, which makes it difficult to be detected. Table 2.1 summarizes the existing works on the control signal attacks.

2.2.3 Measurement Attacks

AGC Attacks

Automatic generation control is a wide-area frequency control application in interconnected power grids. It ensures system frequency remains within acceptable bounds and limits the tie-line power flow between adjacent control areas to their scheduled values. AGC relies on power flow and frequency measurements from remote sensors to calculate the area control error (ACE). The ACE represents the power exchange error and the system frequency error between the real system state and the scheduled state. Based on the ACE, automated control commands on AGC generators are computed once every few seconds. However, existing measurement validation techniques, such as state estimation, typically run once every few minutes, which cannot accommodate the second-level frequency of AGC. Therefore, the lack of measurement validation or attack detection mechanism makes AGC susceptible to measurement attacks. Moreover, AGC is a highly automated system that requires minimal supervision and intervention by system operators. Once compromised, it may rapidly cause a power imbalance in the system[90].

Sridhar et al.[92] inject four adverse measurements, i.e., scaling, ramp, pulse, and random attacks, to demonstrate their impacts on the physical system stability and the market operation. In scaling attacks, measurements are modified to higher/lower values during the entire duration of the attack. Ramp attacks gradually increase or decrease original measurements over time. Pulse attacks modify measurements through temporally spaced short pulses. Random attacks add random values to true measurements. In an attack scenario to jeopardize system stability, the attacker's goal is to cause a rapid decline in the system frequency to trigger under-frequency load shedding. In the other attack scenario, to manipulate the market operation and profit by generating more power, the attack involves modification of generator operating points identified by the security-constrained economic dispatch (SCED). In this case, the attacker is a utility that wants to generate more power than the dispatched schedule without being detected. The attacker injects fabricated tie-line power and system frequency measurements to force ACE miscalculation, forcing generators in the targeted area to ramp down. Meanwhile, the attacker ramps up its own generator, thereby generating more than the operating point suggested by SCED. As an increased generation in the attacker's area compensates for a decrease in the targeted area, the system frequency is kept.

Similarly, the four types of attacks discussed above are studied by Chen et al.[93] to explicitly implement the AGC attack strategy targeting the load frequency control. Tan et al.[90] consider that the grid frequency is a global parameter that can be easily verified. They assume there exist upper and lower bounds, known by the attackers, as stealthiness constraints for any injected attack vector to pass the data quality checks. The stealthiness constraints limit the attack vector magnitude and make the attacker unable to cause an unsafe frequency deviation in a single AGC cycle. Thus, Chen et al.[93] focus on attacks on power flow measurements using a continuous false data injection attack over multiple AGC cycles to overcome the stealthiness constraints. They define a metric to assess the effectiveness of their attacks, i.e., Time-to-Emergency (TTE), as the time from the onset of an attack to the first time instant when the average frequency deviation of the system is out of the threshold (e.g., 0.5 Hz) in their case study. They optimize their proposed attack by minimizing the TTE and following the stealthiness constraints simultaneously, leaving the shortest time period for the system to counteract.

FDI Attacks

FDI attacks against state estimation and bad data detection are one of the hottest topics in the smart grid. It was first presented by Liu et al.[19;94] with DC system models in 2011. The authors assumed that the attacker knows the topology and network parameters of the entire power system and can manipulate the data measurements from the meters. An FDI attack can cheat the power system state estimation, which is the basis of many power system applications, such as contingency analysis, and economic dispatch[95;96]. Falsified state estimation results could potentially mislead the operation and auto-control mechanism of the EMS. The consequences of such attacks include economic loss, unstable system states, and even system voltage collapse[97]. Liang et al.[98] introduce an FDI attack that can induce physical line overflows. By considering the EMS sequential data processing functionalities, their optimized attack vector results in line overload when the false measurements cause generation re-dispatch. Elaborately constructed attack vectors can bypass bad data detection by keeping consistent with physical laws like Kirchhoff's circuit laws. The construction of the FDI attack vector a in DC models obeys (2.1):

$$a = H\dot{x} \tag{2.1}$$

where H is the measurement matrix; $\dot{x}$ is the estimated state deviation due to the attack; and $\hat{x}_{attack} = \hat{x} + \dot{x}$. Therefore, the malicious measurements $M_a = M + a$ will get the same BDD residual r as the original measurements M do.

Hug et al. further investigated the FDI attack in AC state estimation[99] in 2012. Unlike the DC model, where the elements in the measurement matrix H are constant, the relationship between the measurements and the states becomes non-linear in AC systems. The attack vector is derived as:

$$a = H(\hat{x} + \dot{x}) - H(\hat{x}) \tag{2.2}$$

where $\hat{x}$ is the estimated state; $\dot{x}$ is the change in the estimated state. The BDD residual under an AC-FDI attack is determined by the covariance matrix, the malicious measurements, and the estimated states after the attack. Since the attack vector is noiseless, the residual after the attack is not greater than the original residual; thus, the attack is hidden. Note that the construction of AC FDI attacks requires the estimated states, as shown in (2.2).

The state-of-the-art research on FDI attacks is on weakening the assumption that the attacker has the full knowledge of the system network information (i.e., H and $H(\bullet)$ are known to the attackers). However, the attacker has limited ability to hack into meters. In this case, the attacker can only access some specific measurements due to the different physical protection of the meters[100]. The limited access to meters leads to a subset of research works generating attack vectors by minimizing the number of manipulated measurements. For an attacker, minimizing the number of attacked meters, as shown in (2.3), can reduce the risk of being detected and the attack cost.

$$\alpha_k = \min_{x} \|Hx\|_0 \tag{2.3}$$

where α_k denotes the minimum objective value, $\|\bullet\|_0$ is the cardinality of a vector. Such a problem is proven to be NP-hard and non-convex; thus, it is often solved by mixed-integer linear programming (MILP) methods[101]. By exploiting the sparsity of H in the power system on account of physical topology, Sou et al.[101] propose a min-cut polynomial time approximate algorithm, which is faster but still as accurate as the MILP method. Wang et al.[102] simplify the original problem by solving the relaxed L1-norm problem for sparse attack construction. Due to recent studies, the L0-norm minimization can be relaxed to L1-norm minimization for sparse attack evaluation[103;104]. Recall that the construction of a perfect AC FDI attack requires the knowledge of estimated states. In reality, however, an adversary cannot obtain the same estimated state as the operators. To close the gap, Zhao et al.[105]

provide a sufficient condition for an imperfect FDI attack. By satisfying this condition, an imperfect attack vector can avoid being detected.

Blind FDI Attacks

Recently, FDI attacks with little to no information inspired researchers to construct blind FDI attacks without explicit knowledge of the power grid topology. Some researchers proved that such attacks exist and can further decrease the attack cost. In 2015, Kim et al.[106] presented the subspace method to learn the system operating subspace from measurements and launch attacks accordingly. Their subspace method did not require system parameter information and depended on partial sensor measurements. In 2015, Yu et al.[107] studied the problem of blind FDI attack, which makes inferences from the correlations of the line measurements. The construction of the attack utilizes the principal component analysis (PCA)[108] approximation method to transform the observation vector (a set of possibly correlated measurement variables M) into a set of linearly uncorrelated variables, $\tilde{x}$, called principal components. In the proposed attack model[107], attackers first collect some historical measurement data and run the PCA transformation. The PCA matrix, $H_{PCA} \in \mathbb{R}^{m \times n}$, is introduced by the dimensionality reduction of PCA, m is the number of measurements, and n is the number of principal components. The attacker can generate the stealthy blind FDI attack vector $a = H_{PCA}\dot{x}$ with an arbitrary $n \times 1$ non-zero vector $\dot{x}$. The attack is proven stealthy in the noiseless condition, and the noise will slightly degrade the performance of the attack.

In cases where attackers have the topology information needed, in 2012, Rahman and Mohsenian-Rad[109] proved that an attack could estimate H by collecting offline topology data manually (e.g., getting access to the grid topology maps through intruders or utility company employees), and online measurements data (deploying attacker's sensors and PMUs). Another approach exploits the relationship between the publicly available locational marginal prices (LMPs) and the Lagrange multipliers of the network-constrained economic dispatch. Thus, LMP components can unveil the topology information. In 2014, Kekatos et al.[110]

developed a regularized maximum likelihood estimator (MLE) to recover the grid Laplacian from the LMPs. A convex optimization problem was solved using an iterative alternating direction method of multipliers (ADMM) based algorithm. In the scenario where the loads vary within a small range, the topology information can be embedded into the correlations among power flow measurements. Esmalifalak et al.[111] propose an independent component analysis (ICA) algorithm to speculate the matrix H from power flow measurements. Higgins et al.[112] propose a data prepossessing before the ICA process. The proposed data classification is through T-distributed stochastic neighbor embedding (T-SNE) for dimensional reduction. Despite the above cases where attackers can obtain the topology information, attackers are also able to construct FDI attacks with limited topology information. Deng et al.[113] demonstrate that the adversary could launch unobservable FDI attacks to modify the state variable on a bus if they know the susceptance of every transmission line that is incident to that bus.

Meanwhile, attackers can launch effective and unidentifiable FDI attacks based on data-driven strategies[114]. Data-driven methods, especially machine learning-based approaches, are an essential branch of cyber-physical attacks on the smart grid. In 2019, Chen et al.[114] assumed an attacker who has little knowledge of the power system and is unable to estimate important parameters from observations. The attacker can only perform attacks and online learning iteratively to search for an optimal strategy. The optimal attack strategy was modeled as a partially observable Markov decision process (POMDP). Which, however, was impossible to be solved. Thus, the attacker could obtain an approximately optimal strategy through a Q-learning algorithm with the nearest sequence memories (NSM). In 2017, Markwood et al.[115] proposed a measurement matrix estimation attack, which was termed as a topology leaking attack. When the attacker knows the historical bus power injections and relative voltage phase angles, the measurement matrix H can be estimated. In cases where attackers can not distinguish the eavesdropped measurement corresponding to the current system topology, Higgins et al.[112] proposed an unsupervised learning method to cluster the data set via the density-based spatial clustering of application with noise (DBSCAN) algorithm in 2020.

Load Redistribution Attacks

In 2011, Yuan et al.[20] defined a special type of false data injection attacks, namely load redistribution attacks. By considering the characteristics of the power system and the attacker's capability, limited access to specific meters are available to LR attackers. Unlike original FDI attacks with a strong assumption that the attacker has access to all the meters in the system, LR attacks only manipulate the injection measurements of load buses and line power flow measurements. Centralized generator measurements and zero load bus injection measurements are not attackable. In other words, LR attacks are realistic false data injection attacks. In 2014, Liu et al.[116] proposed a local LR attack, which does not require the network parameter information of the whole system. They defined non-attacking regions, attacking regions, and boundary buses that connect these two types of regions. According to their research, an attacker, without knowing the network information of the entire power system, can launch a successful local load redistribution attack with only the knowledge of the network information (topology and line admittance) of the attacking region. This is done by keeping the same phase angle variations at all boundary buses.

Researchers have recently focused on revealing the specific attack consequences. In 2019, Che et al.[117] analyzed the mechanism that the attacker can implicitly identify the targeted initial contingency as a system weak point, then leverage such weak point to implement LR attacks to cause physical damage to the system. Under the impact of the load attack vector, the SCED enforces the line flow limits based on the incorrect power flow state. When the generators are following the dispatch commands sent from the SCED, severe transmission overloads can be caused[125]. In 2017, Xiang et al.[119] quantified the impact of LR attacks on long-term power supply reliability by proposing a power system reliability evaluation model. The proposed Monte Carlo simulation-based assessing method considers LR attacks that can cause load curtailment. In 2018, Fu et al.[120] presented an attacker who does not pursue a temporary profit but the most tripped lines during the cascading process by

Table 2.2: Measurement signal attack

Control signal attack	**Target**	**Objective**	**Means**	**Consequence**	**Mathematical expression**
AGC attack[90;92;93]	Automatic generation control	Rapid decline in the system frequency	ACE manipulation	Under-frequency load shedding	Measurement injection y_t^a
FDI attack[19;94–102]	State estimation based BDD	Incorrect estimated state	Measurement manipulation	CPSG functional failure	Measurement injection y_t^a
Blind FDI attack[106;107;109–111;113–115]	State estimation based BDD	Incorrect estimated state	Measurement manipulation	CPSG functional failure	Measurement injection y_t^a
LR attack[20;21;116–120]	State estimation based BDD	Incorrect estimated state	Realistic measurement manipulation	CPSG functional failure	Realistic measurement injection y_t^a
Topology attack[121;122]	Topology estimation	Incorrect topology estimation	Measurement manipulation	Incorrect topology state	Measurement injection y_t^a
Spoofing attack[123;124]	PMU	Manipulating PMU measurements	GPS signal manipulation	Incorrect location and time stamp	Measurement injection y_t^a

coordinating LR attacks with physical attacks. As the main cause of cascading failure is a physical attack, the system operator will always try to prevent cascading failure by re-dispatching the system back to a security operation point. This is when LR attacks come into play to disrupt and mislead the re-dispatching by causing maximum line overloading. Fu's case study showed that the LR-enhanced coordinated attack is more serious than a single physical attack causing cascading attacks. In 2020, Zhang et al.[21] extended the LR attack to AC distribution systems by presenting a net load redistribution attack, which aims at misleading the distribution system state estimation to observe illusory voltage violations. Measurements from prosumer buses with behind-the-meter distributed energy resources can be manipulated by an NLRA. In 2020, Choeum et al.[118] proposed an LR attack against the conservation voltage reduction (CVR) in distribution systems with DERs. The presented adversary injects malicious load data into the advanced metering infrastructure network and misleads the CVR to develop an abnormal control signal for the voltage regulator and smart inverter set points. The CVR results are consequently distorted, which causes an increase in active power flow from the substation.

Topology Attacks

In 2013, Kim et al.[121] proposed topology attacks in distinguishing from the FDI attack. The main difference between the topology attack and the FDI attack is that the topology attack manipulates the estimated topology state (switch and breaker states) instead of the estimated system state (power injection, power flow). A topology attack is achieved by manipulating both the meter measurement data and the network data, which can be represented as binary bits indicating the on and off states of various switches and line breakers. The attack vector in a DC model is shown in (2.4):

$$a = (\bar{H} - H)x \tag{2.4}$$

where H and $\bar{H}$ are the measurement matrices before and after the attack, respectively. When the measurement is noiseless, the system state x can be replaced with a function of measurements to generate the attack vector. However, the estimated state $\hat{x}$ is required when considering measurement noise.

The DC and AC attack vectors mentioned in this subsection require complete knowledge of network information to construct the measurement matrices and functions. In reality, this may not be possible. Therefore, a topology attack with local network information[121;126] has been studied. Kim et al.[121] consider a weak attacker who only has access to a few local meters. The authors propose line removal attacks, i.e., the adversary tries to remove lines from the actual network topology and mislead the operator that the line is disconnected. Liu et al.[126] observe the existing topology attacking model has two practical issues. The first issue is that there is no limit on the attacking amounts for load measurements at buses. The second issue is that attackers have limited capability to obtain necessary information. Thus, the authors propose a local topology attack model to determine the feasible attack region by obtaining less network information.

Table 2.3: Control-signal-measurement attacks

Control signal attack	**Target**	**Objective**	**Means**	**Consequence**	**Mathematical expression**
Line outage masking attack[127–131]	Topology estimation	Measurement manipulation to mask line-outage	Measurement manipulation	Voltage violation and line overflow	Measurement injection y_t^a
Stuxnet-Like Attack[132;133]	Communication channel	Incorrect control and measurement signal	Control signal and measurement manipulation	Stealthy malicious control commands	Control command injection u_t^a and measurement injection y_t^a

GPS Spoofing Attack

In CPSGs, spoofing attacks on PMUs are conducted via global position systems (GPS), in which the adversary produces artificial GPS signals. Two attack approaches, i.e., source ID mix attacks and time stamp attacks, are studied based on the spatio-temporal characterization of the GPS signals. A source ID mix attack is an attacker exchanging the location information of measurement data among different PMU channels without altering the measurement values. This attack places the measured data in the wrong positions in associated data servers. In 2019, Cui et al.[123] demonstrated the impact of source ID mix spoofing on the wide-area monitoring systems (WAMS) and the wide-area damping control. By swapping the signals of two buses, the WAMS estimated the disturbance at a location far away from the correct location; the damping control failed, and the system frequency kept dropping. The other type of GPS spoofing attack is called time stamp attack, also known as time synchronization attacks (TSAs), which aim to introduce erroneous time stamps maliciously, thereby inducing a wrong phase angle in the PMU measurements[122]. In 2019, Risbud et al.[124] formulated an optimization problem to identify the most vulnerable PMUs to construct a TSA. The vulnerability was quantified by the state estimation error, and a greedy algorithm was utilized to solve the problem.

2.2.4 Control Signal Measurement Attacks

Line Outage Masking Attacks

The recent attack on the Ukrainian power grid[134], which affected both the physical infrastructure and the situational awareness at the control center, is drawing more attention from researchers. A novel line outage masking attack is proposed[127–131], where an adversary attacks an area by physically disconnecting some lines from the attacked area (i.e., remotely open the circuit breakers) to occur short-term damage like voltage violation and line overflow, and then mask the measurements within the attacked area by DoS or FDI attacks. Such attacks combine both control and measurement layer attacks to cause immediate failure and block the operator's awareness at the same time, which may lead to cascading failures.

In 2017, Deng et al.[131] presented two coordinated cyber-physical attacks (CCPAs) to mask the line outage, namely replay and optimized CCPA. To construct the replay CCPA, attackers alter the meter readings on all the branches to force the active power flow measurements after the line outage to be the same as the power flow measurements from a normal state. The replay CCPA is extremely costly, and the actual system state is not consistent with the manipulated measurements, which makes it detectable by independently known-secure PMUs. The optimized CCPA neutralizes the impact of the line outage on the BDD residual. In 2019, Soltan et al.[127] proved that finding the set of line failures after data distortion and data replay masking attack is an NP-hard problem, based on the operator's knowledge of the phase angle measurement before and after the attack as well as the line admittance matrix. In 2016, Li et al.[135;136] proposed to conduct two-step cyberattacks that mask line outages resulting from the physical attacks. The cyberattacks are decomposed into two steps, which include a topology-preserving attack as the first step, followed by the load redistribution attack (if the first step is not feasible). More specifically, the topology attack masks line outages caused by physical attacks, while the load redistribution attack keeps the total load unchanged and redistributes the line flow to bypass the state estimation-based detection. In 2019, Chung et al.[130] further improved the masking approach by deploying a line-removing FDI attack (topology attack) that misled the SCADA system with a fake outage in another position. After the real line outage attack, the topology attack region is then selected to re-dispatch the power flow. The attack vector is generated in an AC model with local network information and the capability to manipulate the measurement within the attacked area.

Stuxnet-Like Attacks

Traditional Stuxnet attacks inject malicious control commands into the actuators and, meanwhile, corrupt the sensor readings to cover the ongoing attack. To avoid being detected, Stuxnet attacks require the attacker's capabilities to replay all the measurements during the steady state of the system. Forensic analysis of Stuxnet attacks[132] has shown the feasibility of a very targeted and highly sophisticated cyberattack. Moreover, with some modifications, Stuxnet can be tailored as a platform for targeting other systems e.g., automobile or power systems.

In 2019, Tian et al.[133] defined Stuxnet-like attacks against secondary voltage control, which assume the attacker has write access to both the control signal and sensor measurement. The cyber-physical system dynamic is described as a discrete-time linear time-invariant (LTI) model. In the presence of an attack, the system dynamics are as follows:

$$x_a(t+1) = Ax_a(t) + Bu_a(t) + w(t) \tag{2.5}$$

$$y_a(t) = Cx_a(t) + v(t) \tag{2.6}$$

where the notations are similar to those in (1.1) and (1.2) with an exception that the subscript a denotes the under attack status. The attacker knows the state transit matrix A, the control matrix B, and the measurement matrix C. Variable u_a is the contaminated control signal received by the actuators; y_a is the manipulated sensor measurement received by the control center; x_a denotes the system state. Functions $w(t)$ and $v(t)$ denote the process and sensor noises, respectively. This Stuxnet-like attack is only implemented on a converged system, where the control center expects unchanged system states. The attacker needs to judge whether the system has converged, according to the eavesdropped control signal and measurement data.

2.3 Cyber-Physical Defense

Cyber-physical defense is absolutely the focus of ongoing research efforts, where a massive number of works have already been published in the literature. This section first categorizes cyber-physical defense approaches into temporally-relevant and spatially-relevant approaches. Further, several state-of-the-art cyber-physical defense approaches in the CPSG, including securing measurement sensors, model and algorithmic enhancement, data-driven approaches, MTD, and watermarking, are reviewed.

2.3.1 Temporally- and Spatially-relevant Detection

In a temporally-relevant detection, the current system state is estimated by prior estimated state, measurement, and control signal. At time t, the estimated measurement $\hat{y}\left(t\right)$ and the residual $\delta\left(t\right)$ are shown as:

$$\hat{y}\left(t\right) = L_1\left(\hat{X}\left(t-1\right), U\left(t-1\right), Y\left(t-1\right)\right) \tag{2.7}$$

$$\delta\left(t\right) \overset{\Delta}{=} y\left(t\right) - \hat{y}\left(t\right) \tag{2.8}$$

where $L_1\left(\bullet\right)$ is an abstract function; $\hat{X}\left(t-1\right) = \left[\hat{x}\left(t-1\right)\cdots\hat{x}\left(0\right)\right] \in \mathbb{R}^{n\times t}$ is the set of the prior estimated state; $U\left(t-1\right) = \left[u\left(t-1\right)\cdots u\left(0\right)\right] \in \mathbb{R}^{l\times t}$; $Y\left(t-1\right) = \left[y\left(t-1\right)\cdots y\left(0\right)\right] \in \mathbb{R}^{m\times t}$. After the estimation, if the calculated residual is larger than a pre-defined threshold, the detection method will signal an alert. Among all temporally-relevant approaches, the most widely used method is the Kalman filter based state estimator and the chi-squared test[45;137;138]. The Kalman filter based estimator minimizes the variance of the estimated state, given the previous observations. The chi-squared test[139] is commonly used to detect anomalies.

The spatially-relevant detection method estimates the system by the correlation between different sensors in one time-interval only. A power system state estimator and the residual-based BDD is an example of the spatially-relevant detection approach. An essential of this estimation is measurement filtering, which utilizes the measurement data redundancy to increase the measurement accuracy. At time t, the estimated system state is calculated based on the measurement from the same time interval,

$$\hat{x}\left(t\right) = L_2\left(y\left(t\right)\right) \tag{2.9}$$

where $L_2\left(\bullet\right)$ is an abstract function. From equation (1.2), the estimated measurement is shown as:

$$\hat{y}\left(t\right) = C\left(\hat{x}\right). \tag{2.10}$$

The residual-based alarm mechanism is also implemented in spatial-relevance detection. One notable difference is that in a temporally-relevant detection, the estimated measurement is calculated from prior system state (2.7); however, in a spatially-relevant detection, the estimation is based on the current state (2.10).

2.3.2 Securing Measurement Sensors

As previously mentioned, the majority of attacks require, more or less, the attacker's knowledge about the system control and measurement signal. An assessment in[72] has shown that the major cybersecurity concerns range from exploiting well-known protocols to the leakage of confidential information. Therefore, one natural approach is to select and protect critical control or measurement signal strategically.

In 2010, Bobba et al.[96] explored the detection of false data injection by protecting a set of critical sensor measurements and a method to verify the values of strategically selected state variables. The authors demonstrated that an attack aims to construct an attack vector such that it avoids specific measurements and state variables that are protected and verified. From the defender's perspective, the operator should select the sets of protected measurements and the verified state to ensure that an adversary cannot find a stealthy attack vector. Thus, FDI attacks could always be detected. The trade-off here is that the protection and verification of a large number of measurements and state variables could be costly.

PMUs have recently attracted researchers' attention due to their ability to provide measurement redundancy and assist in FDI detection. In 2018, Zhao et al.[105] developed a robust FDI attack detection method by checking the statistical consistency of measurements from a limited number of secured PMUs. In the proposed detector, short-term measurement forecasting[140] was advocated to enhance the PMU data redundancy in 2013. In 2011, Giani et al.[141] proposed that it is sufficient to place $p+1$ known secure PMUs at carefully chosen buses to neutralize a collection of p cyberattacks. Since then, the optimal PMU placement has been researched to detect the stealthy FDI attacks with the least PMUs. In 2015, Qi et al.[142] formulated the optimal PMU placement as an optimization problem, which maximizes the determinant of the empirical observability Gramian matrix. In 2017, Pal et al.[143] presented an integer linear programming methodology for the PMUs placement scheme while considering realistic cost and practical constraints. In 2018, Sarailoo et al.[144] adopted synchrophasor availability (SA) on all buses as a constraint and then minimized the number of PMUs. The SA is the fraction of time on average the bus voltage synchrophasor is correctly present. As mentioned in Section III, the synchronization between PMUs requires GPS signals, which are vulnerable and can be attacked[145–147]. In 2015, Fan et al.[148] proposed a cross-layer detection against simultaneous GPS spoofing attacks toward multiple PMUs.

2.3.3 Modeling and Algorithmic Enhancement

Another category of defense approaches is the improvement of detection models and algorithms. In 2011, Huang et al.[149] proposed an adaptive cumulative sum (CUSUM) algorithm, which detects the adversary fast while maintaining a low detection error rate. In 2014, Liu et al.[150] proposed a false data detection mechanism that utilized the intrinsically low-dimensional power grid measurements and the sparse nature of FDI attacks. The detection problem is formulated as a matrix separation problem and is solved by two methods: nuclear norm minimization and low-rank matrix factorization. In 2015, Gu et al.[151] proposed a detection method to detect FDI attacks by tracking the dynamics of measurement variations. They utilized the Kullback-Leibler distance (KLD) to calculate the distance between two probability distributions, i.e., historical measurements and suspicious measurements, to detect the FDI attacks. In 2017, Zhao et al.[152] proposed a short-term state forecasting method considering the temporal correlation to calculate the approximate prior system measurements. The consistency between the forecasted and received measurements is checked by a statistics-based test method. From the consistency test result, a detection metric is constructed by the infinity and the L_2-norm-based measurement residual analysis. In 2018, Ashok et al.[153] showed that the existing CPS defense focuses on either redundant measurements or the cybersecurity of sensors and communication channels. These offline approaches make specific assumptions about the attacks and systems, which are restrictive. One solution of PMUs placement or security mechanism may no longer be adequate under another system configuration. Therefore, the author proposed an online anomaly detection that covers broad attack scenarios. The proposed method leverages online information obtained from load forecasts, generation schedules, and real-time data from PMUs to detect anomaly measurements.

2.3.4 Data-driven Approaches

Another noteworthy category of defense approaches is data-driven machine learning methods that have been gaining traction due to the following two salient advantages:

1. The construction of the data-driven approaches does not depend on the network topology; and

2. This approach is usually sensitive to time-variance measurement, which can be very effective in detecting one time interval stealthy FDI attacks created based on the spatial-relationship of CPSGs.

The use of supervised learning classifiers as alternate FDI detectors was proposed by Ozay et al.[154] in 2015. Supervised machine learning-based binary-classifiers were presented to check the distance between ”secured” and ”attacked” measurements. With the distance information, attacks can be recognized by the learning algorithms. In 2016, Yan et al.[155] proposed to implement the learning based false data classifiers as a secondary detector after the residual-based BDD. They designed FDI detectors with three widely used supervised learning based classifiers, including support vector machine, k-nearest neighbor, and extended nearest neighbor. The proposed detectors are capable of detecting stealthy FDI attacks that can bypass the residual-based BDD. In 2019, Sakhnini et al.[156] tested three classification techniques with different heuristic feature selection techniques. The authors concluded that the support vector machine and the k-nearest neighbor algorithms could get better accuracy than the artificial neural network. However, the artificial neural network is expected to have better performance on larger systems at a higher computational cost. The recent breakthrough in computing provides the foundation for ”deep” neural networks. In 2019, Niu et al.[157] developed a smart grid anomaly detection framework based on a neural network. The recurrent neural network with a long short-term memory cell is deployed to capture the dynamic behavior of power systems. According to the captured behavior, the estimated measurements are calculated and compared with the observed measurements. If the residual between the observed and the estimated measurements is greater than a given threshold, an attack is detected.

As for reinforcement learning based methods, Chen et al.[114] proposed a BDD method based on Kernel density estimation in 2019. By using historical records, the measurements can be estimated. The effectiveness of the proposed detection method relies on the abundance of integrated records of normal operations of the power grid. When an attack vector is injected consistently, the tempered measurements could be used for the Kernel density estimation analysis. Thus, the proposed BDD detection method could fail. Other than the studies that contribute to attack detection, Li et al.[158] proposed a defense methodology that recovers the real measurements to maintain uninterrupted state estimation under FDI attacks in 2020. The proposed method utilized a generative adversarial network based data model which captures the deviations from ideal measurements and then generates correct data to replace the manipulated data. Besides the aforementioned defense approaches that protect the transferred measurement data, the defense on the communication channel is vital. One of the cutting-edge wireless communication technologies used in the smart grid is the cognitive radio, which is motivated by the ever-increasing demand for high data rates in the face of limited spectral resources. In 2013, Ding et al.[159] introduced a spectrum attacker who can inject attack data into the honest spectrum sensor to mislead the fusion center to lower the spectrum utilization. Moreover, the authors show that the kernel K-means clustering (KMC) algorithm yields better performance than the KMC algorithm in the detection of spectrum attacks. However, high-quality clean training data are too expensive or too difficult to obtain in some cases. In 2016, Xie et al.[160] proposed a convex framework to provide robust classification and training in improving the anomaly-resistant against sensor failures (i.e., falsified channel sensing resulting in Byzantine attacks) in which possibly anomalous samples occur in the training set. In 2017, Qin et al.[161] proposed a low-rank matrix completion based malicious user detection framework for the secure cooperative spectrum sensing with a lower data acquisition cost.

2.3.5 Moving Target Defense

The aforementioned operational defense approach is either computationally complex or somewhat passive. As an emerging technique, MTD, is originally proposed to enhance network security[162]. It proactively changes the system configuration so that it reduces the attack surface and increases the uncertainty about the network system. With the properly arranged MTD perturbation, the attacker's knowledge about the system is always outdated. This approach increases the barriers for the attackers to launch stealthy attacks. MTD has recently been introduced in the physical layer of the cyber-physical power system (CPPS) to provide proactive defense, which is an advantage over the traditional remedial defense. Comparing with the MTD in the cyber-layer network system, MTD in CPPS is very complex as it requires the physical dispatch of control, measurements, or device properties.

In 2012, the concept of MTD was first introduced into the physical layer of the power system by Morrow et al.[163] and Davis et al.[164]. In general, MTD utilizes D-FACTS devices to actively modify impedance perturbations to invalidate attackers' knowledge about the power system configurations, which is essential for constructing stealthy attacks. Table 2.4 summarizes the existing works on MTD, where the superscript "AC" or "DC" indicates the corresponding AC or DC model used.

There are two essential steps in the construction of an MTD, namely MTD planning and MTD operation. First, in the MTD planning, a utility needs to install D-FACTS devices on an appropriately identified subset of transmission lines, namely solving the problem of D-FACTS placement. Arbitrary placement and full placement are the two simplest D-FACTS placement strategies. Arbitrary placement randomly selects a subset of lines to install D-FACTS devices[165]. Full placement is the most expensive method in which D-FACTS devices are installed on every transmission line[32]. However, the detection effectiveness of MTDs under these two placements is not considered. Max-rank placement[31;166] can make MTDs achieve the maximum rank of the composite matrix (i.e., max-rank MTDs), a metric of the detection effectiveness. Spanning-tree placement proposed in[167] installs D-FACTS devices on the lines which form a spanning tree of the system. MTDs under spanning-tree placement is effective to detect single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FDI attacks.

After the allocation of D-FACTS devices, the system operator/defender needs to continuously determine the D-FACTS setpoints under different load conditions in the MTD operation. The MTD operation includes four methods. First, random selection is the simplest operation method without any computational overhead, in which the D-FACTS setpoints are randomly perturbed[165]. As D-FACTS devices are originally used to control the power flow, OPF-based operation methods integrate the D-FACTS devices into the optimal power flow model to minimize the system losses or generation costs[31;168;169]. Neither the random selection method nor OPF-based operation methods consider the detection effectiveness.

Table 2.4: Moving target defense in CPSG

MTD Algorithm	**MTD planning**	**MTD operation**	**Characteristics**
Random MTD[165]**DC**	Arbitrary placement	Random selection	Detection effectiveness is not considered
OPF-based MTD[168]**DC**, [169]**AC**	N/A	OPF-based operation	Minimize generation cost and guarantee detection effectiveness[168]
Hidden MTD[170]**DC**, [41]**AC**, [171]**DC**	Placement enumeration[170]; max-rank placement using protected meters[171]	Random selection subject to hidden condition[170]	MTD has max-rank and is hidden to alert attacker, but[171] uses extra protected sensors
Spanning-tree MTD[167]**DC**	Spanning-tree placement	Random selection	Covers all buses, but max-rank MTD is not ensured
Max-rank MTD[32]**DC**, [166]**DC**, [31]**DC, AC**	Full placement[32], max-rank placement[31;166]	Optimization-based operation[32]; ACOPF-based operation[31]	Minimizes system losses[32] or generation costs[31]. Guarantees max-rank MTD based on numerical methods[32;166] or graph-theory methods[31]

Thus, these two methods must be constructed in the D-FACTS placements, which ensure the detection effectiveness, such as the max-rank placements. Second, the optimization-based operation takes both the economic cost and the detection effectiveness into account, in which the metric of detection effectiveness is maximized or taken as constraints[32;166]. Finally, the hidden MTD operation method delicately selects D-FACTS setpoints such that all measurements remain the same after MTD is applied[41;170;171]. In this case, vigilant attackers cannot detect the MTD in place using BDD. To find suitable placement for hidden MTD operation, authors in[170] enumerate all placement combinations, while authirs in[171] use the max-rank placement in[166], with the help of protected meters.

In the literature, there are three important concerns in evaluating the performance of MTD. First, attack detection effectiveness is the most important metric for a defense algorithm. As not all MTDs are effective in detecting FDI attacks, the feasibility and the limitation of MTD are discussed in[167]. Many works focus on improving the attack detection effectiveness of MTDs though the MTD planning[31;166;167;171] and MTD operation[32;168;170]. Two metrics are proposed to measure the detection effectiveness of MTD, namely the Lebesgue measure[168] and the rank of the composite matrix[31;32;166;170]. The composite matrix rank is superior to the Lebesgue measure in the evaluation of MTD detection effectiveness since it demonstrates the inherent nature of MTD on FDI attack detection and provides an explicit objective for constructing an effective MTD. Authors in[31] proved the rank of the composite matrix could be merely determined by D-FACTS placement, as long as no D-FACTS devices work in idle states. In addition, the number of buses covered by D-FACTS devices and the incremental line reactance introduced by D-FACTS devices also impact the MTD detection effectiveness[167]. However, there is no metric proposed to measure this impact.

Second, the cost of the MTD application is a must concern for a utility. The cost consists of the planning cost and the operation cost. In the planning cost, the number of D-FACTS devices used in MTD determines the capital cost and labor fee. Max-rank placement in[31] uses the minimum number of D-FACTS devices to achieve the maximum rank of the composite matrix. In the operation cost, the D-FACTS setpoints impact the generation cost and system losses, as these setpoints can change power flow in the system. Thus, OPF-based operation methods can be used to reduce the MTD operation cost in both AC and DC models. To integrate the OPF-based operation methods into the EMS, an interior-point solver proposed in[169] can solve these methods within seconds.

Third, the hiddenness of MTDs provides a superior function, making the MTD stealthy to attackers. Vigilant attackers use BDD to detect the existence of MTD before launching any attacks. If attackers detect any MTD in place, they may stop FDI attacks and invest more resources to launch data exfiltration attacks to obtain the latest system configuration[170]. Hidden MTDs can mislead these attackers to launch detectable attacks based on incorrect line parameters. In summary, a desirable MTD would be a hidden MTD with maximal detection effectiveness and low cost.

2.3.6 Watermarking

Watermarking is originally used to identify the ownership of noise-tolerant signals such as audio, video, or image data. It also can be used to check the integrity and authenticity of a signal. The first use of watermarking to defend against the replay attack employed in Stuxnet was introduced by[44;137] in 2009, where the physical watermarking as a control-theoretic method to authenticate the correct control operation was proposed. Although existing tools like cryptography can provide authentication, physical watermarking is more effective against physical attacks or insiders who are usually authenticated users. The concept is that by injecting a known noise as a probe input of the system, an expected effect of such input should be found in the actual measurement output due to the system dynamics. Thus, if the attacker is unaware of the watermarking, the injected attack will be detected by a chi-squared detector. In 2014, Weerakkody et al.[45] considered a more adversarial attacker with access to a subset of real-time control and sensing signals. The physical watermarking approach is extended to show the ability to counter a more intelligent adversary. Since introducing a random probe signal into the system could affect the operating cost, Miao et al.[172] proposed an optimization method for the trade-off between cost-centric and security-centric controllers in 2013. Despite the detection capability, the physical watermarking needs to inject perturbation as a probe into the system, which may affect the system's performance. Moreover, the physical watermarking detection sensitivity is usually related to the probe signal magnitude. Thus, to increase the detection performance, the defender has to sacrifice the optimal system performance.

In 2016, Satchidanandan et al.[173] extended the physical watermarking to dynamic watermarking in a noisy dynamical system. The authors introduced independent and identically distributed random variables to actuator nodes, namely privately imposed excitation. The realization of the time-sequence excitation is superimposed on the control input from an honest actuator. The author assumed that the control policy is in place, and the excitation is only known by the honest actuator itself. The proposed dynamic watermarking can ensure that a malicious sensor is constrained to distorting the process noise by at most a zero-power signal by implementing the correlation detector. In 2018, Ferdowsi et al.[174] proposed a deep learning framework for the dynamic watermarking of IoT signals. The framework is based on the long short-term memory blocks to extract stochastic features from IoT signals and watermarks the features inside the original signal. This dynamic extraction enables eavesdropping attack detection since the attacker cannot extract the watermarked information.

Watermarking can also be used for attack identification in CPSG. In 2018, Liu et al.[32] designed a reactance perturbation-based scheme to identify originally covert FDI attacks on power system state estimation. The term originally covert attack refers to the stealth of the attack prior to reactance perturbation. The authors proved that the originally covert attack (constructed with the original measurement matrix H_0) is detectable and identifiable in a reactance perturbation with a new measurement matrix H if and only if the rank of $[H_0 \; H]$ is equal to $2(n-1)$, where n is the number of buses. In 2020, Zhang et al.[175] proposed an attack identification approach for GPS spoofing attacks (GSAs) against PMUs. They performed a parallel probing technique on each PMU to determine the locations of spoofed PMUs and the ranges of GSA phase shifts under the assumption that the PMU in a substation is secure.

The attack models and the defense mechanisms surveyed in this chapter are summarized in Fig. 2.1. The two-layer model in this figure is a graphic form of the CPSG model abstracted in Section 1.2. In Fig. 2.1, each attack on smart grid functionalities is shown with corresponding counter-measurements labeled next to it.

2.4 Opportunities and Challenges

Despite the tremendous research efforts reviewed in this work, cyber-physical security challenges remain to be thoroughly addressed. Critical power system functionalities such as market operation, advanced metering, and network operation may also face attacks. Mean-

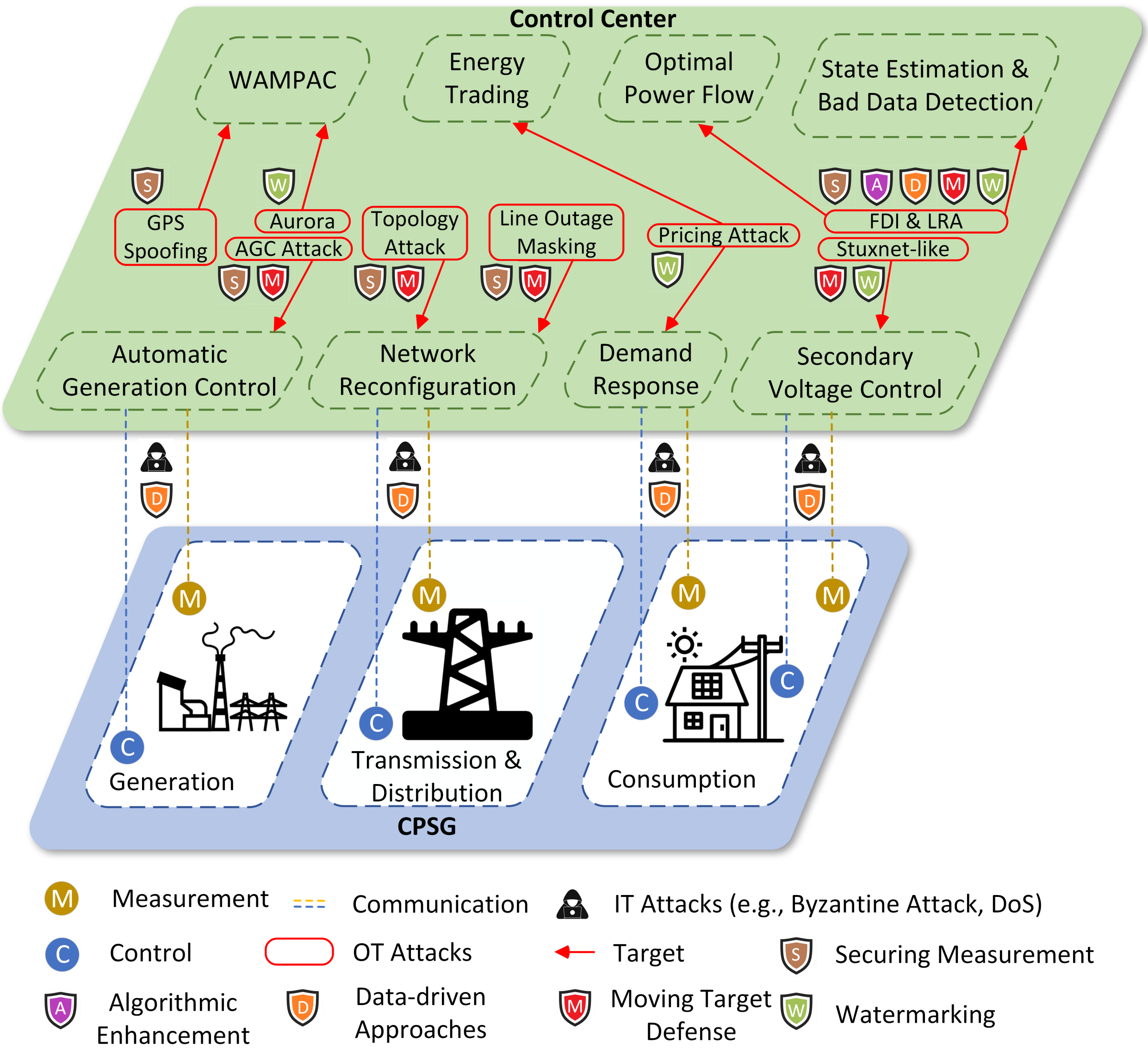


Figure 2.1: Infographic of attack and defense mechanisms in smart grid.

while, the potential implications of these attacks remain to be further investigated. In addition, the emerging applications, including time of use, demand response, and large-scale electric vehicles, will have strong impacts on the smart grid and may also become targets of cyber-physical attacks in the future. This section highlights four critical challenges and opportunities in the field of smart grid cyber-physical security that deserve further research efforts.

2.4.1 Cyber-physical Security in Distribution Systems on The Grid Edge

A CPSG is a critical infrastructure with an enormous number of complicated devices. Cyber-physical attack and defense simulations are necessary to estimate their performance, though it is impossible to implement most experiments on a real-world power grid. However, the existing cyber-physical studies focus primarily on transmission systems, while the work on the three-phase unbalanced distribution systems with low system observability is significantly under-researched. A growing number of distribution systems on the grid edge are experiencing significant penetration of DERs. The emerging power-electronic-device-based electric vehicles, local energy storage, and demand-response have also contributed to the system dynamics and complexity. Fully taking into account the new dynamics and complexity in low-observability distribution systems is quite challenging in the context of cyber-physical security. More research efforts are therefore necessary in distribution systems on the grid edge.

On the other hand, conventional, discrete-time, model-based simulations are accepted by researchers[176]. However, the traditional power system simulation tools may not be suitable for studying the distribution grid with increasing complexity and cyber-physical concerns. There has been a growing need to use continuous-time simulation with hardware in the loop (HIL) capabilities. In[177], the authors develop a SCADA security testbed, which integrates a real-time immersive network simulation environment with PowerWorld. The authors in[178] develop a testbed with PowerWorld and OPNET. A platform equipped with GridLAB-D and NetSim has been used for power systems and communication network simulation in[179]. Due to time-domain analysis complexity, these simulation platforms cannot run in real-time or perform HIL simulation. In a broader sense, the real-time simulation reflects the exact dynamic behavior of a CPSG, and the HIL ensures precise operation as the real devices. While these two functionalities are usually unavailable with the current simulation structure, development on the real-time simulation testbed with HIL largely remain to be conducted.

2.4.2 Interdependence

Studying the CPSG security issues relies on the interdependence of both the cyber layer and the physical layer. The attack detection requires advanced communication technologies to transfer data from the physical devices to the control center. On the other hand, most cyber-physical attack schemes have taken advantage of this interdependence to launch attacks in the cyber layer and induce physical damages. For future research in this area, cyber-physical interdependence needs to be comprehensively explored. For instance, the physical attacks on cybersecurity have been under-investigated, and the threats can be devastating when the dependence of physical systems is exploited by an attacker[2]. Another cyber-physical interdependence that has been largely ignored is simulation software. Traditional software is developed to simulate or emulate either communication networks (e.g., OPNET, NS2, OMNET) or physical power systems (e.g., RTDS, DSATools, PSS/E, PowerWorld). Such software cannot provide realistic cyber-physical environments[180]. Additionally, the interdependence between CPSGs and other critical infrastructures, such as communication, water, and transportation networks, ought to be researched in the context of cyber-physical attacks against CPSGs.

2.4.3 Attack Coordination

In real-word CPSG, sequential outages are the most common causes of blackouts[181], e.g., the 2003 Northeast Blackout[182] and the 2011 Southwest Blackout[183]. If a series of attacks can trigger such events, then an intimidating cyber-physical security risk will be worthy of attention. Section 2.2 discussed the line outage masking attack, one of the popular methods among coordinated attacks. Meanwhile, most researchers assume that the cyberattack vector is injected simultaneously with the physical damage in the existing research. This assumption may be validated in a specific condition, such as the system is in a steady state. However, the general circumstances in which the attackers cannot promise timely cyberattack injection with respect to the system dynamic have remained to be considered. However, the timing and ordering of coordinated attacks can also have an impact on the eventual damages. With an elaborate schedule, not only will concurrence be relaxed, but the damage may be amplified. On the other hand, from a defender's perspective, analyzing the coordinated attacks on CPSG based on temporal-topological correlation can help to restore the complete attack path and identify the intent of the attacks[184].

2.4.4 Attack Identification and Mitigation

In future power systems, an attack detector will be an indispensable tool for detecting and identifying anomalous measurements. Without reliable attack identification, it is hard to implement a mitigation process with pertinence. While detecting attacks is computationally straightforward, identifying the attack location and strategy is computationally challenging[185]. For instance, bad data cannot be identified once belonging to the critical sets of measurements, also known as bad data groups, because they cause the same normalized residuals for each element of the set[186]. Another problem is that existing state estimation based algorithms in transmission systems are not suitable for unbalanced distribution systems with high r/x ratios[187]. With the aforementioned issues, few solutions have been proposed for the identification of attacks. In addition, rather than brutally getting rid of identified compromised measurements, how best to mitigate the adverse effect of those attacks is also a very challenging issue depending on particular operation and controls of a CPSG.

2.5 Conclusion

A CPSG relies on the cooperation of both cyber and physical layer functionalities. The ubiquitous threat to the entire smart grid's large attack surface makes it necessary to comprehensively analyze and classify attacks. This chapter provides a CPPS operation model and addresses the associated vulnerabilities targeted by an attacker. It also classifies the existing attack approaches against different components based on the CPPS model. A review of the cutting-edge operational defense approaches is presented to summarize and categorize the state-of-the-art in the field, ranging from the state estimation based detector to the emerging MTD and watermarking methods. As smart grid technologies become more prevalent and more physical devices are connected to the cyber-physical infrastructures, significant attack surfaces are introduced, as well as a wide range of opportunities and challenges. Four challenges were highlighted in the investigation of smart grid cyber-physical security. This survey provides insights that future research efforts must target a new set of cyber-physical security concerns, including real-time risk modeling and simulation, risk mitigation, and coordinated attack defense.

Chapter 3

Net Load Redistribution Attacks

After reviewing the state-of-the-art cyber-physical security research in smart grids, it is seen that researchers are expected to study the power system cybersecurity from both the attacker's and the defender's perspectives. This chapter proposes a realistic FDI attack model, namely NLRA, which releases two strong assumptions in the existing FDI strategies. Furthermore, a modified AC-OPF problem is constructed to maximize the attack impact while keeping the proposed attack stealthy to the system operator's BDD. To obtain the required system state, WLS-based and machine learning-based state estimators are proposed to be implemented by attackers.

3.1 Introduction

Cyber data attacks are viewed as "the worst interacting bad data injected by an adversary"[188]. In[67], Liu et al. analyze several existing security accidents and introduce the taxonomy of the attacks according to their spatial-temporal characteristics. An attacker with the capability of configuration information can manipulate the measurement data at the smart meters as they are usually physically exposed[4]. Such attacks are defined as FDI[19;94;188]. FDI attacks can result in incorrect state estimation and further undermine the economic and secure operation of power systems. In previous research, it is assumed attackers have the entire power network information. In reality, this is an impractical assumption due to the security and complexity of today's power grid. Therefore, FDI attacks with incomplete information[51;107;109;116;126] are drawing more research attention. Liu et al.[116] demonstrate an attacker could construct an undetectable FDI attack in an AC transmission system with incomplete network information by maintaining the same phase angle increment at the boundary nodes of the attack region. According to the characteristics of transmission systems (i.e., high X/R ratio, meshed network), Liu et al. propose a method to approximately estimate phase angle differences between boundary nodes[51]. Their results show the construction of an FDI attack does not require knowledge of the entire power network. Yuan et al. develop a novel concept of load redistribution attacks[20], which is a more realistic form of the FDI attack in the DC transmission system. To the best of the author's knowledge, the LR attack has not been researched in AC distribution systems with distributed energy resources, wherein the malicious measurement may be disguised by the uncertain DERs' power injection. Existing research on the effect of LR attacks mainly concerns economic consequences. However, little research has been conducted when attackers aim at creating system state violations.

An important reason for this gap is that constructing LR attacks in AC distribution systems remains challenging, as attackers only have limited resources. Even though local FDI attacks[51;116;126] are proposed to reduce the attacker's required knowledge about system configurations, the lack of accurate system state from an attacker's perspective excludes them from being used in the real world. Therefore, this chapter first proposes to use boundary conditions to ensure the stealthiness of the proposed NLRA so that only part of the system configuration information is required for an attacker. Then it is proposed to use power flow enhanced deep learning SE to provide attackers with the necessary system state because attackers usually do not have redundant measurements to implement WLS-based SE.

In summary, this chapter fills the gap by answering the following questions:

- In NLRA construction, how do attackers maximize the attack impact? How to construct NLRA with incomplete knowledge about system configuration?

- How do attackers get accurate system states without enough redundant measurements to implement WLS-based SE?

The first question is solved in the NLRA model section, and the second question is solved in the PFEDL-based SE section.

3.2 NLRA Model

Cyberattacks on a certain type of power system measurements can easily expose themselves. For example, a control center can straightforwardly detect a cyberattack on the measurements of a utility-scale wind or solar farm through direct communication between the system control center and the generating resource control room[189]. In the NLRA model, a generator bus with utility-scale generators is not attackable. A nodal net load, calculated as the total load minus the total local generation, is measured at a specific node in the power system. Nodal net load measurements would become highly uncertain with a greater amount of behind-the-meter DERs in the distribution system, giving rise to cyberattacks as the attacker can disguise an attack vector as uncertainties.

Before introducing the attack model, this section defines the attack region and non-attack region for clarity. As shown in Figure 3.1, a connected power network is separated into the attack region in the left ellipse and the non-attack region in the right ellipse by a set of tie lines. The buses on each end of a tie line are called boundary buses. The attack region consists of all the buses whose measurements can be manipulated by the attacker, excluding the boundary buses. The non-attack region consists of all the buses that are outside the attack region.

This section investigates a stealthy FDI attack model, termed as NLRA, in which measurements on the net power injection at a load node and related line power flow measurements can be compromised. With some mild assumptions on the attackers' capability, attackers can precisely control the errors injected into these measurements (attack vector) in a coordinated manner to mislead the estimation of nodal voltage magnitudes in the attack region.

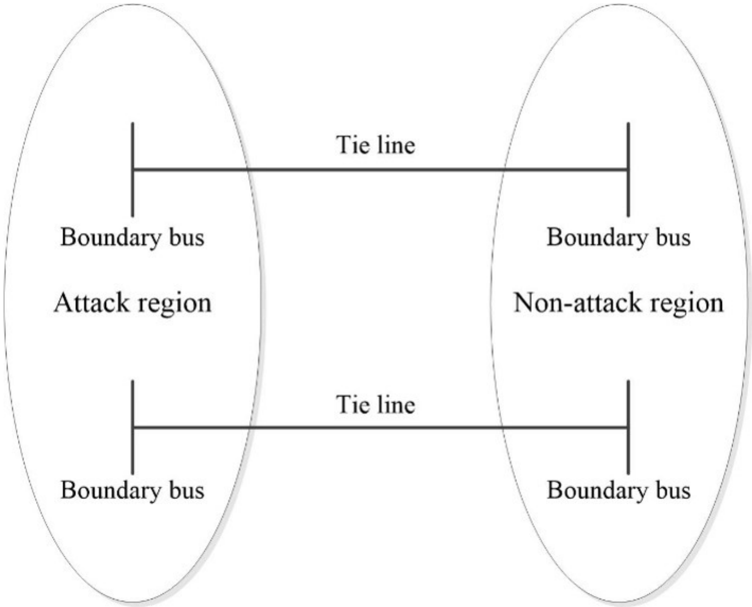


Figure 3.1: Illustration of NLRA attack and non-attack region.

The attackers must also maintain the sum of all net power injection measurements in the attack region unchanged to make the attack stealthy.

3.2.1 Stealthy Conditions

To launch a stealthy AC FDI attack with incomplete network information, two conditions need to be satisfied. First, the attack vector itself should bypass the BDD tests. Liu et al.[19] proved that if an attacker injects false data that are consistent with the physical characteristics of power systems into the measurements, the attack vector can bypass the BDD. Second, the injected attack vector should not cause any state or measurement changes outside of the attack region since the attacker does not have access to the non-attack region. Liu et al.[116] demonstrated that if an attack vector ensures that all boundary nodes between the attack and non-attack regions have the same incremental phase angle, the additional power flow due to the injected false power will not flow out of the attack region. Therefore, the power system states and measurements in the non-attack region would remain unchanged.

Recently, an FDI attack strategy against AC SE with incomplete network information is proposed in transmission systems[51]. The authors utilize the high X/R characteristic in transmission systems to approximate the necessary voltage angles on boundary buses to ensure the above two stealthy conditions. However, this approximation is not accurate in distribution systems due to the low X/R ratio. Meanwhile, attackers usually cannot obtain enough redundant measurements to implement WLS-based SE. To address this issue, this dissertation proposes to use PFEDL-based SE as an attacker to obtain accurate system states with a small amount of measurements. The detailed PFEDL-based SE is introduced in Section 3.3

3.2.2 Attacker's Capabilities

To meet the two conditions discussed in Section 3.2.1 and launch a stealthy NLRA attack, an attacker must have the following capabilities:

1) Knowledge of line impedance in the attack region. In reality, line impedance may not be directly accessible to attackers. They need to launch data exfiltration attacks to obtain the line impedance. Several methodologies[190–193] have been proposed to estimate the line impedance;

2) Read & write access to power injection and line flow measurements in the attack region. The attacker can eavesdrop on those measurements and perform man-in-the-middle attacks; and

3) Knowledge of the voltage magnitude and angles at boundary nodes, as well as read access to the tie-line power flow between the attack and non-attack regions.

The attacker's capabilities required to launch a stealthy NLRA are summarized in Table 3.1. Note that the attacker's cost (e.g., resources invested) to launch a successful NLRA is highly related to the scale of an attack region. For an attacker, attacking a large region requires a higher cost than a small region, but may result in more severe consequences in the distribution system.

Table 3.1: Attacker's capabilities for NLRA

Measurements	**Capabilities**
Line impedance	Attack region (Knowledge)
Power injection	Attack region (Read & Write)
Line power flow	Attack region (Read & Write), Tie lines (Read-only)
Voltage magnitude	Boundary nodes (Read-only)
Voltage angle	Boundary nodes (Read-only)

3.2.3 Attack Objectives

With the above attacker's capabilities, the proposed NLRA is modeled as a modified AC-OPF problem. The attacker's goal is to mislead the distributed system operator (DSO) to observe under-voltage issues in AC state estimation by injecting an attack vector into the measurements. Let A, B, T represent the set of nodes in the attack region, boundary nodes, and tie lines, respectively. The objective of the NLRA attack is formulated below.

$$min \sum_{a \in \mathcal{A}} c_a V_a \tag{3.1}$$

Here, subscript a denotes the targeted nodes in the attack region A; V_a represents the intended voltage magnitude measurements of node a; c_a is a non-negative weight coefficient assigned to a node representing the attacking emphasis, the summation of which equates to 1. To achieve the most desirable under-voltage violation, an attacker can assign a large weight to the most critical node, and a zero weight for nodes of no interest.

3.2.4 Stealthiness Constraints

To make NLRA stealthy, an attacker needs to ensure measurements of the tie line's power flow S_T, the voltage magnitude on the boundary nodes V_B, and the voltage phase difference between the boundary nodes $\Delta\theta_B$ remain unchanged after the attack as defined in Equation 3.2.

$$\begin{bmatrix} S_T \\ V_B \\ \Delta\theta_B \end{bmatrix} = \begin{bmatrix} S'_T \\ V'_B \\ \Delta\theta'_B \end{bmatrix} \tag{3.2}$$

Here, superscript $(\bullet)'$ denotes the measurements before an attack. Constraints 3.3 and 3.4, showing the essence of an NLRA, indicate the sum of all net load changes should be equal to zero, and the net load's change at each node is within a reasonable range, respectively.

$$\sum_{a \in \mathcal{A}} \Delta D_a = 0 \tag{3.3}$$

$$-\delta S_a^{l'} \leq \Delta D_a \leq \delta S_a^{l'} \tag{3.4}$$

In 3.3, ΔD_a is the attack magnitude, i.e., net load change, at each node in the attack region. In 3.4, δ is a percentage of allowable change on the original load (apparent power) measurement $S_a^{l'}$. Constraint 3.4 is imposed because the DSO can check the sensor measurements when an under-voltage condition occurs. In this case, unrealistic injected data can be easily exposed.

With local information of the attack region and boundary information between the attack and non-attack regions, NLRA is modeled as a modified ACOPF problem, in which the prevailing ACOPF constraints 3.5-3.9 hold for the proposed NLRA.

$$S^l = P^l + iQ^l \tag{3.5}$$

$$g_P(\theta, V, P^l) = 0 \tag{3.6}$$

$$g_Q(\theta, V, Q^l) = 0 \tag{3.7}$$

$$h_f(\theta, V) \leq 0 \tag{3.8}$$

$$h_t(\theta, V) \leq 0 \tag{3.9}$$

Here, voltage angle θ, voltage magnitude V, real and reactive power load P^l and Q^l are decision variables. In 3.6, g_P is the nonlinear equality constraint of nodal real power balance. In 3.7, g_Q represents the nonlinear equality constraints of nodal reactive power balance. h_f in 3.8 and h_t in 3.9 are nonlinear inequality constraints of power flow limits at the “from node” and “to node”, respectively. Attack vectors generated by the proposed NLRA model obey Kirchhoff’s current & voltage laws, implying they follow the inherent characteristics of the distribution system.

3.2.5 Attack Framework

The flow chart of an NLRA in the distribution system is shown in Figure 3.2. The attacker has the capability of eavesdropping on compromised sensors in the attack region. The attacker can generate an attack vector by running the NLRA model formulated in 3.1-3.9 based on the eavesdropped measurements. Further, the attacker can inject the calculated attack vector back into the corresponding communication links through man-in-the-middle attacks.

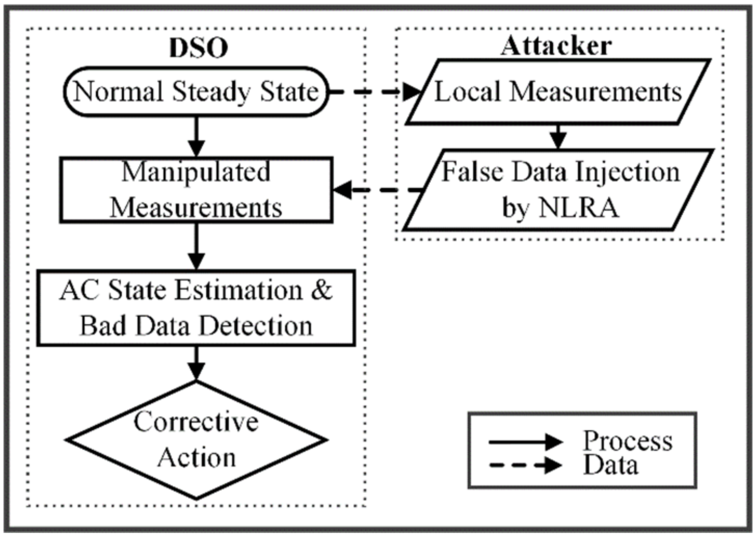


Figure 3.2: The flowchart of NLRA.

In this study, it is assumed that DSO is equipped with an AC state estimator and BDD. The AC state estimator utilizes compromised measurements combined with the measurements in the non-attack region to check the existence of a cyberattack. If the attack bypasses the BDD test, the system will respond to the estimated system states with corrective actions. This section focuses on demonstrating the impact of the NLRA attacks in the distribution system. Corrective actions, and most importantly, defense approaches for the DSO in response to NLRA are out of the scope of this chapter.

3.3 Attacker's Power FLow Enhanced Deep Learning-based SE

The last challenge for an attacker to construct the proposed NLRA is to deal with the estimated voltage states. The approximation method[51] in transmission systems is not feasible here in distribution systems, and the attackers usually do not have enough redundant measurements to implement WLS-based SE. Thus, this chapter proposes to use PFEDL models[194] to estimate the system states. In power systems, redundant factors (RF) are utilized to indicate system observability. The RF is defined as the ratio between the number of available measurements and the number of system states. Typically, WLS-based SE requires an RF of 2.5 to work correctly. Experiment results in this chapter show that in the 69-bus system when the RF is 1.5, the WLS-based SE cannot converge, while the PFEDL-based SE can provide the attackers with accurate system states.

A PFEDL model can incorporate historical measurement data and the power flow model. The hybrid model utilizes deep neural networks to learn the state correlations while considering physical laws in a power system. Inspired by the autoencoder in the artificial intelligent area, the PFEDL model consists of an encoder and a power flow enhanced decoder, as shown in Figure 3.3. The autoencoders are trained to copy their inputs to their outputs with internal layers. Such an internal hidden layer divides the neural network into two parts which are the encoder and decoder. When using a PFEDL-based autoencoder as a state estimator, the measurement $\mathbf{M}$ is fed into the encoder, and the hidden layer output is the estimated system state $\hat{\mathbf{x}}$. To improve the accuracy of the PFEDL state estimator, the estimated state then goes through the decoder to output the restored measurements $\hat{\mathbf{M}}$. Once the restored

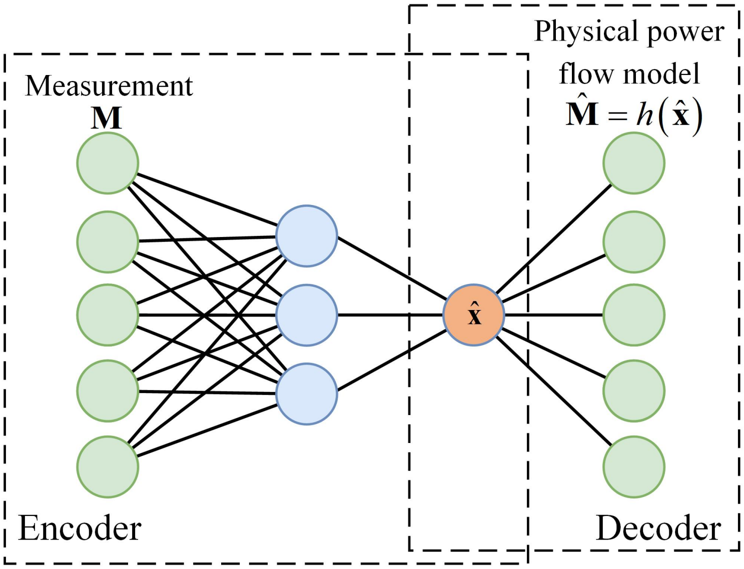
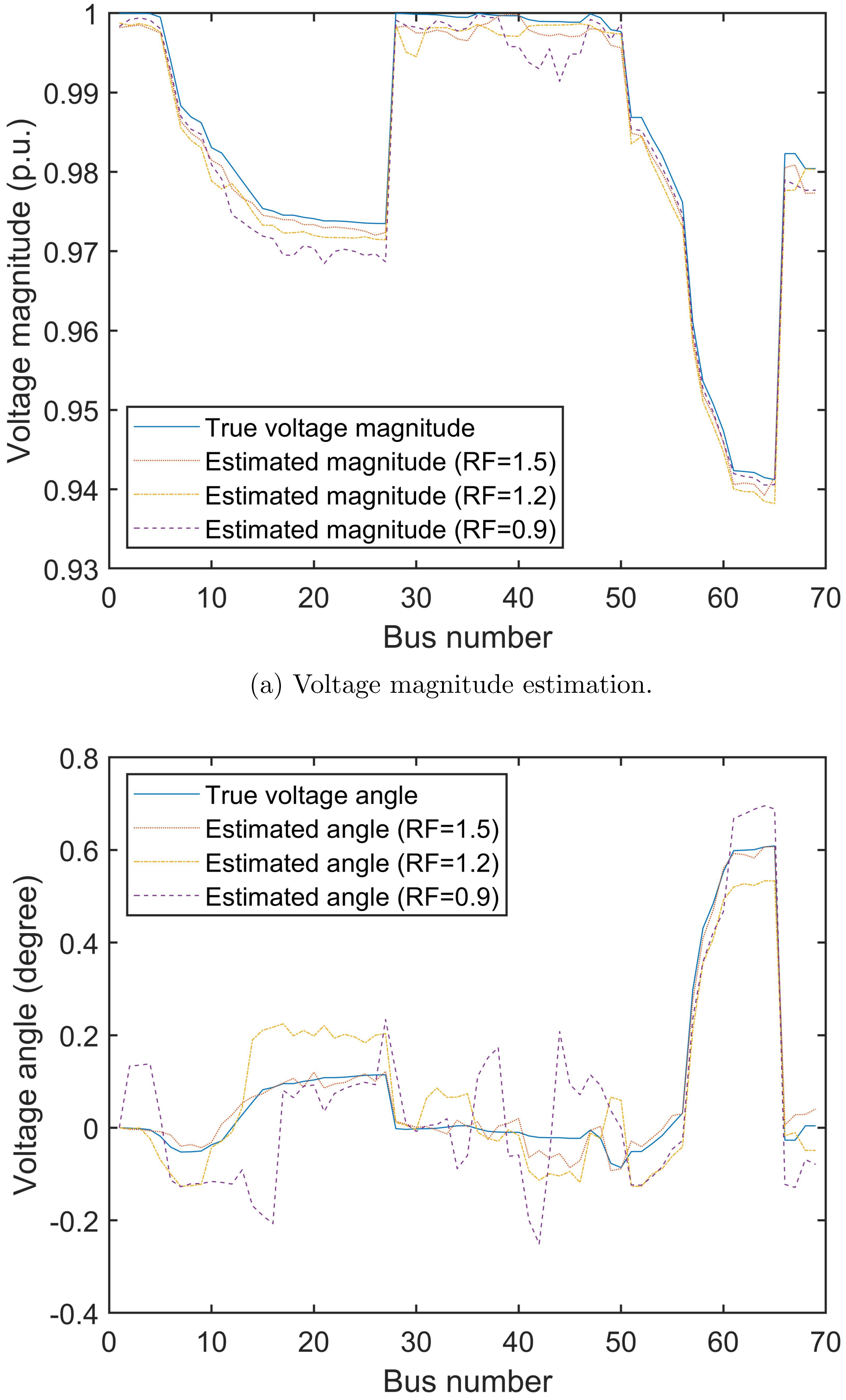


Figure 3.3: Architecture of the power flow enhanced deep learning state estimator

(estimated) measurements are generated, the cumulative error between the actual and the estimated measurement vector is calculated as the loss function to train the PFEDL neural network. The loss function is then back-propagated, thereby adjusting the weights and biases of the neural network accordingly.

Since neural networks can approximate nonlinear functions[195], deep neural networks are utilized as the encoder to estimate the system states given the measurements $\mathbf{M}$. In the experiment section, the attackers' PFEDL-based SE utilizes long short-term memory (LSTM) neural network and a feed-forward deep neural network (FNN). The LSTM neural network is one type of recurrent neural networks. It can map the nonlinear relationship of measurements and states while learning the temporal correlations of the load change.

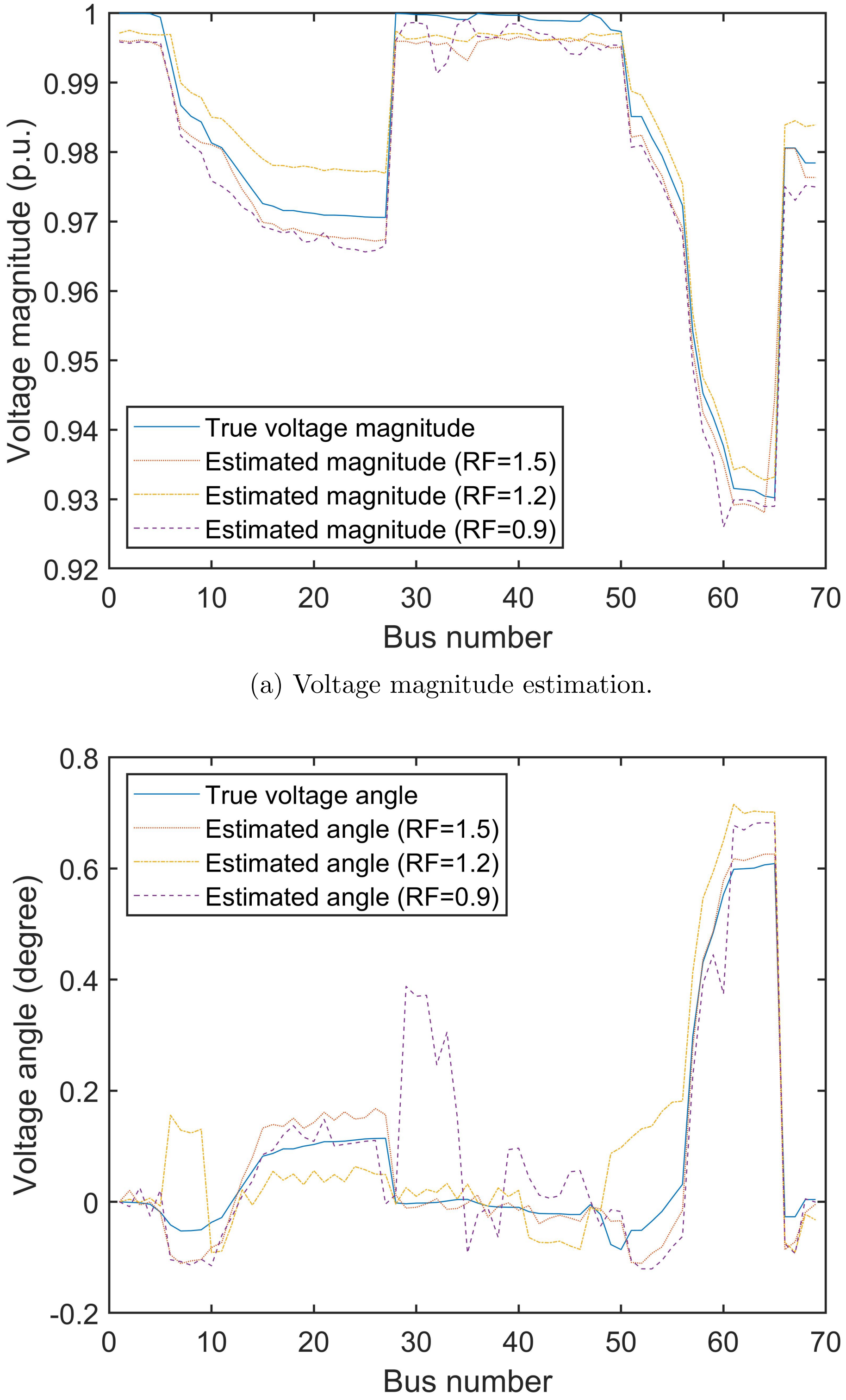
The attacker's PFEDL-based SE results under various RF are compared with the actual system state in Figs. 3.4 and 3.5. One year of hourly load data from ERCOT is used to train and test the PFEDL-based SE. The standard deviation for the measurements is 0.004. The largest redundant factor in the following cases is 1.5, which is still too low for the WLS-based SE to converge. As a comparison, it is seen that the PFEDL-based SE works well, and the estimated states are accurate when RF is 1.5 (dotted red lines). By observing the PFEDL-based SE under various RF, it can be seen that a lower RF for the PFEDL-based



(a) Voltage magnitude estimation.

(b) Voltage angle estimation.

Figure 3.4: PFEDL (FNN) state estimation results under low observability.



(a) Voltage magnitude estimation.

(b) Voltage angle estimation.

Figure 3.5: PFEDL (LSTM) state estimation results under low observability.

Table 3.2: Average MAEs for PFEDL-based SE in 69-bus system

Redundant factors	1.5	1.2	0.9
LSTM	0.0034	0.0037	0.0053
DNN	0.0013	0.0017	0.0020

SE will lead to a worse SE accuracy for both the LSTM and FNN models. The accuracy of the two PFEDL-based SE is evaluated in Table 3.2. The accuracy measure is the mean absolute error (MAE):

$$MAE = \frac{1}{N}\sum_{k=1}^{N} |\hat{x}_k - x_k| \tag{3.10}$$

where N is the number of the estimated states. A lower MAE value means better estimation accuracy.

The average MAE under each RF is calculated from 1,000 cases. For the PFEDL-based SE, the RF values of 1.5, 1.2, and 0.9 are used to calculate the MAEs. Since the WLS-based SE cannot converge when RF is 1.5, this chapter calculates the MAE of the WLS-based SE under a fully measured (RF=3.5) system as a reference. The reference MAE value is 0.0052. When comparing this reference MAE with the MAE results in Table 3.2, it shows that the PFEDL-based SE under a low observable system can perform as well as the WLS-based SE under a fully measured system. The PFEDL-based SE greatly reduces the attack cost, which is an advantage for NLRA attackers. In contrast to WLS-based SE, an attacker's required number of measurements reduces by 57% if the PFEDL-based SE is used to replace the WLS-based SE.

3.4 Experiment Results

In this section, the proposed NLRA model is simulated on a modified PG&E 69-node radial distribution system. The DSO, equipped with the AC state estimator and BDD, has full access to all sensor measurements and global information of the entire distribution system. This section simulates NLRA on this system and assesses its attack consequences while accounting for the impact of attack regions, attack timing, and the sizes of attack regions. The NLRA, AC state estimation, and BDD are all performed in MATPOWER[196].

3.4.1 Test Distribution System

The one-line diagram of the modified 69-node system is shown in Figure 3.6. This section modifies the original 69-node radial distribution system by adding aggregated behind-the-meter DERs at certain nodes (in Figure 3.6). The proposed NLRA is simulated in two regions (i.e., main feeder and lateral) and three time periods (i.e., valley, shoulder, and peak hours). The range of allowable voltage magnitudes is between 0.95 p.u. and 1.05 p.u. in this system.

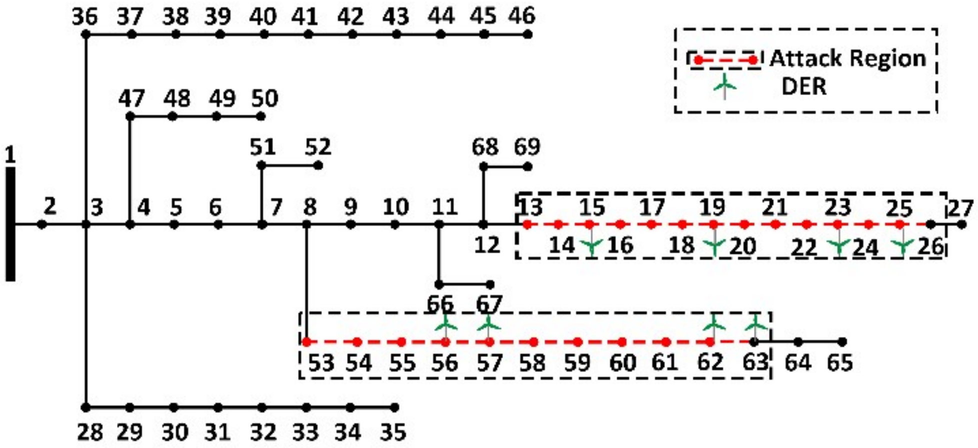


Figure 3.6: The modified PG&E 69-node system.

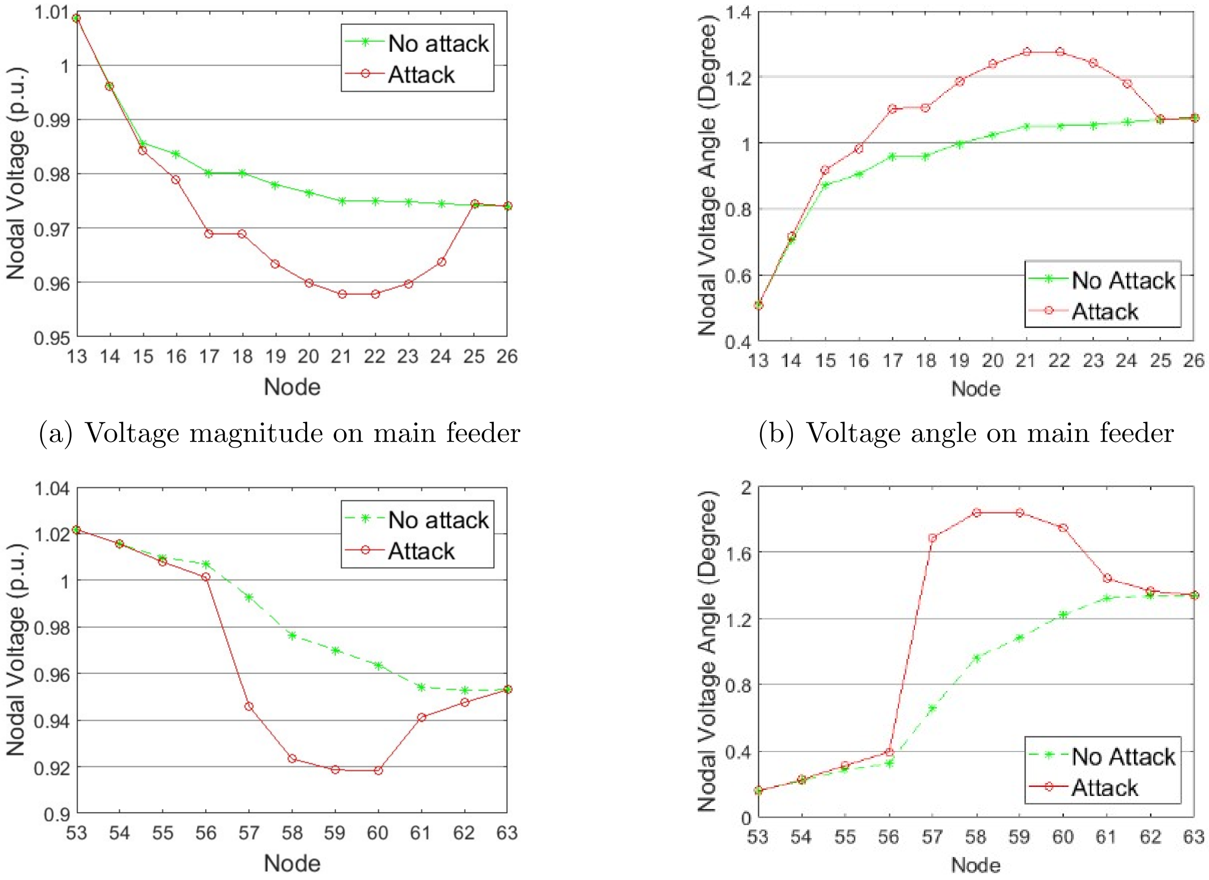
3.4.2 Impact of Attack Regions

This subsection compares the system impact of the attack at two regions, i.e., one on the main feeder (Nodes 13 to 26) and the other on the lateral (Nodes 53 to 63). To demonstrate the flexibility of attacking different numbers of nodes, Nodes 22 to 25 on the main feeder and Node 59 on the lateral are chosen to be the targets, that is, the weight coefficients on these nodes are non-zero. While attackers may pick their target nodes of interest, this section randomly selects the target nodes in the middle of the attack region.

Such a selection allows the voltage magnitude profile to drop first and then rise to satisfy the boundary condition. This is a unique characteristic of the NLRA attacks in the radial distribution system. Figure 3.7 shows nodal voltage magnitudes and angles at a peak hour before and after the attack. In order to make the NLRA stealthy in the attack region, the voltage magnitudes and the incremental phase angles on the boundary nodes remain the same after the attack. When the attack is on the main feeder, the largest voltage drop occurs at Node 21, which is 0.017 p.u. No voltage magnitude of any node drops below the secure range and the DSO observes no under-voltage issue. When the attack is on the lateral, the DSO perceives six voltage violations below the lower limit of 0.95 p.u. at Nodes 57 to 62. The largest voltage drop of 0.057 p.u. occurs at Node 58. The difference in attack consequences between the two attack regions is largely attributed to the difference in line impedance. Specifically, an attack region with higher line impedance would more likely experience larger voltage drops under an NLRA. Therefore, the optimal strategy for an attacker is to launch an NLRA on laterals, where the line impedance is higher than that of the main feeder.

3.4.3 Impact of Attack Timing

This subsection investigates the impact of attack timing on the distribution system by implementing NLRA in different time periods. Figure 3.8 compares the profiles of voltage magnitudes after NLRA on the main feeder and the lateral at peak, shoulder, and valley hours. It is seen larger voltage drops occur during the peak hour in both attack regions. As shown in Figure 3.8b, on the lateral occur five and one nodal voltage violations at the shoulder and the valley hours, respectively. The most severe attack consequences occur on the lateral during the peak hour when the under-voltage condition occurs on six nodes, i.e., Nodes 57-62. This result can be explained by comparing the net load differences at those time periods. A higher load condition provides NLRA with more freedom to manipulate and redistribute the nodal net loads in the attack region.



(a) Voltage magnitude on main feeder

(b) Voltage angle on main feeder

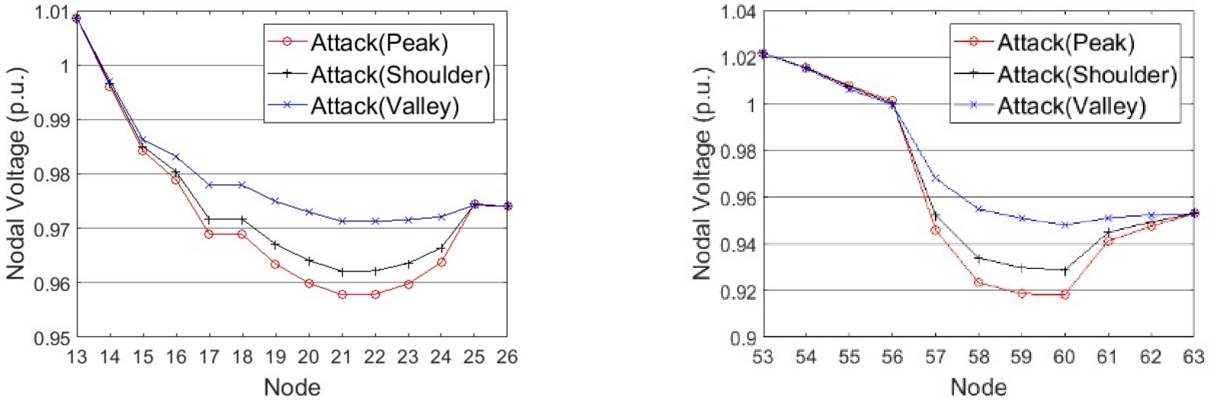
(c) Voltage magnitude on lateral

(d) Voltage angle on lateral

Figure 3.7: Attack consequences on a peak hour.

3.4.4 Impact of System the Size of an Attack Region

This subsection studies the impact of the size of an attack region on the stealthiness of the NLRA. The size of the attack region is reduced by randomly selecting 4 to 7 nodes out of the total 11 nodes (Nodes 53-63) on the lateral as the attack region. For each attack region, this subsection simulates NLRA, the AC state estimation, and the BDD tests 1000 times. In order to show the stealthiness of the NLRA, an attack stealthiness probability (ASP) metric is used, which is defined as the probability of the attack vector bypassing the BDD test. Due to the meter measurement noise, the residual check for a normal system fails with a probability of 0.02 to 0.05, which gives the non-attack case an ASP between 0.95 to 0.98. Figure 3.9 shows the ASP of the four attack sizes. It is seen ASP increases as the size of the attack region increases. This is because the larger the attack region is, the more



(a) Attack on main feeder.

(b) Attack on lateral.

Figure 3.8: Attack consequences in different time periods.

noise-free measurements there will be. The noise-free measurements are the injected false data, which strictly obey Kirchhoff's current and voltage laws and thus reduce the residual in the AC state estimation.

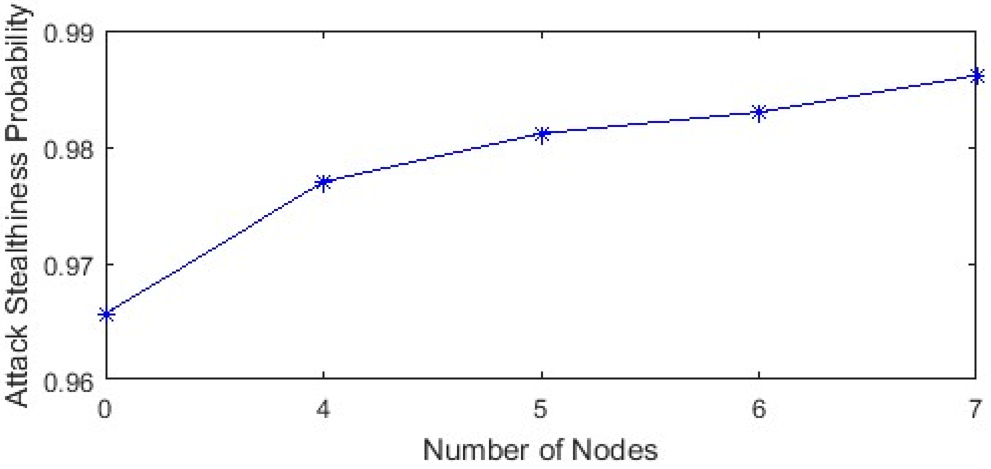


Figure 3.9: ASP versus the number of nodes in the attack region.

3.5 Summary

Based on the concept of the LR attack, this chapter proposes a stealthy NLRA against the AC distribution system with behind-the-meter DERs using local network information. The proposed method is stealthy to BDD in the state estimator and can mislead DSO with illusory under-voltage issues. The numerical results show that the estimation residual after the NLRA is smaller than that before the attack, which can ensure the stealthiness of NLRA. The simulation results demonstrate that a larger attack area in NLRA further reduces the residual in BDD. The NLRA attacks aiming at the region with higher line impedance (e.g., laterals) and larger total net load (e.g., the peak load) lead to larger voltage drops in the attack region. The future work will focus on the development of a defense framework in which the proposed NLRA will be simulated. Attack sequences with a high level of DERs as well as other attack goals in the distribution system will be also researched.

Chapter 4

Voltage Stability Constrained Moving Target Defense

In previous sections, FDI attacks are reviewed and constructed. Evaluation results show that these attacks can mislead the power system SE and cause economic losses or even failures. Many defense strategies are proposed to detect FDI attacks. As an emerging proactive defense strategy, MTD utilizes D-FACTS devices to actively perturbs the branch equivalent impedance to invalidate attackers' knowledge about the power system configuration. Since D-FACTS devices can control power flows and reduce system operational costs, their add-on cybersecurity benefits via MTD have drawn increasing attention in the research society.

4.1 Introduction

The majority of MTD strategies in the literature are designed to detect FDI attacks against state estimation[39;41;165;197]. In 2021, Liu et al.[43] first propounded that there are two intertwined and essential problems associated with MTD, i.e., MTD planning and MTD operation. The MTD planning refers to optimally installing MTD devices (e.g., D-FACTS devices) on an appropriately identified subset of the system (e.g., transmission lines). The

MTD operation determines how to optimally dispatch MTD device setpoints in real-time. In 2014, a random MTD (RMTD) operation[165] was proposed to randomly change the reactance of D-FACTS equipped transmission lines without considering the detection effectiveness. In 2020, a DCOPF-based MTD operation[168] was proposed to minimize the generation cost while ensuring MTD detection effectiveness. An ACOPF-based optimized MTD (OMTD) strategy that minimizes the system loss is introduced in[52]. In[133], Stuxnet-like attacks, which can compromise the control signals to mislead the system to unsafe conditions and inject false sensor measurements to cover the ongoing attack, are detected by MTD. In 2018, Liu et al.[32] defined the "hidden" MTD (HMTD), which optimally changes the branch reactance in AC network to minimize the system loss as well as line power flow differences. An HMTD is stealthy to attackers, even when the attackers are capable of checking the activation of D-FACTS[39]. In[54], Cui et al. propose an HMTD strategy for three-phase unbalanced distribution systems. Lakshminarayana et al.[55] propose to actively perform MTD so that the attacker's knowledge to mask the effects of the physical attack is outdated.

However, MTD operations may deviate the steady-state operating point of a power system from its optimal one, causing massive economic and stability impacts[52]. In[53], voltage stability is defined as the ability of a power system to maintain steady voltages at all buses in the system after being subjected to a disturbance. One of the most common disturbances is the load increases which occur due to the peak load period. To maintain stability after such disturbance, the system needs the preserved capabilities of the transmission network for power transfer. The action of MTD perturbation, which changes the transmission line impedance, may degrade the power transfer capability and cause voltage instability or even voltage collapse during the peak load period. In 2015, Wang et al.[198] proposed an online line switching methodology for increasing load margins to the static stability limit of a look-ahead power system. Cui et al.[199] propose a voltage stability constrained OPF model utilizing a sufficient condition on power flow Jacobian nonsingularity. In 2018, Wang et al.[200] proposed voltage stability constrained OPF by using the minimum singular value of the power flow Jacobian as a voltage stability index. To the best of the author's knowledge, there is no research on MTD operations to detect FDI attacks while guaranteeing system stability.

This chapter aims to fill the gap by proposing a novel voltage-stability-constrained MTD framework against highly structured FDI attacks, especially in the presence of stressful system conditions. One important consideration here is that the voltage-stability improvement ought to be minimally "invasive", meaning such an enhancement should not significantly degrade the attack detection effectiveness of the MTDs or incur a prominent increase in the system operating cost. In summary, this chapter answers the following research questions:

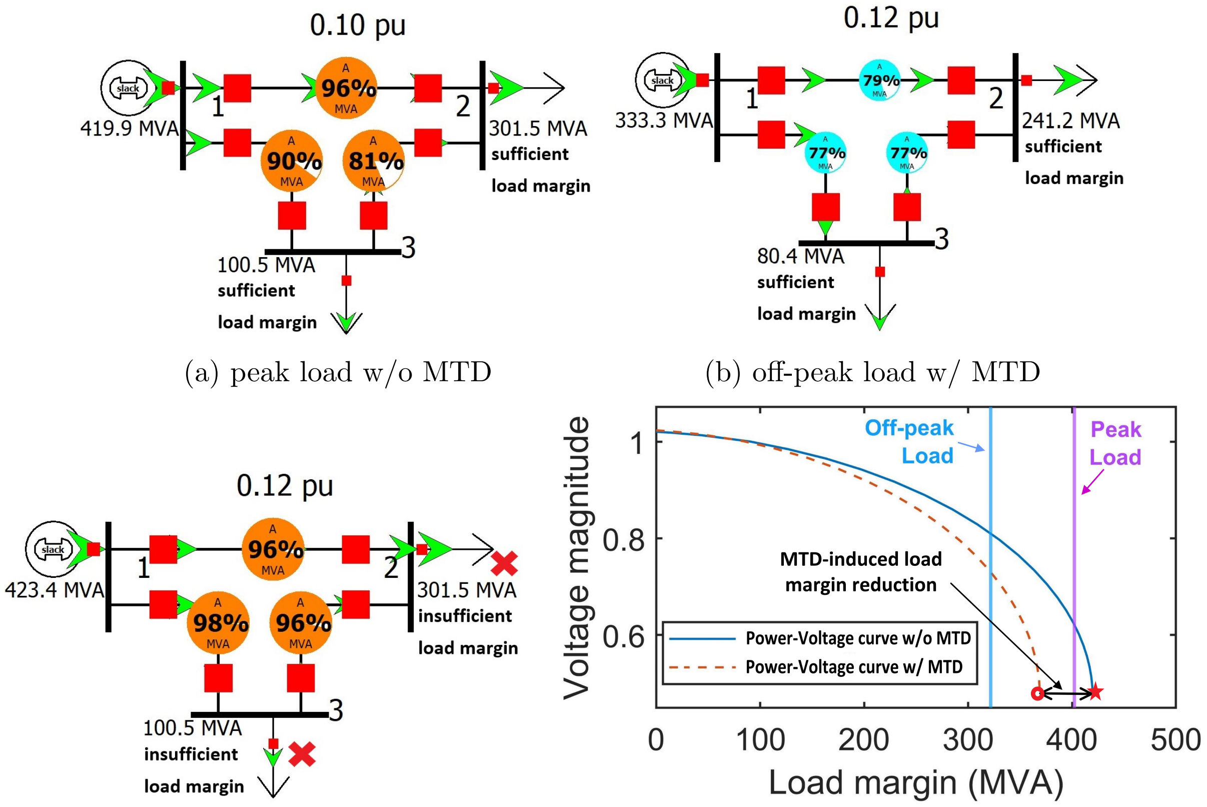
- Can a system with the existing MTD operation methods suffer voltage instability or even experience voltage collapse at the peak load?
- If an MTD can induce voltage instability, how to enhance the existing MTD strategies to improve the voltage stability of a power system?

The first question is answered by simulating the MTD-induced voltage instability on three test systems with real-world load profiles. The second question is answered in Section 4.4.

4.2 MTD-Induced Voltage Instability

This section first shows the voltage stability issue induced by MTD in a 3-bus demo system. Then, a comparative analysis is conducted to show the likelihood of such issues in two complex power systems. A definition of the demonstrated voltage stability issue is given. Meanwhile, the theoretical connection between voltage instability and myopic MTD operations is explained. Last, this section explains why the conventional power flow control methods are unsuitable for solving MTD-induced voltage instability.

Figure 4.1 illustrates the 3-bus system, in which Bus 1 is the slack bus with a generation capacity of 500 MVA. Buses 2 and 3 are the load buses whose off-peak loads are 241.2 MVA and 80.4 MVA, respectively. The load increases by 25% at the peak hour with a fixed power factor. The power flow limits of Lines 1-2, 1-3, and 2-3 are 220 MVA, 215 MVA, and 105



(a) peak load w/o MTD

(b) off-peak load w/ MTD

(c) peak load w/ MTD

(d) power-voltage curves with and without MTD

Figure 4.1: MTD-induced voltage instability in the 3-bus system.

MVA, respectively. The system without MTD at the peak load is in a normal steady-state as shown in Fig. 4.1a. When an MTD is introduced at the off-peak load in Fig. 4.1b, e.g., the impedance of Line 1-2 changes from 0.1 to 0.12 per unit, the MTD would not cause any instantaneous voltage stability issues since the system after the MTD has enough load margin for the off-peak load. However, this is not the case when it comes to the peak load in Fig. 4.1c, where the originally sufficient load margin at the off-peak becomes insufficient. Figure 4.1d illustrates the MTD-induced voltage instability by using the CPF P-V curve. As seen, the solid curve and the dashed curve represent the CPF P-V curves without and with the MTD, respectively. The saddle node bifurcation (SNB) point of the P-V curve without MTD (star marker) is to the right of both the peak and off-peak loads. This indicates that the system without MTD has enough load margin and no voltage stability issue for both the off-peak and the peak loads. However, when the MTD is introduced, the SNB point (circle marker) of the dashed curve is between the off-peak and peak loads. The amount of reduced load margin attributed to the MTD is labeled in Fig. 4.1d as the MTD-induced load margin reduction. Figure 4.1d indicates that the MTD would not cause any instantaneous voltage stability issues at the off-peak load; however, the system will not have sufficient load margin at the peak load and will thereby undergo voltage stability issues.

To facilitate the presentation, this section defines MTD-induced voltage instability as voltage instability under stressful conditions due to the load margin decrease by an MTD previously implemented. More specifically, the system originally had a sufficient load margin to cope with the stressful condition, but the very sufficiency no longer exists after an MTD is implemented.

To clearly show the likelihood of MTD-induced voltage instability in more complex power systems, this section further tests the following two systems: 1) the Western Electricity Coordinating Council (WECC) 240-bus system[201] and 2) the 300-bus system[196]. Both systems have sufficient load margins if no MTD is implemented. For each system, 3,000 RMTD scenarios are carried out at the given default load. The MTD magnitude is within 20% of the original line impedance. In the WECC 240-bus system, there is no voltage instability at the default load immediately after all 3,000 RMTD scenarios are executed. However, when it approaches the peak load, 313 out of 3,000 scenarios, i.e., 10.43% of the total scenarios, induce voltage instability due to the load margin reduction by MTD. Another interesting case to show the MTD-induced voltage instability is the 300-bus system. Due to a lack of load profile in this system, this section only studies the system under the default load[196], where an ACOPF problem with MTD successfully converges. Nevertheless, 98 out of 3,000 RMTD scenarios (i.e., 3.26%) incur voltage instability immediately after executing the MTD because of the load margin reduction. These simulation results show that existing MTDs are likely to induce voltage instability in larger-scale power systems. In reality, power systems are much more complicated than the examples of 3-bus, 14-bus, WECC 240-bus, and 300-bus systems. Thus, MTD-induced voltage instability ought to be systematically addressed for any realistic MTD applications, particularly in the presence of drastic net load variations caused by an increasing amount of renewable generation.

The MTD-induced voltage instability resides in existing MTD methods that myopically perturb the line impedance without looking-ahead capabilities for reserving sufficient load margin. The lack of such capability in existing MTD models may lead to insufficient load margin[53], which can cause voltage instability or even voltage collapse. The connection between voltage instability and MTD can be demonstrated by using the CPF P-V curve, as shown in Fig. 4.1d. The SNB point of this curve depends on the transmission line impedance and is used to calculate the system load margin given the current load condition. Once the system load increases beyond the SNB point, the power flow problem becomes ill-conditioned. This is because the Jacobian matrix of the power flow problem has a sufficiently large condition number[202]. A large condition number is generally associated with small singular values or eigenvalues of a matrix. The voltage instability is also related to small eigenvalues[203]. Therefore, such an ill-conditioned power flow problem is proven to be equivalent to the system voltage instability[204]. Since an MTD can alter the SNB point through the line impedance perturbation, it may reduce the voltage stability of a power system and even result in voltage collapse.

It is worth mentioning that the MTD-induced voltage instability cannot be solved by the conventional MTD-based power flow control methods. Existing D-FACTS enabled power flow control methods[32;43;54;168] only consider a snapshot in power system operation, meaning that the system voltage stability is guaranteed right after the D-FACTS setpoint is changed. That might be enough for traditional power flow control methods, in which the D-FACTS setpoints are less frequently altered (e.g., seasonal changes or as needed). However, as MTD is introduced as a by-product of D-FACTS devices to enhance the cybersecurity of the power grid[52], their setpoint changes become much more frequent (e.g., intra-day changes), which the traditional D-FACTS control methods are not prepared for. Therefore, the first objective of this work is to provide existing MTD strategies with look-ahead capabilities to prevent MTD-induced voltage instability. Furthermore, existing D-FACTS enabled power flow control strategies cannot provide guidance on how to adjust the D-FACTS setpoints, if deemed necessary, while simultaneously improving the voltage stability and maintaining the MTD performance. This represents another objective of this work.

4.3 Preliminaries

This section introduces the background knowledge of MTD planning, MTD detection effectiveness, power injection to impedance sensitivity matrix, and voltage stability index t as preliminaries.

4.3.1 MTD Planning and Detection Effectiveness

Two D-FACTS placement strategies, i.e., max-rank placement[52] and graph-based placement[43], are used in this chapter to guarantee the MTD detection effectiveness against net load redistribution attacks. To date, it is extremely challenging to quantitatively measure the MTD detection effectiveness in nonlinear, full AC models. Thus, it is customary to utilize the DC model[43;56;168;197;205] to theoretically analyze the detection effectiveness. The placement strategies[43;52] are developed in DC models, and their performance is verified in their AC counterparts. The MTD detection effectiveness of the proposed methods can be theoretically explained as follows. In[52], a sufficient condition for the max-rank placement is proposed by using a graph theory-based topology analysis. By ensuring that there exists no loop in the D-FACTS graph and the non-D-FACTS graph, the max-rank placement is designed to utilize the minimum number of D-FACTS devices to achieve the maximum rank of its composite matrix, which is indicative of the MTD detection effectiveness. In[43], the graph-based placement is proposed as an enhanced max-rank placement to retain the maximum rank of its composite matrix while eliminating the unprotected buses by using additional D-FACTS devices. By following these two placement methods, all the constructed MTDs have the max rank of their composite matrices if no D-FACTS device is in the idle state. Thus, the MTD detection effectiveness is largely ensured by the D-FACTS placement strategies[43;52].

4.3.2 Power System Quantities to Impedance Sensitivity

The power injection to impedance (PII) sensitivity is originally proposed, as an intermediate step in the chain rule of calculus, to determine the relationship between the state variables and line impedance[206]. In this chapter, the PII is utilized to calculate how much the system load margin can be increased due to the adjustment of the original MTD setpoints, when the system is near the power flow singularity (i.e., SNB point). The sensitivities of power injections to a change in line impedance are denoted as:

$$\left[\ \Delta S\ \right] = [PII]\,[\Delta x_l] \tag{4.1}$$

$$PII \triangleq \left[\frac{\partial S_i}{\partial x_l}\right] \tag{4.2}$$

The apparent power injection S_i at Bus i is differentiated with respect to x_l for all lines that connect Bus i and the adjacent Buses. With the help of PII, the necessary MTD setpoints adjustment can be calculated under the most stressful system condition.

It should be noted that the Power Flow to Impedance (PFI) matrix and the voltage State to Impedance (SI) matrix proposed in[206] are not directly suitable for the MTD setpoint adjustment. As discussed in Section 4.2, the MTD-induced voltage instability is directly related to insufficient load margin rather than steady-state power flow and voltage. In contrast, with the desired load margin as the objective, the PII matrix can be straightforwardly used to calculate the MTD adjustment such that the MTD-induced voltage instability is circumvented.

4.3.3 Voltage Stability Index (t-index)

The CPF method uses an iterative process involving predictor and corrector steps that require high computational cost for large systems. A different strategy to represent the voltage instability is by using the minimum singular value of the power flow Jacobian. Cui et al.[199] proposed a voltage stability margin index to quantify the power flow Jacobian nonsingularity. The proposed voltage stability index is derived from a sufficient condition for the nonsingularity of power flow Jacobian[207]. A voltage stability index t_i for each load bus i is defined as:

$$t_i = |V_i| - \sum_{j=1}^{n} \frac{|Z_{ij} S_j|}{|V_j|}, \quad i, j \in \mathcal{N} \tag{4.3}$$

where $|V|$ is the voltage magnitude, S is the apparent power injection, Z_{ij} is the bus impedance matrix element, and $\mathcal{N}$ is the set of n load buses. A larger t-index value indicates a better voltage stability performance at a load bus. Contrasted with the CPF method, the t-index calculation does not require an iterative process which could greatly save computational efforts for a large system. As opposed to the CPF method, the t-index method is more suitable when the system operator is only concerned about power flow Jacobian singularity, while the tracing of the power flow solution path is not necessary.

4.4 Voltage-stability-constrained MTD Framework

This section proposes two voltage stability constrained MTD methods, i.e., a t-index optimization method and a load margin constrained method, to ensure the system voltage stability with sufficient load margin. To distinguish the system operation point with or without MTD, this chapter defines hereinafter the D-FACTS operation point before MTD as *pre-MTD*, while the operation point after MTD as *post-MTD*.

4.4.1 t-Index Optimization Method

This subsection first derives t-index to impedance sensitivity matrix (TII) and then forms an optimization problem to maximize the t-index for the most critical forecasted load $S' = max([S_{t_1}, S_{t_2}, S_{t_3}, ..., S_{t_N}])$, where t_1 to t_N are the time indices of the look-ahead time periods within an MTD window. The basic idea of the t-index optimization method is to maximize the lowest t-index among all the load buses of a system implemented with an original MTD. The proposed method is a post-MTD method that adjusts the original MTD setpoints.

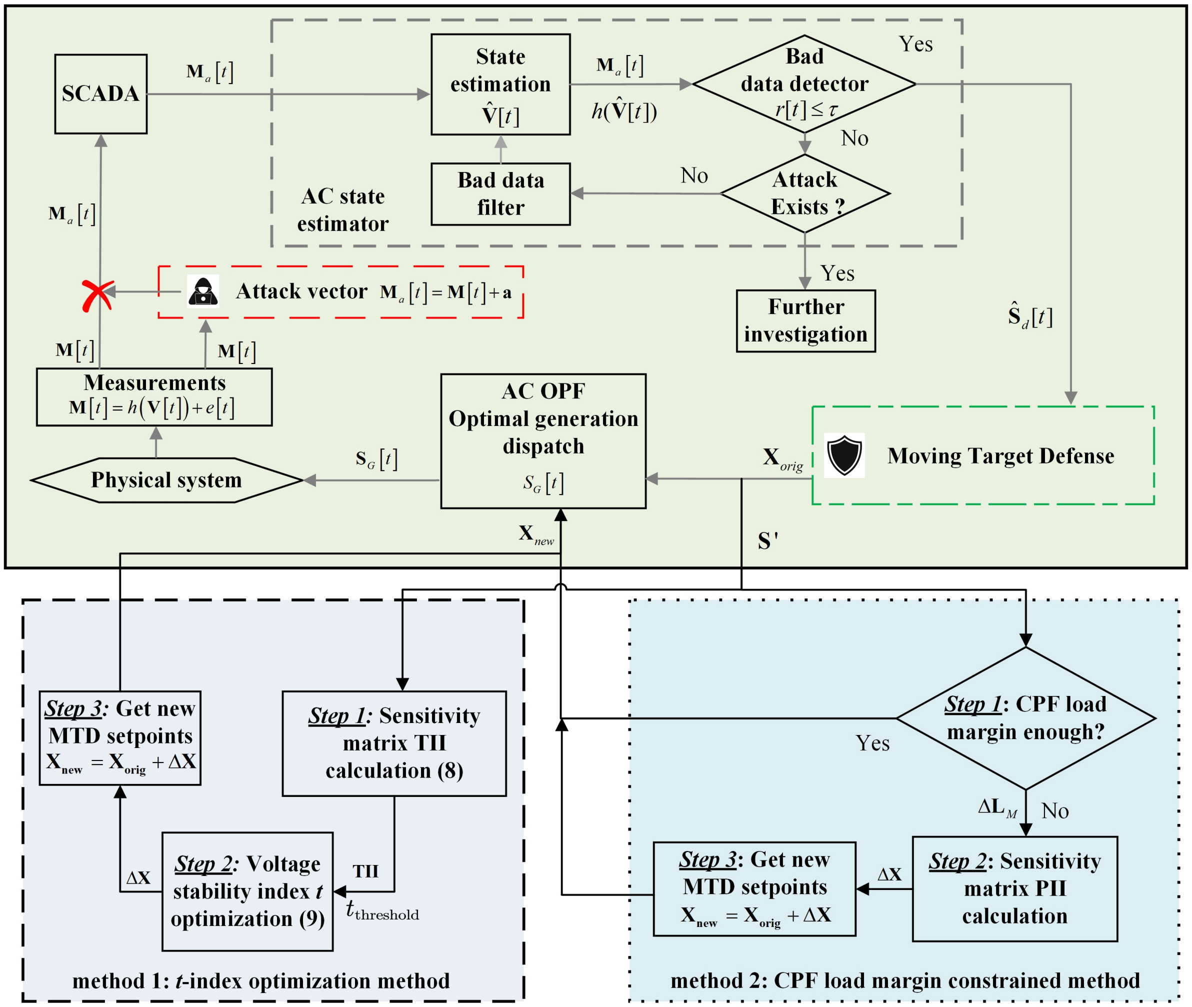

SCADA
$\mathbf{M}_a[t]$
State estimation
$\hat{\mathbf{V}}[t]$
$\mathbf{M}_a[t]$
$h(\hat{\mathbf{V}}[t])$
Bad data detector
$r[t] \le \tau$
Yes
No
Bad data filter
No
Attack Exists ?
AC state estimator
$\mathbf{M}_a[t]$
Attack vector $\mathbf{M}_a[t] = \mathbf{M}[t] + \mathbf{a}$
Yes
Further investigation
$\hat{\mathbf{S}}_d[t]$
$\mathbf{M}[t]$
$\mathbf{M}[t]$
Measurements
$\mathbf{M}[t] = h(\mathbf{V}[t]) + e[t]$
AC OPF Optimal generation dispatch
$S_G[t]$
Moving Target Defense
$\mathbf{X}_{orig}$
Physical system
$S_G[t]$
$\mathbf{X}_{new}$
S'
Step 3: Get new MTD setpoints
$\mathbf{X}_{new} = \mathbf{X}_{orig} + \Delta\mathbf{X}$
Step 1: Sensitivity matrix TII calculation (8)
Step 2: Voltage stability index t optimization (9)
ΔX
TII
$t_{\text{threshold}}$
method 1: t-index optimization method
Step 1: CPF load margin enough?
Yes
$\Delta\mathbf{L}_M$
No
Step 3: Get new MTD setpoints
$\mathbf{X}_{new} = \mathbf{X}_{orig} + \Delta\mathbf{X}$
ΔX
Step 2: Sensitivity matrix PII calculation
method 2: CPF load margin constrained method

Figure 4.2: Flowchart of the new MTD framework with the proposed methods built-in.

TII Sensitivity Matrix

TII sensitivity matrix represents the relationship between the change of t-index ΔT and the change of MTD setpoints ΔX on the branches equipped with D-FACTS devices. The TII sensitivity matrix is described as follows:

$$\Delta T = TII \times \Delta X \tag{4.4}$$

$$TII \triangleq \left[\frac{\partial t_i}{\partial x_l}\right] = \left[\frac{\partial t_i}{\partial Z_{ij}} \times \frac{\partial Z_{ij}}{\partial x_l}\right], \quad i, j \in \mathcal{N}, \quad l \in \mathcal{L} \tag{4.5}$$

where TII is an $\mathcal{N} \times \mathcal{L}$ matrix, $\mathcal{L}$ is the set of D-FACTS equipped transmission lines l.

From (4.3), it is shown that the t-indices at load buses are functions of the bus impedance matrix elements. To get the derivative of the t-index, the t-index at each load bus i is firstly differentiated with respect to the bus impedance matrix Z. Then, chain rule can be used to combine $\partial t_i / \partial Z_{ij}$ with $\partial Z_{ij} / \partial x_l$. During the derivative of t-index, the net power injection can be assumed as constant. Thus, t is a function of Z and V, $t_i = f(Z, V)$. For each load bus i, the derivative of t_i over Z is calculated by

$$\begin{aligned}
\frac{\partial t_i}{\partial Z_{ij}} &= \frac{\partial \left|V_i\right|}{\partial Z_{ij}} - \frac{1}{2} \sum_{j=1}^{n} \left(\frac{Z_{ij} S_j Z_{ij}^* S_j^*}{V_j V_j^*} \right)^{-\frac{1}{2}} \frac{\partial}{\partial Z_{ij}} \left(\frac{Z_{ij} S_j Z_{ij}^* S_j^*}{V_j V_j^*} \right) \\
&= \frac{\partial \left|V_i\right|}{\partial Z_{ij}} - \frac{1}{2} \sum_{j=1}^{n} \left(\frac{Z_{ij} S_j Z_{ij}^* S_j^*}{V_j V_j^*} \right)^{-\frac{1}{2}} S_j S_j^* \frac{\partial}{\partial Z_{ij}} \left(\frac{Z_{ij} Z_{ij}^*}{V_j V_j^*} \right) \\
&= \frac{\partial \left|V_i\right|}{\partial Z_{ij}} - \frac{1}{2} \sum_{j=1}^{n} \frac{\left|S_j\right|}{\left|Z_{ij}\right| \left|V_j\right|} \left(\frac{\partial \left(Z_{ij} Z_{ij}^*\right)}{\partial Z_{ij}} - \frac{\left|Z_{ij}\right|^2}{\left|V_j\right|^2} \frac{\partial \left(V_j V_j^*\right)}{\partial Z_{ij}} \right)
\end{aligned} \tag{4.6}$$

Note that for complex number C, $|C|^2 = CC^*$ holds, where C^* is the conjugate of C. For a normal complex derivative, Z_{ij}^* is not differentiable. This is because, for a complex limit calculation, a conjugate function variable can approach zero from different directions in the complex domain and results in different solutions, which is against the Cauchy–Riemann equations. Since $X \gg R$ in transmission systems, the line resistance can be ignored, and assume Z consists of pure imaginary variables. Then, $\frac{dZ_{ij}^*}{dZ_{ij}} = -1$ and

$$\frac{\partial t_i}{\partial Z_{ij}} = \frac{\partial \left|V_i\right|}{\partial Z_{ij}} + \sum_{j=1}^{n} \frac{\left|S_j\right|}{\left|V_j\right|} + \sum_{j=1}^{n} \frac{\left|S_j\right| \left|Z_{ij}\right|}{\left|V_j\right|^2} \frac{\partial \left|V_j\right|}{\partial Z_{ij}} \tag{4.7}$$

When substitute (4.7) into (4.5), $\partial \left|V_i\right| / \partial Z_{ij}$ in (4.7) will turn to $\partial \left|V_i\right| / \partial x_l$ after the chain rule (4.5). Since the derivative of voltage magnitude over line impedance is equivalent to the SI sensitivity in[206], $\partial \left|V_i\right| / \partial x_l$ can be replaced with the SI elements. The $(i, l)^{th}$ element in the TII matrix can be calculated as,

$$\begin{aligned}
TII_{il} &= \left[\frac{\partial t_i}{\partial x_l} \right] \\
&= SI_{il} + \sum_{j=1}^{n} \frac{\left|S_j\right|}{\left|V_j\right|} \frac{\partial Z_{ij}}{\partial x_l} + \sum_{j=1}^{n} \frac{\left|S_j\right| \left|Z_{ij}\right|}{\left|V_j\right|^2} SI_{il}
\end{aligned} \tag{4.8}$$

For each transmission line equipped with D-FACTS devices, $\partial Z_{ij}/\partial x_l$ is calculated with respect to a unit step impedance change Δx_l.

t-Index Optimization Model

Based on the aforementioned TII sensitivity matrix, this section further proposes a t-index maximization model (4.9) to adjust the MTD setpoints. To facilitate the presentation, let subscript *orig* denote an original post-MTD system state without using any voltage stability enhancement methods, and subscript *new* represent the state adjusted by using the proposed voltage stability methods. As original MTD operation methods optimize D-FACTS setpoints to achieve MTD hiddenness, maximize attack detection effectiveness, minimize power generation costs, and to minimize system losses[52], any adjustment on the original MTD setpoints would deviate from the optimal values. Therefore, the proposed model (4.9) also minimizes the MTD setpoints adjustment for a minimal impact on the original MTD performance.

$$\min_{\Delta X, t_{\text{threshold}}} \quad \delta_1 \|\Delta X\|_2 - \delta_2 t_{\text{threshold}} \tag{4.9}$$

$$\text{s.t.} \quad t_{\text{threshold}} \le T_{\text{orig}} + \Delta T \tag{4.9a}$$

$$\text{LB} \le X_{\text{orig}} + \Delta X \le \text{UB} \tag{4.9b}$$

$$\Delta T = TII \times \Delta X \tag{4.9c}$$

where ΔX is the MTD setpoint adjustment which will be added to the setpoints in the original MTD X_{orig}. The final output of the proposed model is the optimized setpoints $X_{\text{new}} = X_{\text{orig}} + \Delta X$. δ_1 and δ_2 are the weighted coefficients to balance the trade-off between the impact on the performance of the original MTD and the t-index increase. The first component of the objective function (4.9) minimizes the adjustment of the MTD branch impedance which ensures the adjustment will not significantly affect the performance of the original MTD. The second component of (4.9) maximizes (i.e., minimize negative) the t-index threshold $t_{\text{threshold}}$, which is equivalent to maximizing the t-index at the most critical load bus. T_{orig} is the vector of t-index in the system with the original MTD at the peak net load S'. Constraint (4.9a) is the t-index threshold constraint to ensure the lowest t at the most critical load bus is greater than the t-index threshold. Constraint (4.9b) aims to ensure the total impedance change after the adjustment is within the physical capacity of D-FACTS devices. LB and UB are the lower and upper bounds of line reactance perturbation, where UB and LB are equal to ±20% of the transmission line impedance which is generally used in MTD[32;39;41;52;165;197]. In (4.9c), ΔT is the vector of the incremental t-index at all load buses calculated based on TII. In this chapter, the original MTD strategies are the ACOPF-based OMTD[52] and RMTD, in which all the D-FACTS devices have been properly placed according to[43] to ensure the MTD detection effectiveness. The prior work in[43] showed that the detection effectiveness of the original MTDs can be guaranteed as long as each of the D-FACTS device placed is operated in a non-idle state. Based on the salient feature of the original MTD, the L-2 norm is used in (4.9) to minimize the Euclidean distance from X_{orig} to X_{new} for each D-FACTS device such that the feature of the original MTD (e.g., optimized generation cost, MTD detection effectiveness) can be preserved to the largest degree.

The steps of the proposed t-index optimization method are shown in Algorithm 1. In Step 1, Algorithm 1 calculates the TII sensitivity matrix of the system with original MTD setpoints. By following (4.8), the TII matrix is constructed to reveal the relationship between the t-indices on load buses and the impedance of the D-FACTS lines. In Step 2, the TII-matrix is then used in (4.9) to determine the t-index constrained MTD setpoints. In Step 3, the solution ΔX from (4.9) is the expected MTD setpoints adjustment. The new MTD setpoints in the proposed t-index constrained MTD are calculated by adding ΔX to the original MTD setpoints, i.e., $X_{\text{new}} = X_{\text{orig}} + \Delta X$.

```
Algorithm 1 t-index optimization method
Input: $X_{\text{orig}}$, $S'$
Output: $X_{\text{new}}$
1: Calculate $TII$ from (4.8)
2: Solve the $t$-index optimization problem (4.9)
3: $X_{\text{new}} = X_{\text{orig}} + \Delta X$
4: return $X_{\text{new}}$
```

In general, TII is calculated and the t-index optimization method is carried out for the most critical net load condition S' within an MTD window. Since the proposed model (4.9) maximizes the t-index at the most critical load bus, the t-index at the load bus with high voltage stability may degrade. However, this is typically acceptable as the entire system remains voltage stable under the most stressful condition. Note that the weight coefficients can be finely tuned to find the trade-off between the MTD's performance and the voltage stability. For instance, when higher variability and uncertainty of renewable generation are considered, a higher weight can be placed on the voltage stability rather than maintaining a small impact on the original MTD's performance. One contribution of this work is the development of the TII matrix that can be used to model a linear relationship between the t-index and the D-FACTS setpoint for enhancing the system voltage stability. The linear model can be seamlessly integrated with another objective function (e.g., L-0 norm) for different original MTD strategies or to serve distinct purposes of the system operator.

4.4.2 Load Margin Constrained Method

Load margin L_M is another noteworthy metric for measuring the system voltage stability. It is defined as the maximum amount of load that the system can support given a system configuration. With a specific system configuration and peak load forecast, the load margin is calculated by CPF with a predictor-corrector method. As previously discussed, all existing MTD methods fail to consider the system load margin that is very likely to degrade by MTDs. This motivates us to develop a load margin constrained MTD method. The proposed method is demonstrated in the dotted box of Fig. 4.2 and the steps, shown in Algorithm 2, are described as follows:

- *Step 1*: Algorithm 2 checks the original D-FACTS setpoints X_{orig} by using the CPF method. The load margin L_M is calculated given the original MTD and forecast peak load. If the load margin of X_{orig} is able to satisfy the most critical load forecast within an MTD window, i.e., $S' \leq L_M$, these setpoints can be applied to the system without adjustment. Otherwise, the expected incremental load margin can be calculated by

Algorithm 2 load margin constrained method

```
Input: $X_{\text{orig}}$, $S'$
Output: $X_{\text{new}}$
Solve CPF problem to get load margin $L_M$ of $X_{\text{orig}}$
if $S' \leq L_M$ then (as they satisfy the load margin constraint for the most critical condition)
    return $X_{\text{orig}}$
else
    Calculate the expected load margin increase $\Delta L_M$
    Compute the $PII$ from (4.2)
    Solve $\Delta X = PII^{-1} \times \Delta L_M$, subject to LB $\leq X_{\text{orig}} + \Delta X \leq$ UB
    $X_{\text{new}} = X_{\text{orig}} + \Delta X$
end if
return $X_{\text{new}}$
```

$\Delta L_M = S' - L_M$.

- *Step 2*: Algorithm 2 computes the sensitivity matrix PII in (4.2). PII reveals the relationship between the expected incremental load margin ΔL_M and the line impedance change ΔX on the branches equipped with D-FACTS devices. After computing PII, Algorithm 2 calculates the expected MTD setpoint adjustment by $\Delta X = PII^{-1} \times \Delta L_M$. ΔX must ensure the MTD setpoint after the adjustment is still within the physical limits of the D-FACTS devices, i.e., LB $\leq X_{\text{orig}} + \Delta X \leq$ UB. LB and UB are the same as used in the t-index optimization method.

- *Step 3*: The load margin constrained MTD setpoints are calculated by adding ΔX to the original MTD setpoints, i.e., $X_{\text{new}} = X_{\text{orig}} + \Delta X$. The new MTD setpoints are then returned by Algorithm 2.

Notice that the scope of this chapter is on the voltage stability issue induced by MTD only. In other words, the pre-MTD system state is voltage stable even under the most stressful conditions without MTD. In view of the MTD setpoint that is perturbed around the pre-MTD system state ΔL_M calculated in *Step 2* should be comparatively small. However, if the system is not pre-MTD voltage sable ΔL_M can be large. In such a case, feasible X_{new} may not exist due to the physical limits of the D-FACTS devices. Other methods[208;209] including the potential load shedding as a last resort, need to be considered to ensure the pre-MTD voltage stability of the system.

4.4.3 Proposed Voltage Stability Constrained MTD Framework

Figure 4.2 illustrates the flowchart of the proposed voltage-stability constrained MTD framework that integrates Algorithm 1 and 2 proposed in this section. These two methods lie in the post-MTD process where original MTD setpoints are calculated and can be adjusted if deemed necessary. The core idea in designing this framework is that the proposed methods ought to greatly enhance the system voltage stability within an MTD rolling window, but should not significantly degrade the attack detection effectiveness of the original MTD setpoints or incur a prominent increase in the system operating cost.

By comparing the two proposed algorithms, the t-index optimization method has an advantage over the load margin constrained method that the adjusted MTD setpoints X_{new} are typically closer to the original MTD setpoints X_{orig} since the minimization of the setpoint deviation in (4.9). This is much desirable when the original MTD is an OPF-based MTD strategy (e.g., OMTD and HMTD) with specific objectives including system cost minimization, attack detection probability maximization, and/or MTD hiddenness requirement. The t-index optimization method can improve the voltage stability while maintaining the original MTD performance as much as possible. Both of the proposed methods are computationally efficient since they only involve matrix computations, solving a series of power flow problems, and solving a nonlinear minimization problem with all linear constraints. The numerical tests show that the proposed methods can solve an IEEE 118-bus case with 60 D-FACTS devices within 20 seconds on a desktop computer. More comparative numerical results will be shown in the next section. In comparison to the original MTD, the proposed methods make the following novel modifications: 1) A load profile, especially the forecasted peak load, is taken into consideration when the proposed methods are executed. This enhancement provides the original MTD with the look-ahead capability to effectively prevent MTD-induced voltage instability; 2) By adjusting the original MTD setpoints, the proposed methods can improve the voltage stability while maintaining the original MTD's objectives such as the generation cost minimization[52] and detection effectiveness maximization[43]. In addition, the proposed methods are computationally-efficient post-processing approaches for adjusting the original MTD setpoints. This feature enables the proposed methods to be seamlessly integrated with existing MTD strategies.

4.5 Experiment Results

This section presents the case study and simulation results on the proposed methods and framework. The net load redistribution attack against MTD detection cases are tested on the IEEE 14-bus and 118-bus systems available from MATPOWER[196]. The t-index optimization problem is solved by the FMINCON toolbox in MATLAB. The load margin constrained method is implemented by using the CPF toolbox in MATPOWER. In order to compare the performance of the proposed methods with various MTDs, this section uses two D-FACTS placement methods, i.e., max-rank[52] and graph-based placement[43], as well as two MTD operational strategies, i.e., RMTD and OMTD in the case study. The max-rank placement solution for the IEEE 14-bus system and the IEEE 118-bus system can be found in[52]. Specifically, the placement uses 7 and 62 D-FACTS devices in the IEEE 14-bus and the IEEE 118-bus system, respectively. The graph-based placement uses 9 and 97 D-FACTS devices in the IEEE 14-bus system and the IEEE 118-bus system, respectively. The graph-based placement solution for the IEEE 14-bus and the IEEE 118-bus system can be found in[43]. The simulations are performed on a desktop with an Intel Core i5 processor and 8 GB RAM. The line impedance change in all the cases are set to be within 20% of the original impedance.

4.5.1 Accuracy Analysis of TII Sensitivity Matrix

This section analyzes the error of the TII sensitivity matrix under several IEEE systems and MTD magnitudes. Given an MTD setpoint, the estimated t-index can be calculated by adding the ΔT (4.4) to the original t-index. Figure 4.3 shows the errors between the real

t-index and the t-index estimated by the TII sensitivity matrix on all the load buses in the IEEE 14-bus system. It is seen that the errors between the real and calculated t-index are small, meaning that the TII sensitivity matrix (4.8) is accurate. In addition, this chapter evaluated the mean absolute percentage errors of the calculated t-index in five IEEE systems under various MTD magnitudes (from 5% to 20% of the original line impedance). Each mean absolute percentage error is calculated under 1,000 RMTD scenarios. From Table 4.1, it is seen that the largest mean absolute percentage error is 2.288%, which indicates the TII sensitivity matrix can accurately derive the t-index.

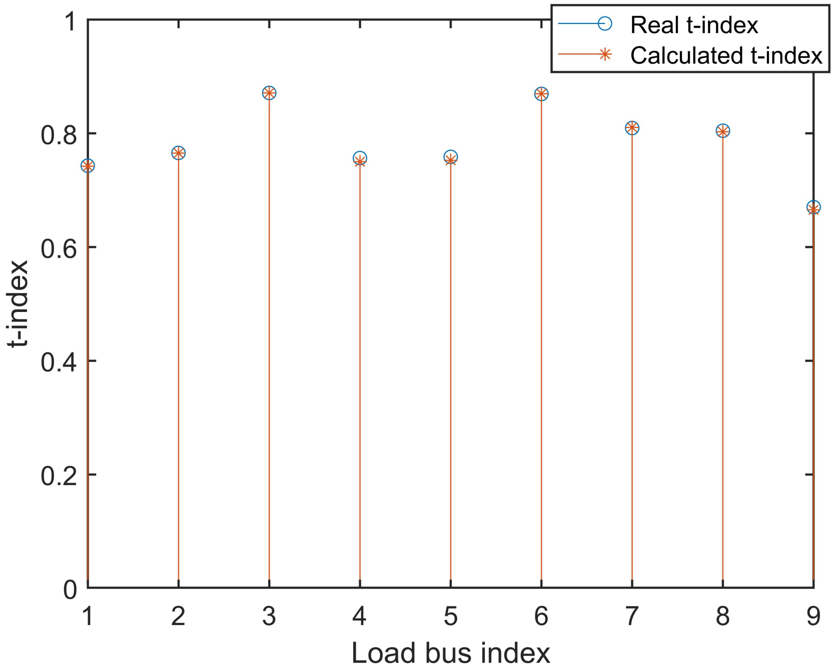
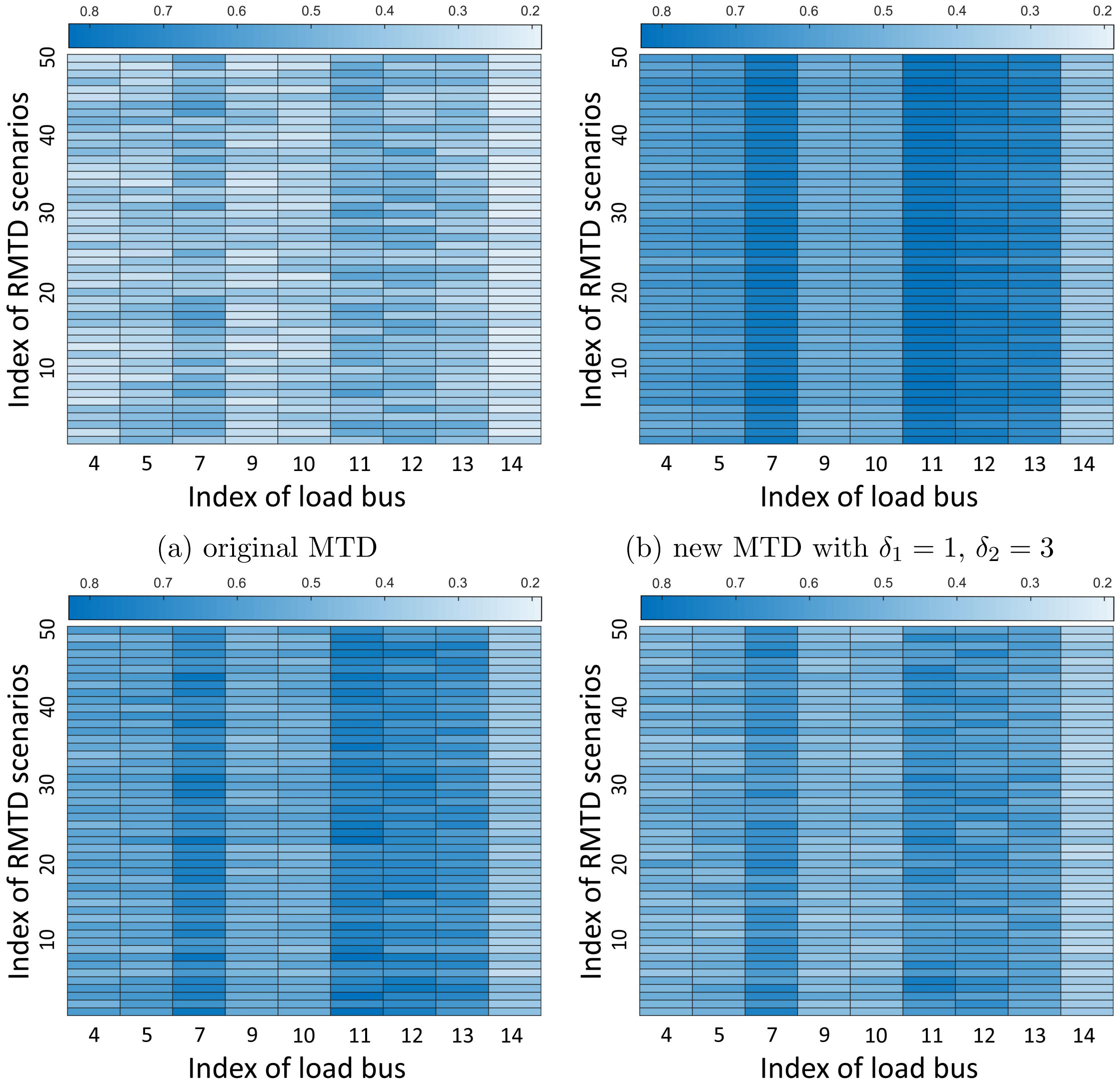


Figure 4.3: TII accuracy in IEEE 14-bus system.

Table 4.1: Mean absolute percentage errors of t-index

RMTD magnitude	0∼0.05	0.05∼0.10	0.10∼0.15	0.15∼0.20
14-bus system	0.028%	0.076%	0.133%	0.180%
37-bus system	0.002%	0.006%	0.010%	0.015%
39-bus system	0.333%	0.916%	1.445%	2.113%
69-bus system	0.056%	0.161%	0.271%	0.320%
118-bus system	0.360%	1.030%	1.549%	2.288%



(a) original MTD

(b) new MTD with $\delta_1 = 1$, $\delta_2 = 3$

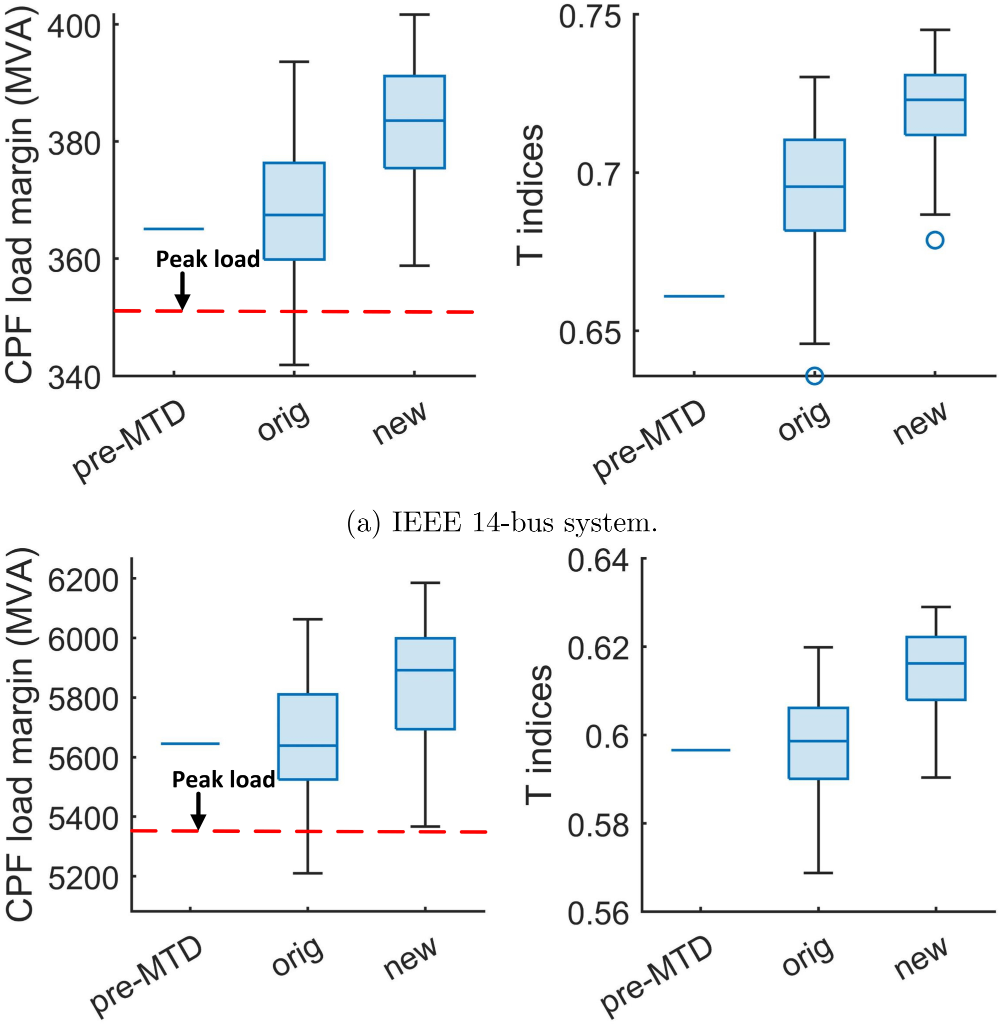
(c) new MTD with $\delta_1 = 1$, $\delta_2 = 1$

(d) new MTD with $\delta_1 = 1$, $\delta_2 = 0.5$

Figure 4.4: Heatmaps of t-indices before and after the t-index optimization method with various δ_2 in the 14-bus system

4.5.2 Impact on Voltage Stability Metrics

To compare and evaluate the performance of the two proposed methods, this section constructs 1000 RMTDs to form a defense pool. The load of the two systems are scaled up by 1.35 times to create a very stressful load condition. Figure 4.4 shows the t-indices of all the load buses in the IEEE 14-bus system obtained under the original MTDs and the new MTDs after the t-index optimization. In this figure, 50 scenarios out of the 1,000 RMTDs with the lowest t-indices are selected and indexed on the y-axis, whereas all load buses of

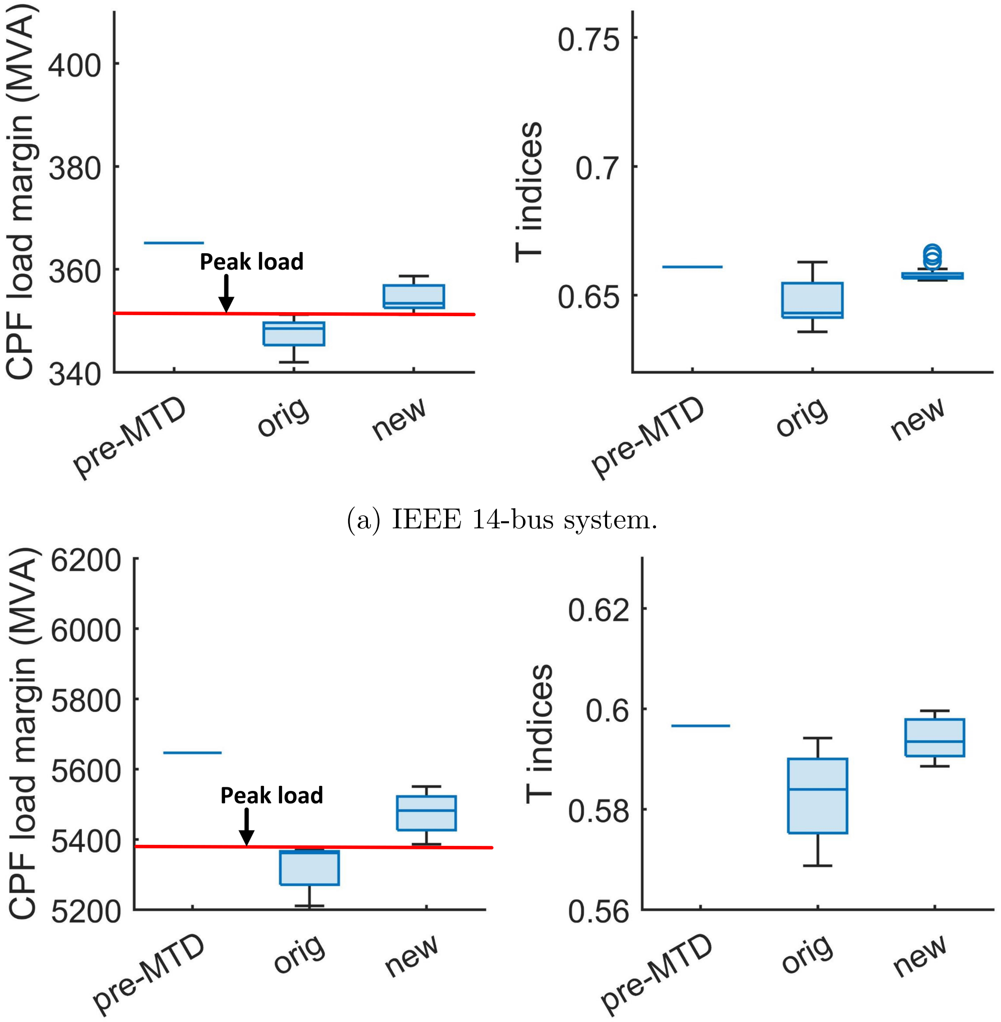


(a) IEEE 14-bus system.

(b) IEEE 118-bus system.

Figure 4.5: Voltage stability metrics before and after t-index optimization method

this system, i.e., Buses 4,5,7,9-14, are plotted on the x-axis. In Fig. 4.4a, the t-indices vary distinctively by scenarios due to the random D-FACTS setpoints of the RMTDs. As seen in Fig. 4.4b, all 50 MTD scenarios result in higher t-indices, signifying better voltage stability after the t-index optimization method is implemented. An intriguing observation in Fig. 4.4b is that the t-index of each load bus obtained are quite similar across all the 50 scenarios. This can be explained by investigating the new D-FACTS setpoints in those scenarios. It is found that the new setpoints X_{new} obtained are similar across the scenarios under $\delta_1 = 1$ and $\delta_2 = 3$. Recall δ_1 and δ_2 are the weights of the MTD adjustment and the t-index lower bound increase in the optimization model (4.9). As δ_2 is comparatively larger



(a) IEEE 14-bus system.

(b) IEEE 118-bus system.

Figure 4.6: Voltage stability metrics before and after load margin constrained method

than δ_1 in Fig. 4.4b, all 50 new MTD scenarios converge to a similar optimal setpoint, where the increase in the t-index lower bound is emphasized. To further explore the impact of both weights on the new MTD, t-indices under two more combinations of the weights are shown in Figs. 4.4c and 4.4d. When δ_2 is decreased to 1 in Fig. 4.4c and further reduced to 0.5 in Fig. 4.4d, the scenario difference in the t-index of each load bus gets larger, suggesting that the t-index optimization problem is more sensitive to the original RMTD setpoints in the scenarios. In addition, the average t-index in Figs. 4.4a is 0.39. It is elevated by the proposed method to 0.66, 0.59, and 0.54 in Figs. 4.4b, 4.4c, and 4.4d, respectively. The percentage increases are 69% , 51%, and 38%, respectively. By examining all the RMTDs in the defense pool, it is found that 16% of the original RMTDs undergo voltage collapse at the peak load. In comparison, all the failed scenarios are saved from voltage collapse by implementing the t-index optimization method.

Figure 4.5 demonstrates the box plot of two voltage stability metrics, i.e., load margin and the minimum t-indices value, for all the load buses in the two systems before and after implementing the proposed t-index optimization method. In each box, the central mark indicates the median, and the bottom and top edges suggest the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points, exclude outliers, and the outliers are plotted individually. Three system states are compared, including the pre-MTD state, the original MTD state, and the new MTD state. It is observed that when the system is transitioning from the pre-MTD to the original MTD state, the CPF load margin and the t-indices may increase or decrease. This is because the RMTDs in the defense pool are constructed randomly without considering the voltage stability. As seen, from the original to the new MTD state, the proposed t-index optimization method elevates both the t-indices and the CPF load margin, indicating increased system voltage stability. A similar trend, shown in Fig. 4.5b can be observed in the IEEE 118-bus system. Two peak loads are labeled as dashed red lines in the load margin figures. These peak loads, which are not used here in Algorithm 1, are added to be consistent with Fig. 4.6. The results in Fig. 4.5 show that the t-index optimization method can promote both of those metrics, which in turn increases the voltage stability of the system.

Analogously, Fig. 4.6 shows the box plots of those voltage stability metrics before and after using the proposed load margin constrained method. According to Algorithm 2, this method only makes adjustment if the system cannot support the forecasted peak load. Therefore, Fig. 4.6 only shows the original MTDs that fail to do so, which is why the body of the box plot in Fig. 4.6 is much shorter than that in Fig. 4.5. In the IEEE 14-bus system, the forecasted peak load is 351.2 MVA labeled by a horizontal red line. In the left plot of Fig. 4.6a, the load margin of the pre-MTD state is 365.1 MVA, which is greater than the forecasted peak load. Hence, the pre-MTD system state is capable of supporting the forecasted peak load. For all RMTDs whose original load margin is less than the forecasted peak load (i.e., "problematic" RMTDs), Lines 5-10 in Algorithm 2 are executed. It is seen in the left plot of Fig. 4.6a that the proposed load margin constrained method significantly brings up the load margin of those problematic RMTDs. As a result, the load margin of all new MTDs are equal to or greater than the forecasted peak load. The right plots in Fig. 4.6a shows the minimum values of t-indices among all the load buses. As seen, the t-indices of the system also increase by using the proposed load margin constrained method. Nevertheless, the improvement is not as significant as that in Fig. 4.5a since the t-indices are not directly maximized in the load margin constrained method. Similar plots for the IEEE 118-bus system are displayed in Fig. 4.6b. The results in Fig. 4.6 demonstrate that the proposed load margin constrained method can significantly increase the load margin of original MTDs and ensure ample load margins to support the forecasted peak load.

PSS/E simulations are further carried out on the IEEE 14-bus system. The dynamic voltage responses of this system under the original MTDs and the proposed load margin-constrained MTDs are compared in Fig. 4.7. As seen at the beginning of the simulation, the system is at an off-peak load without any MTDs. At 1s, both RMTD and the load margin-constrained MTD are implemented. Compared with the RMTD, the load margin constrained MTD decreases the impedance on 11 transmission lines and increases the line impedance on the rest 9 transmission lines. The voltage is stable in both cases after the MTD operations. However, this is not the case when it comes to the peak load (the total load increases by 60% instantaneously) starting from 4s. The system with the original MTD undergoes drastic and short spikes of voltage oscillations and the voltage collapses at 4.6s. Such oscillations indicate that the system generators strive to increase their generation to prevent the system from voltage collapse, but unfortunately they fail to do so due to the insufficient power transfer capability of the transmission lines. In contrast, the system with the load margin-constrained MTD undergoes a smaller voltage drop right after the load increase due to the power mismatch, but the system voltage remains stable at around 0.75 p.u.. Although the voltage magnitude does not meet the ANSI requirement, the proposed method still shows a much better dynamic voltage response, which, in turn, can provide the system operator with sufficient time to implement AC-OPF or dispatch other voltage supporting devices[208;209]. However, this is out of the scope of this chapter and will be studied in the author's subsequent efforts.

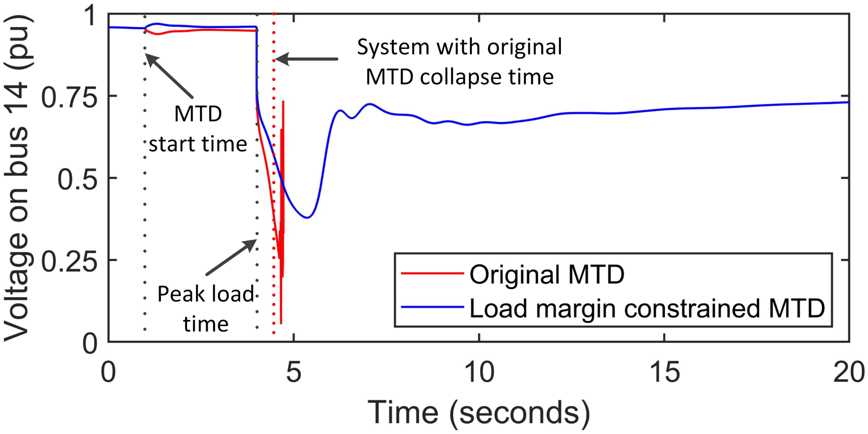


Figure 4.7: Dynamic voltage magnitude response simulated by PSS/E.

4.5.3 Impact on Generation Cost and Attack Detection

This subsection evaluates the impact of the two proposed methods on the system generation cost and MTD performance with various MTD settings. The generation costs in the pre-MTD, load margin constrained MTD, and t-index optimization cases are the corresponding optimal system generation cost by running single-period ACOPF problems on MATPOWER[196], where the quadratic production cost curve of each generator is given. In such problems, the generators are optimally dispatched (i.e., an economic dispatch problem) with the setpoints of the D-FACTS devices given as *a priori*. An exception exists in the case of the OMTD operation (i.e., the third row of Table 4.2), which utilizes a novel ACOPF-based MTD model[52;210] to jointly dispatch the D-FACTS setpoints and the generators while minimizing the system generation cost. Table 4.2 illustrates the system generation costs for the peak load in four cases. The ACOPF in MATPOWER is used to optimally dispatch the generation for each case as illustrated in Fig. 4.2. In Table 4.2, the first case shows the pre-MTD generation cost of the systems, while the second case represents the original-MTD generation cost when an ACOPF-based OMTD[52] is executed. The last two cases show the new-MTD generation costs after each proposed method is implemented. As seen, for both the IEEE 14-bus and 118-bus systems, the lowest generation costs are associated with the OMTD operation in the original MTD state. This is expected since the OMTD operation without considering the voltage stability is solely dedicated to the cost minimization. The second lowest generation costs are pertaining to the new MTD state after the t-index optimization method is implemented. A relatively small cost increase is induced by this method. This is because the t-index optimization method optimally adjusts the original OMTD setpoints to improve the t-indices of load buses and the resulting new MTD setpoints are close to the OMTD ones. The largest generation cost emerges when the load margin constrained method is applied due to much larger MTD setpoint deviation from the OMTD ones. The generation cost results in Table 4.2 show that the load margin method is able to guarantee the system voltage stability at a higher system generation cost. In contrast, the t-index optimization method can ensure the voltage stability with a negligible increase in system generation cost.

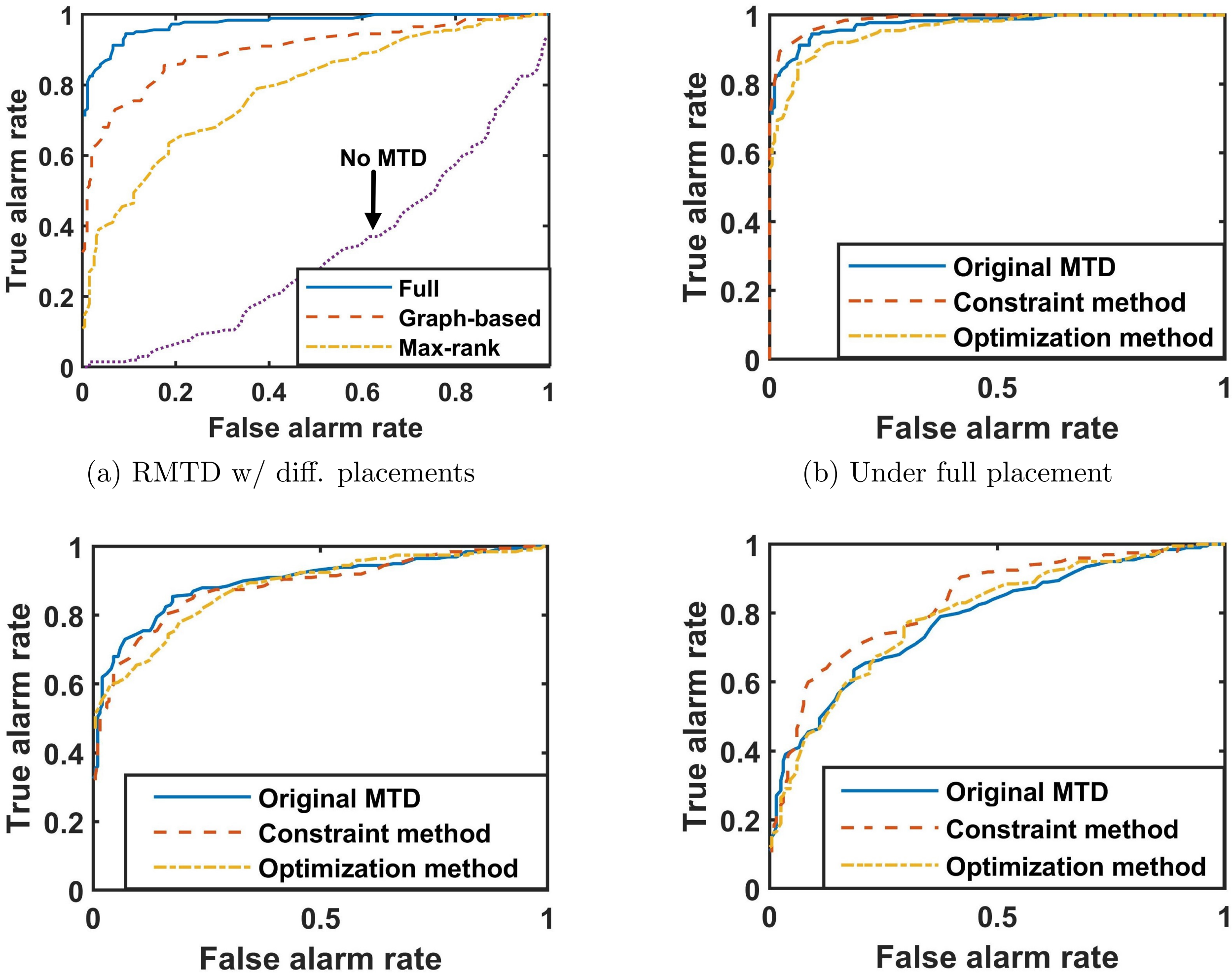
Table 4.2: Comparison of generation costs at the peak load

Cases	14-bus ($/hr)	118-bus ($/hr)
Pre-MTD	8,083.2	129,725.8
OMTD operation	8,076.4	129,714.6
Load margin method	8,358.0	129,906.3
***t*-index method**	8,083.9	129,718.8

Furthermore, simulations are carried out to test the MTD effectiveness against net load redistribution attacks using AC SE-based BDD. Four different D-FACTS placements are considered including zero placement (No-MTD), full placement, max-rank placement[52], and graph-based placement[43]. The measurement noise is assumed to be Gaussian distributed with zero mean and the standard deviation as 1% of the actual measurement. For each D-FACTS placement, this section again constructs 1,000 RMTDs as the corresponding defense pool. Further 1,000 net load redistribution attack vectors are constructed to form an attack pool. Figure 4.8 shows the receiver operating characteristic (ROC) curves of MTDs. These

ROC curves are created by plotting the true positive rate (TPR) versus the false positive rate (FPR) at various BDD thresholds. Figure 4.8a compares the attack detection effectiveness of the original MTD under different D-FACTS placements. As seen, the ROC curve without MTD passes through the bottom right of the graph, leading to the smallest area under the curve (AUC) among all the placement. A smaller AUC indicates a worse performance in attack detection effectiveness. Again, the results in Figure 4.8a demonstrates: 1) the net load redistribution attack is stealthy against AC SE-based BDD; 2) instead of increasing the BDD residual, the net load redistribution attack decreases the residual[21], leading to a smaller TPR than the FPR at a given threshold; and 3) the MTD detection effectiveness varies according to the D-FACTS placement. Theoretical explanation of the MTD detection effectiveness can be found in Section 4.3.1.

Further, this section tests the impacts of the two proposed methods on the attack detection effectiveness under the other three D-FACTS placements, whose attack detection effectiveness is compared in Figs. 4.8b to 4.8d. It is seen that both the load margin constrained method and the t-index optimization method will maintain similar attack detection effectiveness as the original RMTD under the full and graph-based D-FACTS placement. A larger AUC difference between the original MTD and the load margin constrained MTD emerges under the max-rank placement. This can be explained by examining the line impedance change in percentage induced by an MTD, which is indicative of the average absolute MTD magnitude. The average MTD magnitudes of the load margin constrained MTD is 10.42%, which is larger than that of the other two MTDs, i.e., 9.70%. Here, the observation that the attack detection effectiveness increases with the MTD magnitude is consistent with other MTD works[133;211]. The results in Fig. 4.8b to 4.8b indicate that both of the proposed methods can maintain similar attack detection effectiveness as the original MTD. Moreover, by minimizing the MTD adjustment in (4.9), the t-index optimization method has a relatively smaller impact on the attack detection effectiveness performance of the original MTD compared with the load margin method.



(a) RMTD w/ diff. placements

(b) Under full placement

(c) Under graph-based placement

(d) Under max-rank placement

Figure 4.8: ROC curves of BDD residual in IEEE 118-bus system.

4.6 Summary

This chapter addresses a critical issue induced by existing MTDs that myopically perturb the transmission line impedance and result in system voltage instability for varying (net) load. A 3-bus example system is used as an example to illustrate this issue and two methods are further proposed to address it. For the first method, namely the t-index optimization method, this chapter derives the t-index to impedance sensitivity matrix. By utilizing this matrix, this chapter maximizes the lowest t among all the load buses with the minimum impedance adjustment such that the system voltage stability is guaranteed while keeping the performance of the original MTD strategy. The second method, i.e., a load margin constrained method, is developed based on CPF to ensure the load margin is beyond the forecast peak load and thus keeps the system voltage stable during the most stressful time period. Furthermore, this chapter proposes a new MTD framework that seamlessly integrates the proposed two methods.

Extensive simulation results show that both methods can significantly improve the load margin and the voltage stability of a system with an original MTD in critical net load conditions. Moreover, the t-index optimization method can maintain the objectives close to the original OMTDs. The load margin constrained method may induce the new MTD setpoints further away from the original MTD, which is acceptable when RMTD is originally implemented. In reality, system operators can choose either of the two proposed methods to enhance the system voltage stability for RMTDs. When OMTD is originally implemented in the system, a better choice is the t-index optimization method. In future work, the author of this dissertation would like to explore implementing the proposed MTD voltage stability constrained methods under other advanced MTD strategies, including inverter-based MTDs, to change system configurations equivalently. The future work of this chapter will address the challenge that no closed-form metric exists in the AC model to quantify the MTD detection effectiveness directly. A sensitivity analysis will be conducted to approximate the impact of reactance perturbation on the BDD residual.

Chapter 5

Smart Inverter Enabled Coding Scheme for Detecting FDI Attacks in Distribution System State Estimation

The previous section proposes the voltage stability constrained MTD framework as a proactive detection method. However, implementing MTD in large power systems only for cybersecurity purposes is not cost-effective. Although MTD planning strategies[31;40;43] have been proposed to minimize the number of required D-FACTS devices, these methods are only designed for transmission systems. Installing D-FACTS devices in distribution systems is still not economically realistic. This chapter proposes a smart-inverter-based meter encoding framework to detect FDI attacks in distribution systems. The proposed encoding scheme is cost-effective and hidden from alert attackers with both WLS-based SE and PFEDL-based SE.

5.1 Introduction

Existing literature has revealed that if the attackers construct FDI attacks based on the encoded measurements, such attacks can be detected by defenders equipped with SE-based BDD. The majority of current meter encoding focuses on implementing their proposed schemes on conventional meters. Additional ZigBee modules and cellular network communication devices must be deployed for the conventional meters within different substations to support data transmission between meters and encoders. Meanwhile, microprocessors are required to implement meter encoding because conventional meters are not programmable and can only report correct measurements. These additional devices are expensive and can dramatically increase the system's operational cost. A coding scheme is proposed in[46] to increase the BDD residual under FDI attacks by encoding all the measurements with an invertible coding matrix. In[47], an proactive data modification method is proposed to detect manipulation of measurements or control signals. This method encodes all the measurement and control signals with time-varying invertible matrices, respectively. Liu et al.[48] proposed an optimal encoding scheme that considers the cost of meter coding. The strategies above focus on designing invertible coding matrices to encode legitimate measurements. These encoded measurements can be detected by alert attackers who use BDD before launching attacks because the encoded measurements are not consistent with physical laws like Kirchhoff's circuit laws or power flow laws. When attackers detect the unhidden meter encoding, they will not launch attacks until they crack the coding matrices. Trevizan et al.[49] proposed a hidden meter encoding scheme that is undetectable to alert attackers. With their hidden meter encoding, the attackers will not notice that the measurements have been encoded, which means the system operators can better flag the FDI attacks. Although the hidden meter encoding can detect FDI attacks without arousing the attacker's suspicion, a deficiency still exists. The hidden meter encoding can only protect the buses whose measurements are encoded. For FDI attack[116] that targets a small attack area, it will have a chance of not being detected if there is no encoded bus inside the attack area. In this case, the number of protected buses is linearly dependent on the number of encoded buses. Thus, to ensure a high percentage of protected buses in a large system, the encoding cost (i.e., the number of encoded buses) of the hidden meter encoding can be high. Furthermore, existing meter encoding researches only consider transmission systems. How to take advantage of the topology of a radial distribution system to implement low-cost meter encoding is still an open question.

This chapter aims to fill the gap by designing a smart-inverter-based meter encoding scheme in distribution systems to detect FDI attacks. One important consideration here is that the newly designed meter encoding scheme ought to be more cost-effective than the existing ones. More specifically, when protecting the same number of buses, the proposed encoding should require less number of encoded meters compared with existing methods. In addition, unlike the existing meter encoding schemes that require additional devices to encode measurements from conventional meters, the proposed meter encoding should only encode the measurements on inverter buses using the current programmable smart inverters. The programmable smart inverters have been widely used in anomaly detection[212–214], which means that it is achievable to implement the proposed smart-inverter-based meter encoding scheme in distribution systems. The main contributions of this chapter are outlined as follows.

- A smart-inverter-based meter encoding scheme is proposed to detect FDI attacks in distribution systems. The proposed meter encoding can mislead the attacker's SE while not being detected by alert attackers.
- It is proved that if an inverter bus is encoded, all the downstream buses on that lateral will be protected by the proposed smart-inverter-based meter encoding.
- A comprehensive evaluation is constructed to test the detection effectiveness of the proposed smart-inverter-based meter encoding against strong attackers. In this chapter, FDI attackers can obtain the necessary system state using either the WLS-based state estimator or the deep learning-based state estimator.

5.2 Preliminary

This section introduces the background of FDI attacks, their vulnerability, and the SE in distribution systems.

5.2.1 False Data Injection Attacks and Their Vulnerability

As the communication channels between meters in the field and control centers are vulnerable to cyberattacks, attackers can launch man-in-the-middle attacks to eavesdrop and manipulate the measurement signals. FDI attack, as one of the most infamous cyberattacks, can stealthily mislead the power system state estimation by injecting false data into legitimate measurements. To bypass the defender's bad data detector, the attacker needs to elaborately construct the attack vector a by following the equation[99]:

$$a \triangleq h(\hat{x}+c) - h(\hat{x}) \tag{5.1}$$

where $h(\bullet)$ are the measurement functions, $c \in \mathbb{R}^n$ denotes the bias vector that the attacker intends to mislead the state estimation, n is the number of system states, and $\hat{x}$ is the estimated state. When the legitimate measurement $\mathbf{M}$ is altered by the manipulated measurement $\mathbf{M}_a = \mathbf{M}+a$, the BDD residual after the attack will be less or equal to the residual before the attack, i.e., the attack is stealthy.

Although current SE-based BDD cannot detect stealthy FDI attacks, these attacks have one vulnerability: an attacker needs the correct system states to construct attack vectors. By observing the well-known FDI construction equations[21;99;215], it is seen that an attacker is required to know the correct system state x or $\hat{x}$ to derive the stealthy attack vector. However, obtaining the system state information is not effortless for an attacker. Existing research[215;216] assumes that attackers can run local SE to obtain the required state information. In this case, the stealthiness of the constructed attack vector is highly dependent on the accuracy of the attacker's state estimation.

5.2.2 State Estimation in Distribution Systems

In an AC SE using noisy measurement, the power flows are non-linearly dependent on the system states. The non-linear dependencies can be mathematically expressed as follows:

$$\mathbf{M} = h(x) + E \tag{5.2}$$

where $\mathbf{M} \in \mathbb{R}^m$ is the vector of measurement data, including bus power injection and line power flow. m is the number of measurements. $x \in \mathbb{R}^n$ is the vector of the system state. E is the measurement noise vector, which is usually assumed to obey a Gaussian distribution. $h(\bullet)$ are the non-linear measurement functions that reveal the relationship between the measurement and the system state. The WLS-based SE mechanisms determine the system state by solving the following problem:

$$\min_X J(x) = [\mathbf{M} - h(x)]^T \mathbf{W} [\mathbf{M} - h(x)] \tag{5.3}$$

where $\mathbf{W}$ is the weight matrix given as the inverse of the covariance matrix of measurement noise. The most common approach to solve (5.3) is the iterative procedure[217].

State estimation is one of the most fundamental tasks in power systems. From a cybersecurity point of view, system operators need the SE-based BDD to detect bad data or FDI attacks; FDI attackers require the estimated state to construct stealthy FDI attack vectors as shown in (5.1). One difficulty of implementing WLS-based SE in distribution systems is that redundant measurement is necessary to satisfy the observability requirement. However, limited real-time measurements are available in distribution systems, which makes (5.2) under-determined. To address this issue, low-observability SE techniques have been proposed. Some methods attempt to improve system observability by optimally placing additional meters[218;219], or deriving pseudo-measurements from historical data[220;221]. Other methods propose to replace the conventional WLS-based state estimator with new state estimators, including matrix completion-based estimator[222], approximation-based estimator[223], and machine learning-based estimator[194]. In this chapter, besides the WLS-based SE, PFEDL-based SE is utilized by both attackers and defenders to address the low-observability issue in distribution systems.

5.3 The smart-inverter-based Meter Encoding

This section proposes to use smart inverters as encoders to report encoded measurements so that only system operators will get the correct measurements. In contrast, attackers will get the wrong measurements. The proposed encoding method is inspired by symmetric cryptography, where the same measurement bias (encryption key) is used to encrypt and decrypt the measurement data. After encoding the output data from smart inverters, the encoded measurements and legitimate measurements from conventional meters will be transmitted through the communication channels. If the attackers construct a formerly stealthy FDI attack based on the encoded measurement, their attack will induce BDD residuals large enough to trigger the BDD alarm.

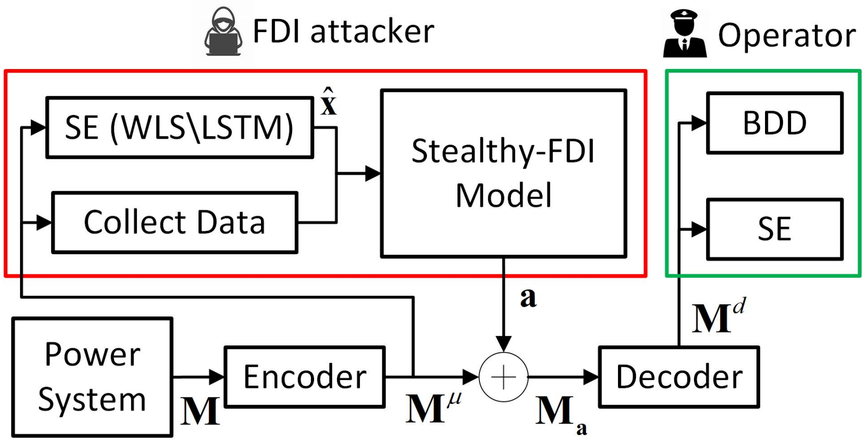


Figure 5.1: Flowchart of the smart-inverter-based meter encoding.

The flowchart of the proposed smart-inverter-based meter encoding against stealthy FDI attacks is shown in Fig. 5.1. The measurement vector $\mathbf{M}$ from the physical power system is processed by the proposed encoding scheme before being transmitted by the communication network. Given a secret encoding vector μ and the encoding function $f(\mathbf{M}, \mu)$, the output

$\mathbf{M}^{\mu}$ of an encoder consists of the encoded measurements from smart inverters and legitimate measurements from conventional meters. Both encoder and decoder have access to the secret encoding vector μ, which is the bias between the encoded and original measurements. The encoding function f in the encoder adds the encoding vector to the original measurements from smart inverters. Correspondingly, a decoding function g substracts the encoding vector from the encoded measurements. After receiving $\mathbf{M}^{\mu}$, the control center will decode the received data by using the decoding function $g(\mathbf{M}^{\mu}, \mu)$. The decoded measurements $\mathbf{M}^{d}$ will be tested by the state estimation-based BDD first to check if the measurements contain bad or manipulated data. When the received measurement passes the BDD, the measurement and the corresponding estimated system state will be used by an energy management system for advanced power system functions, including contingency analysis, optimal power flow, automatic generation control, etc. The recovered measurements $\mathbf{M}^{d}$ will be the same as the original measurements $\mathbf{M}^{d} = \mathbf{M}$ if there is no FDI attack.

$$\mathbf{M}^{\mu} = f(\mathbf{M}, \mu) \tag{5.4a}$$

$$\mathbf{M}^{d} = g(\mathbf{M}^{\mu}, \mu) \tag{5.4b}$$

The encoding function (5.4a) and the decoding function (5.4b) would be implemented at the smart inverter and the control center, respectively. Existing researches widely assume that attackers can bypass data protection strategies implemented within the communication network, e.g., packet encryption, and eavesdrop on data-in-flight between the power system and the control center. In this case, the attackers will get the wrong (encoded) measurements. How to ensure that alert attackers will not suspect the encoded measurements is still an open question.

5.3.1 Meter Encoding Construction and Hiddenness

When alert attackers obtain $\mathbf{M}^{\mu}$ from the communication channel, they will run their SE-based BDD before constructing stealthy FDI attack vectors $\mathbf{a}$ based on their estimated system state. To ensure an alert attacker does not suspect the measurements are encoded, $\mathbf{M}^{\mu}$ should be elaborately encoded by the system operator to be consistent with power flow laws. This process can be viewed as a defender constructing a stealthy FDI attack by encoding the measurement from smart inverters to mislead an attacker's SE. A hidden meter encoding $\mathbf{M}^{\mu} = \mathbf{M} + \mu$ can be achieved by constructing an encoding vector μ, similar to a stealthy FDI attack vector:

$$\mu = h(\hat{x} + B) - h(\hat{x}) \tag{5.5}$$

Where $\hat{x}$ is the correct estimated state, and B is a bias between the attacker's estimated state and the true system state, which is defined as the encoding magnitude. Let $x_{\mu} = \hat{x} + B$ denotes the false state vector that the defender intends to mislead the attackers after meter encoding. When alert attackers run their SE-based BDD with the encoded measurements, the residual after the meter encoding r_{μ} remains the same as the original residual r without encoding:

$$\begin{aligned} r_{\mu} &= \|\mathbf{M} + \mu - h(x_{\mu})\|_2 \\ &= \|\mathbf{M} + h(\hat{x} + B) - h(\hat{x}) - h(x_{\mu})\|_2 \\ &= \|\mathbf{M} - h(\hat{x})\|_2 \\ &= r \end{aligned} \tag{5.6}$$

Thus, the proposed meter encoding is hidden from alert attackers.

5.3.2 Smart-inverter-based Meter Encoding

In reality, it is cost-prohibitive to implement existing meter encoding schemes on conventional meters in the field. This chapter proposes only using smart inverters to implement the encoding schemes to address this issue. Unlike the existing meter encoding methods that can only protect the encoded buses, the proposed smart-inverter-based encoding can protect all the downstream buses by encoding one inverter bus, which will be explained in Section 5.3.3. In this case, the proposed encoding scheme extraordinarily decreases the number of encoded meters when protecting multiple buses, i.e., decreases the defense cost.

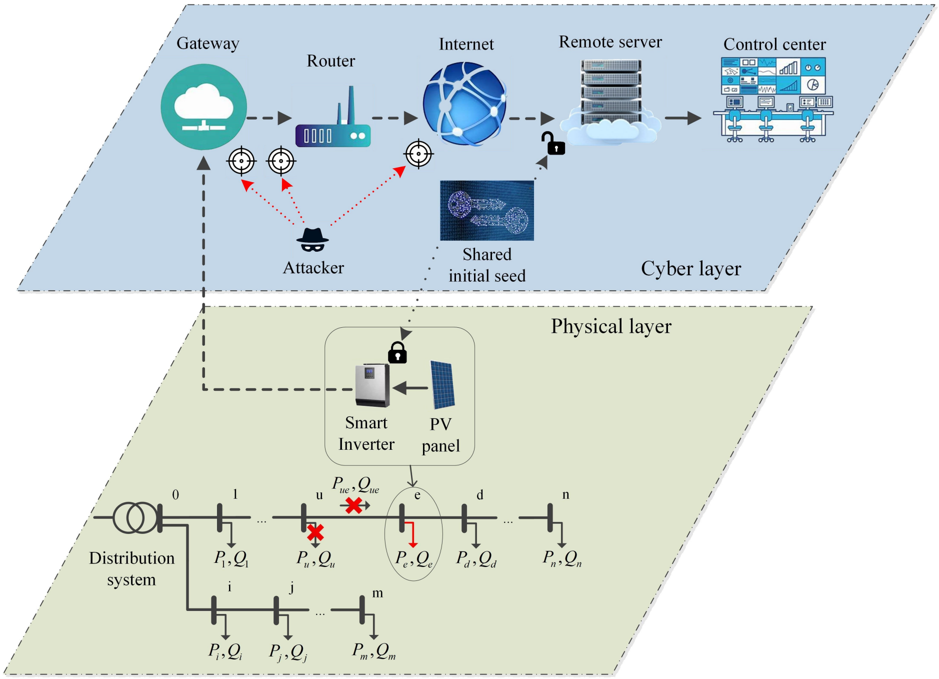


Figure 5.2: Cyber-physical configuration of a distribution system.

The cyber-physical configuration of a distribution system with the proposed smart-inverter-based encoding is given in Fig. 5.2. Assuming a smart inverter is installed at Bus e in a radial distribution system. The solid lines going into the smart inverter and control center represent that the correct measurements are transmitted, while the dashed lines denote that the transmitted measurements contain encoded data. Nowadays, smart inverters are usually programmable[212;224]. Thus defenders can use them to implement the proposed encoding scheme. Instead of sending the actual measurements, the smart inverter sends encoded measurements, which will be transmitted along with the legitimate measurements from other meters. After the remote server receives the encoded measurements, a decoder will be activated to restore $\mathbf{M}^d$ and send the decoded measurements to the control center. To ensure the decoder can correctly restore the encoded measurements, pseudo-random generators[225] can be implemented in both the encoder and the decoder, where the same pseudo-random sequences can be generated using a shared initial seed. In this configuration, when attackers construct FDI attacks, they will eavesdrop on the measurements from the cyber layer by implementing man-in-the-middle attacks[72].

This chapter considers the attackers know the line parameters in a part of a power system. Therefore, the attackers can launch local FDI attacks after implementing their local state estimation with the eavesdropped measurements[215;216]. To prevent attackers from obtaining the correct system state to construct stealthy FDI attacks, encoding vector μ is added to the legitimate smart inverter measurements. The easiest way to construct the encoding vector is to randomly choose the magnitude of μ. However, alert attackers can detect such arbitrary encoding vectors by implementing their state estimation-based BDD. When the attackers identify a large mismatch between the system model and the measurements encoded by an arbitrary encoding vector, they will suspect the measurements are encoded and may devise ways to circumvent the meter encoding. This section proposes the following encoding functions (5.7) to (5.10), which the system operators can use to remain hidden from alert attackers. The encoding vector consists of real and reactive power injections P_e^* and Q_e^* at a smart inverter bus (denoted as the red right-angled line at Bus e in Fig. 5.2).

$$P_{eu}^* = P_e^* - \sum_{d \in \mathcal{N}} P_{ed} \tag{5.7}$$

$$Q_{eu}^* = Q_e^* - \sum_{d \in \mathcal{N}} Q_{ed} \tag{5.8}$$

$$P_{eu}^* = -(V_e^*)^2 g_{eu} + V_e^* V_u (g_{eu} cos(\theta_e^* - \theta_u) + b_{eu} sin(\theta_e^* - \theta_u)) \tag{5.9}$$

$$Q_{eu}^* = (V_e^*)^2 g_{eu} + V_e^* V_u (g_{eu} sin(\theta_e^* - \theta_u) - b_{eu} cos(\theta_e^* - \theta_u)) \tag{5.10}$$

where e denotes the encoded inverter bus, u and d are the corresponding upstream and downstream buses. $\mathcal{N}$ is the set of all the directly connected downstream buses. The superscripts $(*)$ represent the variables that the meter encoding will impact. These variables include the encoded measurements and the system state that the system operators intend to mislead the attackers. The target of the encoding scheme is to mislead the attackers into believing the wrong system state V_e^*, θ_e^* instead of the true system state V_e, θ_e. By calculating (5.7)-(5.10), the encoded measurements $P_{eu}^*, Q_{eu}^*, P_e^*, Q_e^*$ can be derived, given V_e^*, θ_e^*.

Similar to the stealthy FDI attacks, the general rule for a hidden encoding scheme is that the defender encodes the transmitted data so that the measurements are consistent with the power flow laws. The alert attackers will detect the encoding scheme if there is no feasible power flow solution after meter encoding. The proposed smart-inverter-based meter encoding requires that some measurements should not be measured, including the upstream bus power injection measurements P_u^*, Q_u^* and the power flows P_{eu}^*, Q_{eu}^* (denoted as the red crossings in Fig. 5.2) on the transmission lines that connect the invert bus and the upstream bus. This requirement is consistent with the fact that distribution systems are not fully measured. In case such upstream bus injection and power flows are measured, additional effort is needed for the defender to encode these data to make the attacker's eavesdropped measurements consistent with Kirchhoff's circuit laws. Consider the example demonstrated in the physical layer of Fig. 5.2 for clarification. The goal of the system operator is to mislead the attacker's SE by changing the perceived bus power at Bus e while not arousing the attacker's suspicions, i.e., bypassing the attacker's BDD. Assuming the system operators achieve this by using (5.7)-(5.10), they will also need to process the related measurements at Bus u and on the transmission line $u - e$ to hide the encoding.

5.3.3 Protected Region

The proposed meter encoding can protect a region of buses by utilizing the topology of a radial distribution system. The definition of a protected region is given as follows. In a radial distribution system, the region that consists of an encoded inverter bus and all its downstream buses is the protected region of the meter encoding. For clarity, let us assume there is a smart inverter on a lateral of a radial distribution system, as shown in Fig. 5.2. The system operator chooses to encode the measurement from this smart inverter by changing the estimated state x_e on the inverter Bus e as discussed in Section 5.3.2. In this case, all the buses on such lateral after the smart inverter bus is inside a protected region.

In a protected region, the proposed meter encoding scheme can mislead the system state estimated by attackers. This conclusion can be proved by analyzing the power flow equations inside the protected region because SE results are usually nearly identical to the power flow solutions if the state estimator works properly. Since the power flow measurements on transmission line $e-d$ are not encoded, the misled estimated states at Bus d can be calculated by solving the power flow equations on this transmission line, given the misled system states $\hat{V}_e^*, \hat{\theta}_e^*$ (as described in Section 5.3.2) and the correct power flow measurements P_{ed}, Q_{ed}.

$$\hat{V}_e^{*2} g_{ed} - \hat{V}_e^* \hat{V}_d^* (g_{ed} cos(\hat{\theta}_e^* - \hat{\theta}_d^*) + b_{ed} sin((\hat{\theta}_e^* - \hat{\theta}_d^*)) = P_{ed} \tag{5.11}$$

$$-\hat{V}_e^{*2} b_{ed} + \hat{V}_e^* \hat{V}_d^* (b_{ed} cos(\hat{\theta}_e^* - \hat{\theta}_d^*) - g_{ed} sin((\hat{\theta}_e^* - \hat{\theta}_d^*)) = Q_{ed} \tag{5.12}$$

Because all the power flow measurements inside the protected region are not encoded, the misled system states on all the downstream buses (from Bus d to the end Bus n of this lateral) can be calculated accordingly, similar to (5.11) and (5.12). Then, these incorrect states will be used by attackers to construct FDI attacker vectors.

5.4 Experiment Results

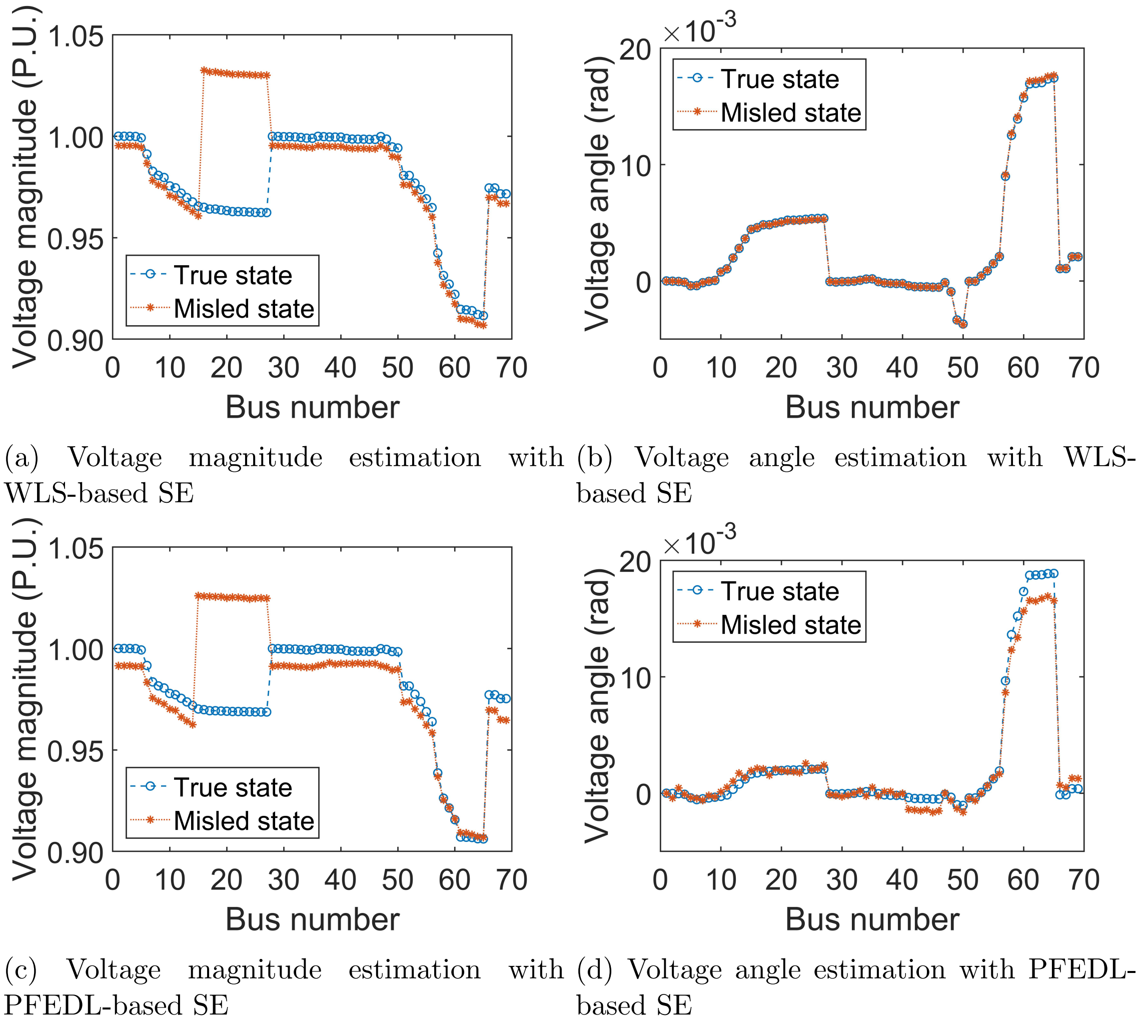
This section evaluates the detection effectiveness of the proposed smart-inverter-based meter encoding against stealthy FDI attacks. The simulations are implemented on the IEEE 69-bus system using the open-source MATLAB tool MATPOWER[196]. The WLS-based SE and BDD are implemented on MATLAB, while the machine learning-based SE is performed using the Python package Pytorch[226]. To train and test the attacker's machine learning-based state estimator, PV generation and one-year load profiles are obtained from ERCOT[227] and eGauge[228], respectively. Among the hourly measurement data, the first 7760 measurement vectors are used to train the attacker's PFEDL model, and the rest measurement vectors are used to construct and test FDI attacks. In this case study, both the training and testing measurements are encoded by the proposed meter encoding scheme. It is assumed that all the measurements contain Gaussian noise with a standard deviation of 2% of the accurate measurement value. The chi-square statistic tests are performed with a 95% confidence level for the SE-based BDD.

This chapter mainly focuses on proposing a low-cost and hidden smart-inverter-based meter encoding scheme to detect FDI attacks in distribution systems. The state estimator observability issue is out-of-scope for this chapter. The case studies are concentrated on an observable distribution system under stealthy FDI attacks. The distribution system is hereinafter assumed to be single-phase balanced.

5.4.1 Protected Region, Encoding Hiddenness, and Encoded Costs

This sub-section evaluates the protected region consisting of all the buses on which the attacker's SE is misled by the meter encoding. Bus 15 is chosen as the encoded bus equipped with a smart inverter. According to the definition of the protected region, the proposed meter encoding will protect the encoded Bus 15 and its downstream Buses 16-27. After a defender implements the encoding process, the attacker's estimated system state based on the eavesdropped measurements is shown in Fig. 5.3. Figures 5.3a and 5.3b illustrate the scenario in which the attackers use the WLS-based state estimator, while Figs. 5.3c and 5.3d demonstrate the scenario where the attackers use the PFEDL state estimator. In these scenarios, the system operator aims to mislead the attackers' state estimation by an encoding magnitude B equal to 0.065 p.u. It is seen that the biases between the true voltage magnitudes and the attackers' estimated voltage magnitudes are around 0.065 p.u. as expected by the defender. The misled state estimation verifies that a protected region consists of the encoded inverter bus and all the downstream buses.

In addition, the proposed encoding is hidden from attackers when they implement their SE-based BDD test on the encoded measurements. This section runs 1,000 BDD tests from an attacker's perspective. The BDD residuals of the encoded measurements are identical to the residuals of the legitimate measurements without meter encoding, as shown in Fig. 5.4. This result indicates that if the legitimate measurements can pass the attacker's BDD, the encoded measurements can also bypass the same BDD, i.e., the proposed meter encoding is hidden from alert attackers. The hiddenness of the proposed meter encoding scheme is a unique benefit as existing meter encoding schemes[46–48] can be detected by an alert attacker.



(a) Voltage magnitude estimation with WLS-based SE

(b) Voltage angle estimation with WLS-based SE

(c) Voltage magnitude estimation with PFEDL-based SE

(d) Voltage angle estimation with PFEDL-based SE

Figure 5.3: Attacker's estimated state with/without meter encoding

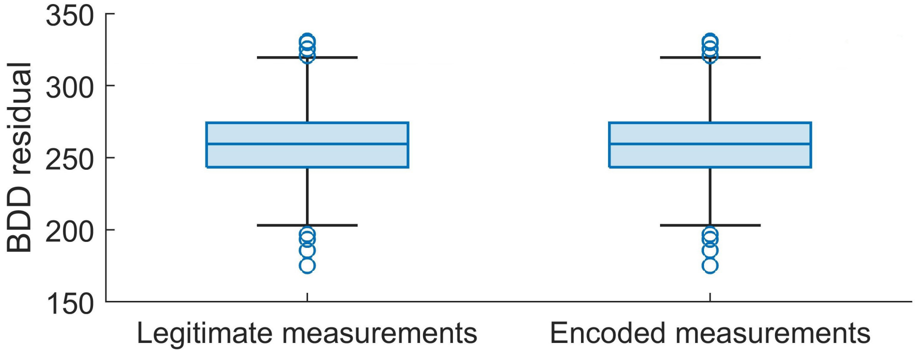


Figure 5.4: Attacker's BDD residuals after meter encoding

For the above example in Fig. 5.3, two measurements (the real and reactive power injections at Bus 15) are required to be encoded to protect the 13 buses in the protected

Table 5.1: Comparison of meter encoding costs

Encoding schemes	Proposed scheme	Existing encoding[49]	Existing optimal encoding[48]
No. of encoded meters per one protected bus	0.15	3.86	1.93

region. As a comparison, the meter encoding scheme proposed in[49] requires the defender to encode 54 measurements to protect 14 buses. An optimal encoding scheme proposed in[48] needs to encode 27 measurements to protect 14 buses. The average required encoded meters to protect one bus is compared in Table 5.1. The comparison result shows that the proposed meter encoding scheme has the lowest encoding cost.

5.4.2 Detection of FDI Attacks

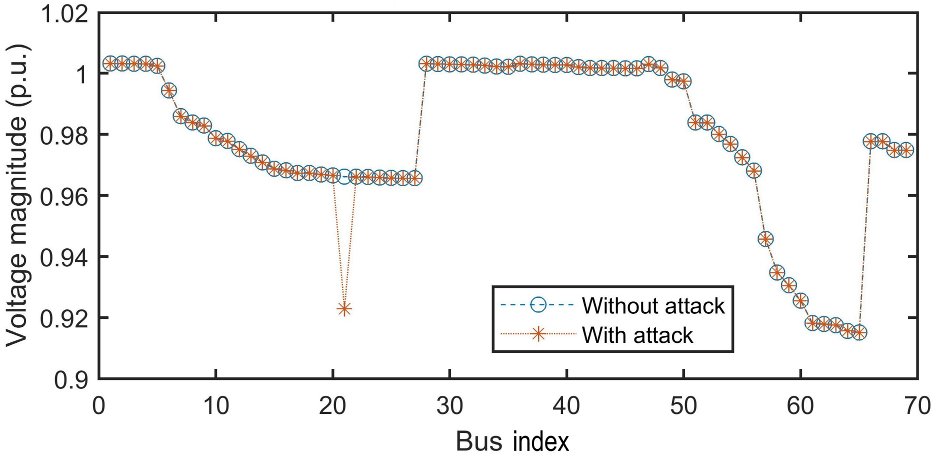


Figure 5.5: FDI attack on voltage magnitude estimation

This sub-section evaluates the detection effectiveness of the proposed meter encoding in detecting stealthy AC FDI attacks. Suppose the attackers aim to decrease the estimated voltage magnitude at bus 21 by 0.04 p.u. To achieve this, the attackers can obtain $\hat{x}$ by using the WLS or machine learning-based state estimators after eavesdropping on the original measurements. By solving (5.1), the attack vector can be constructed, given c equals 0.04 p.u. Without the proposed meter encoding, the FDI attack will mislead the system operator's

SE at Bus 21, as shown in Fig. 5.5. This attack is proved to be undetectable for conventional SE-based BDD[99].

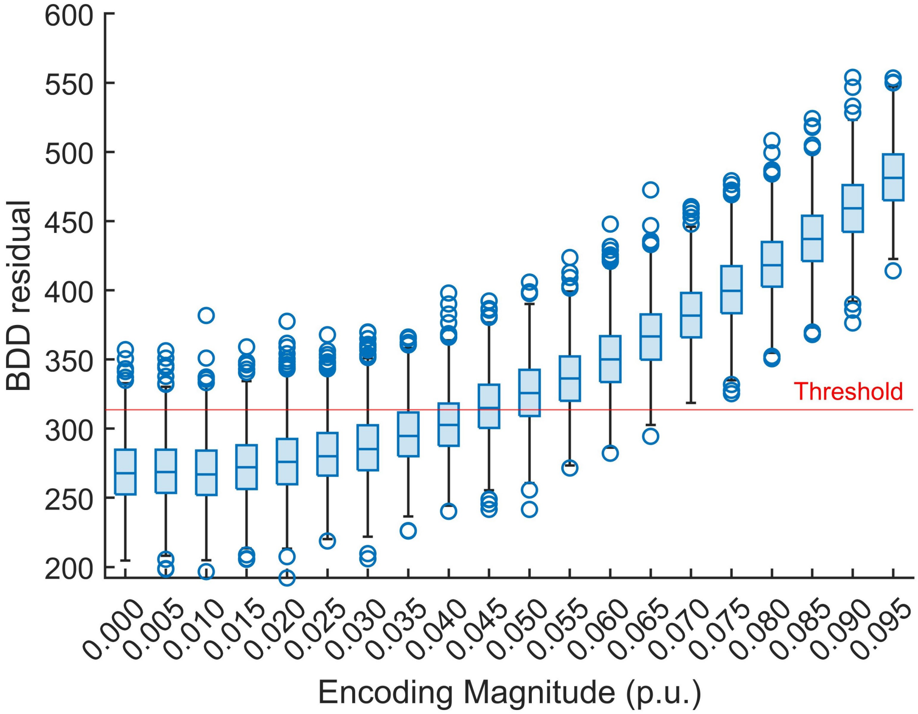


Figure 5.6: BDD residuals with meter encoding after stealthy FDI attacks

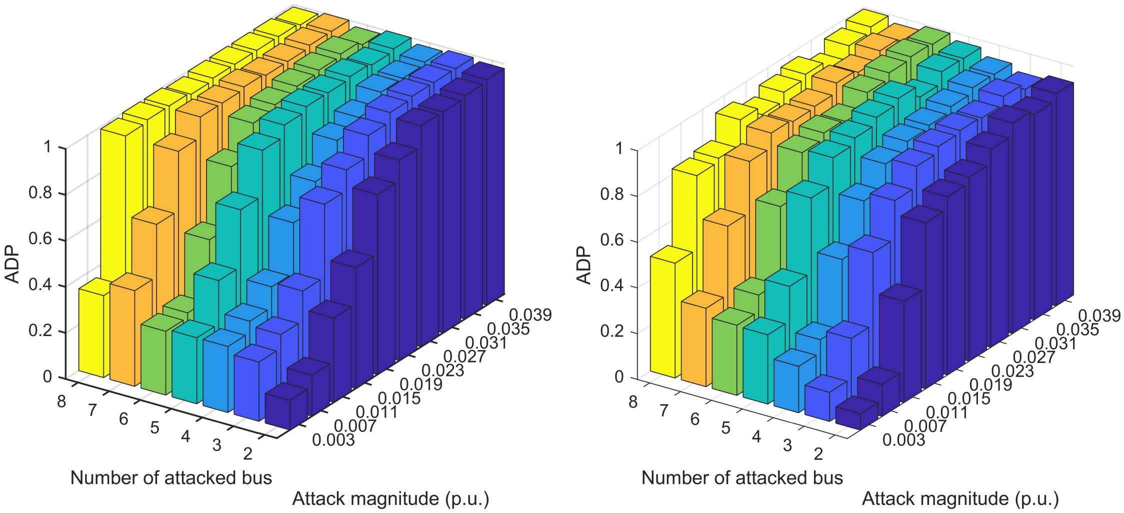
AC SE-based BDD is used from a system operator's perspective to evaluate the detection effectiveness of the proposed smart-inverter-based meter encoding. Bus 15 is the inverter bus whose measurements are encoded. After a BDD test, a residual can be derived to indicate if an attacker has manipulated the received measurements. In this chapter, the BDD threshold is chosen to be 313 with a 95% confidence level. If the BDD residual exceeds the threshold, the system operator can declare a 95% probability that the received measurements are manipulated. The box plot in Fig. 5.6 shows the BDD residuals after an attacker launches FDI attacks under different encoding magnitudes $B = V_{15}^{*} - V_{15}$. These attack vectors are constructed based on the encoded measurements and the corresponding system states estimated by the attacker's WLS SE. The attack magnitude (recall 1.3) is $c = 0.05$p.u. at bus 20. On the x-axis of Fig. 5.6, various encoding magnitudes are tested to evaluate the detection capability of the proposed encoding scheme. The first box from the left illustrates that when attackers construct FDI attacks using their estimated states from the legitimate

Table 5.2: Attack detection effectiveness using the proposed meter encoding

Encoding magnitude (p.u.)	ADP	Percentage of ADP increase
0.00	0.038	0.00%
0.01	0.042	10.53%
0.02	0.059	55.00%
0.03	0.142	273.42%
0.04	0.354	830.53%
0.05	0.640	1685.26%
0.06	0.948	2395.00%
0.07	0.999	2528.95%

measurements (i.e., encoding magnitude is 0), the attacks are stealthy to SE-based BDD. As a comparison, when the attacker's eavesdropped measurements are encoded with increasing encoding magnitudes, the constructed FDI attacks will increase the BDD residual, as shown in Fig. 5.6. When the encoding magnitude is greater than 0.07 p.u., all the FDI attacks are detectable in this example.

In this chapter, the attack detection probability (ADP) is introduced to evaluate the effectiveness of the BDD with the help of the proposed encoding scheme. The ADP is defined as the ratio of the number of detected attacks to the total number of attacks. An ADP value of 1 means all the attacks are detected by BDD, and an ADP value of 0 means no attack is detected. Using the ADP of the BDD without meter encoding as a reference, Table 5.2 lists the ADP with various encoding magnitudes and the corresponding improvement in ADP. Recall that the encoding magnitude is defined as the bias between the attacker's estimated state and the true system state. For each scenario, 1,000 attack vectors are constructed based on the encoded measurements and are tested by BDD. The improvement in the BDD detection effectiveness depends on the encoding magnitude. A larger encoding magnitude leads to greater inconsistency between the attack vector and the physical power flow laws and thus increases the BDD residual and the ADP. From Table 5.2, it is recommended to use a larger encoding magnitude to provide higher detection effectiveness. Although using a large encoding magnitude for defenders does not increase their defense cost (as the meter encoding is a software-based method), a larger encoding magnitude may result in a significant measurement change and raise an attacker's suspicion. In this case, an alert attacker may suspend the attack and invest more resources to obtain the correct measurements, which further increases the cybersecurity risk of the smart grid. Therefore, how to optimally choose the encoding magnitude to balance the trade-off between the detection effectiveness and the encoding magnitude is an important topic, which will be addressed in future work.



(a) WLS-based FDI

(b) PFEDL-based FDI

Figure 5.7: ADP of the proposed meter encoding against FDI attacks under different attack parameters

A fixed encoding magnitude $B = 0.065$ p.u. is used in this simulation to investigate the ADP under different numbers of attacked buses and various attack magnitudes. The ADP of the proposed encoding scheme against the attackers who use the WLS-based SE to construct their stealthy FDI is shown in Fig. 5.7a. The simulations are run in 70 cases with various numbers of attacked buses and attack magnitudes. In each case, 1,000 attack scenarios are constructed based on the encoded measurements and the corresponding attacker's estimated states. The ADP for each case is calculated using the BDD residuals after the 1,000 attack scenarios. As shown in the x- and y- axis, a range of 2 to 8 buses in the protected region is attacked, while the attack magnitude changes from 0.003 p.u. to 0.039 p.u. It is seen that with the increase of attack magnitudes, the ADP of the proposed encoding scheme also increases. This observation is true for both the FDI attacks constructed based on the attacker's WLS state estimator (Fig. 5.7a) and the PFEDL state estimator (Fig. 5.7b). When the attack magnitude reaches 0.023 p.u., the ADPs are over 90% regardless of which state estimator the attacker uses.

5.5 Summary

This chapter proposes a smart-inverter-based meter encoding scheme to detect stealthy FDI attacks in single-phase distribution systems. The proposed scheme can protect all the downstream buses against FDI attacks by encoding a bus equipped with a programmable smart inverter, which decreases the encoding cost compared with existing meter encoding strategies. The case study has demonstrated that the proposed scheme can mislead the attacker's SE on the protected buses and induce the attack vector to trigger the defender's BDD. A comprehensive evaluation, which considers alert attackers equipped with WLS-based SE and PFEDL-based SE, has been conducted to test the proposed scheme's detection effectiveness against FDI attacks. In addition, the case study shows the proposed low-cost meter encoding is hidden from alert attackers who can implement SE-based BDD to detect arbitrary defense strategies.

Chapter 6

Conclusion and Future Work

This chapter concludes the dissertation by summarising the key research findings in relation to the research targets and questions and discussing directions for possible future work. Chapter-specific conclusions are included in their respective chapters.

6.1 Conclusion

In this dissertation, the cyber-physical security of smart grids is comprehensively reviewed. After investigating the impact of various cyber-physical attacks, this dissertation focuses on the infamous data integrity attack, namely FDI attacks. With the provided background, this dissertation address three major research questions that aim to enhance the detection of FDI attacks in distribution systems. Therefore, a novel NLRA as a more realistic FDI attack is proposed to mislead the system operator's SE. Using NLRA as an adversary, a voltage stability constrained MTD framework is proposed to detect such attacks while ensuring long-term voltage stability. Further, a low-cost meter encoding method that does not require additional hardware devices is explored. The answers to these research questions can help distribution systems enhance cybersecurity and improve situational awareness. The accomplishments of this dissertation are summarized as follows.

Chapter 3 proposes an NLRA framework, which targets to mislead system operator's SE in distribution systems with DERs. By following the boundary conditions, attackers can launch NLRA with limited configuration information inside an attack region. Further, deep neural network-based SE is utilized by attackers to get the system state for NLRA construction under low observable distribution systems where conventional WLS-based SE cannot converge. The numerical results show that the proposed NLRA is stealthy to the defender's BDD and can mislead DSO with illusory under-voltage issues.

In Chapter 4, a long-term voltage stability issue under existing MTD strategies is revealed. Steady-state and dynamic simulations show that when MTDs perturb line impedance, the system load margin will be degraded and may induce voltage instability in various power systems. To address this problem, a voltage-stability-constrained MTD framework is proposed to detect FDI attacks while ensuring voltage stability. Two methods, i.e., t-index optimization and load margin constrained method, are utilized to re-dispatch the original MTDs. The NLRA model from Chapter 3 is used to evaluate the detection effectiveness of the proposed MTD framework. Extensive simulation results show that both methods can significantly improve the load margin and the voltage stability of a system with an original MTD in critical net load conditions. Moreover, the t-index optimization method can maintain the objectives close to the original OMTDs.

Finally, Chapter 5 proposes a smart-inverter-based meter encoding method to detect FDI attacks in distribution systems. The advantage of this method over MTD is that no additional D-FACTS devices are needed. Meanwhile, since meter encoding is a software-based detection method, it will not induce system stability issues.

6.2 Future Work

This section summarizes potential research directions. Chapter 3 improves the existing FDI attacks so they can be launched in distribution systems. Interesting future work includes developing defense frameworks in which the proposed NLRA will be simulated, studying attack sequences with a high level of DERs, and investigating other attack goals in distribution systems. A comparison of existing PFEDL-based SE (LSTM and DNN) and pure data-driven SE would be an exciting area of research. Extending the NLRA model to three-phase unbalanced systems could be another area of further inquiry.

Chapter 4 considers voltage stability constrained MTD framework. This work can be improved in the future by implementing the proposed voltage stability constrained MTD methods under other advanced MTD strategies. A machine learning model to address load margin uncertainty during peak load could be another extension of this work. Further, work can be done to address the challenge that there exists no closed-form metric in the AC model to quantify the MTD detection effectiveness directly. A sensitivity analysis can be conducted to approximate the impact of reactance perturbation on the BDD residual.

Chapter 5 investigates the viability of a smart-inverter-based meter encoding that detects FDI attacks in distribution systems without installing additional hardware devices. The encoded measurements include load injections and power flows. However, phase measurements from the most advanced phasor measurement units (PMUs) are not considered. Improving the proposed detection scheme by encoding PMUs could be interesting future work. Further, this work can be better evaluated on a hardware-based testbed consisting of measurement, control, and communication functions.

This dissertation studies cyber-physical security in smart grids. It is necessary to investigate attack detection, identification, and mitigation under more sophisticated adversary models in the future. In addition, investigating the cooperation of many interdependent detection methods could be vital in improving the system operator's situational awareness. Further, investigating the attack and defense strategies in dynamic emulations is worthwhile to enhance the robustness and resilience of smart grids.

Bibliography

[1] Cyber-Physical Systems Public Working Group et al. Framework for cyber-physical systems: Volume 1, overview, version 1.0. *NIST Special Publication*, pages 1500–201, 2017.

[2] Haibo He and Jun Yan. Cyber-physical attacks and defences in the smart grid: a survey. *IET Cyber-Physical Systems: Theory Applications*, 1(1):13–27, 2016. ISSN 2398-3396. doi: 10.1049/iet-cps.2016.0019.

[3] George Loukas. 1 - a cyber-physical world. In George Loukas, editor, *Cyber-Physical Attacks*, pages 1 – 19. Butterworth-Heinemann, Boston, 2015. ISBN 978-0-12-801290-1. doi: https://doi.org/10.1016/B978-0-12-801290-1.00001-1. URL http://www.sciencedirect.com/science/article/pii/B9780128012901000011.

[4] Siddharth Sridhar, Adam Hahn, and Manimaran Govindarasu. Cyber–physical system security for the electric power grid. 100(1):210–224. ISSN 0018-9219, 1558-2256. doi: 10.1109/JPROC.2011.2165269. URL http://ieeexplore.ieee.org/document/6032699/.

[5] Thomas C Reed. *At the abyss: an insider's history of the Cold War*. Presidio Press, 2005.

[6] Terry L Hardy. *Software and System Safety*. AuthorHouse, 2012.

[7] Paulo Shakarian. Stuxnet: Cyberwar revolution in military affairs. page 11.

[8] James P. Farwell and Rafal Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2 2011. ISSN 0039-6338, 1468-2699. doi: 10.1080/00396338.2011.555586.

[9] Gaoqi Liang, Steven R. Weller, Junhua Zhao, Fengji Luo, and Zhao Yang Dong. The

2015 ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, 32(4):3317–3318, 7 2017. ISSN 0885-8950, 1558-0679. doi: 10.1109/TPWRS.2016.2631891.

[10] Hassan Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine*, 8(1):18–28, 1 2010. ISSN 1558-4216. doi: 10.1109/MPE.2009.934876. event: IEEE Power and Energy Magazine.

[11] J. Rost and R. L. Glass. *Disgruntled Employees and Sabotage*, pages 189–212. 2011.

[12] F. M. Cleveland. Cyber security issues for advanced metering infrastructure (ami). pages 1–5. 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 7 2008. doi: 10.1109/PES.2008.4596535. ISSN: 1932-5517.

[13] Patrick McDaniel and Stephen McLaughlin. Security and privacy challenges in the smart grid. *IEEE Security Privacy*, 7(3):75–77, 5 2009. ISSN 1558-4046. doi: 10.1109/MSP.2009.76. event: IEEE Security Privacy.

[14] Himanshu Khurana, Mark Hadley, Ning Lu, and Deborah A. Frincke. Smart-grid security issues. *IEEE Security Privacy*, 8(1):81–85, 1 2010. ISSN 1558-4046. doi: 10.1109/MSP.2010.49. event: IEEE Security Privacy.

[15] Le Xie, Yilin Mo, and Bruno Sinopoli. False data injection attacks in electricity markets. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 226–231. IEEE, 2010.

[16] S. N. Islam, Z. Baig, and S. Zeadally. Physical layer security for the smart grid: Vulnerabilities, threats, and countermeasures. *IEEE Transactions on Industrial Informatics*, 15(12):6522–6530, 2019.

[17] Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé. *Power Generation, Operation, and Control*. John Wiley Sons, 12 2013. ISBN 978-1-118-73391-2.

[18] None, None. National Electric Sector Cybersecurity Organization Resource

(NESCOR). Technical Report 1163840, June 2014. URL http://www.osti.gov/servlets/purl/1163840/.

[19] Yao Liu, Peng Ning, and Michael K Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, 14(1):1–33, 2011.

[20] Yanling Yuan, Zuyi Li, and Kui Ren. Modeling load redistribution attacks in power systems. *IEEE Transactions on Smart Grid*, 2(2):382–390, 6 2011. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2011.2123925.

[21] Hang Zhang, Bo Liu, and Hongyu Wu. Net load redistribution attacks on nodal voltage magnitude estimation in ac distribution networks. In *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 46–50. IEEE, 2020.

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

[23] Alexandru G Bardas, Sathya Chandran Sundaramurthy, Xinming Ou, and Scott A DeLoach. Mtd cbits: Moving target defense for cloud-based it systems. In *European Symposium on Research in Computer Security*, pages 167–186. Springer, 2017.

[24] Matthew Dunlop, Stephen Groat, William Urbanski, Randy Marchany, and Joseph Tront. Mt6d: A moving target ipv6 defense. In *2011-MILCOM 2011 Military Communications Conference*, pages 1321–1326. IEEE, 2011.

[25] Thomas E Carroll, Michael Crouse, Errin W Fulp, and Kenneth S Berenhaut. Analysis of network address shuffling as a moving target defense. In *2014 IEEE international conference on communications (ICC)*, pages 701–706. IEEE, 2014.

[26] David Evans, Anh Nguyen-Tuong, and John Knight. Effectiveness of moving target defenses. In *Moving target defense*, pages 29–48. Springer, 2011.

[27] Stephen W Boyd, Gaurav S Kc, Michael E Locasto, Angelos D Keromytis, and Vas-

silis Prevelakis. On the general applicability of instruction-set randomization. *IEEE Transactions on Dependable and Secure Computing*, 7(3):255–270, 2008.

[28] Deepak Divan and Harjeet Johal. Distributed facts-a new concept for realizing grid power flow control. In *2005 IEEE 36th Power Electronics Specialists Conference*, pages 8–14. IEEE, 2005.

[29] Smart Wires Inc. A mobile unit tours europe, smart wires in india and more, 2019. URL https://www.smartwires.com/portfolio-item/8245/.

[30] Gabriela Hug and Joseph Andrew Giampapa. Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid*, 3(3):1362–1370, 2012. doi: 10.1109/TSG.2012.2195338.

[31] Bo Liu and Hongyu Wu. Optimal D-FACTS placement in moving target defense against false data injection attacks. *IEEE Transactions on Smart Grid*, 11(5):4345–4357, 2020.

[32] Chensheng Liu, Jing Wu, Chengnian Long, and Deepa Kundur. Reactance perturbation for detecting and identifying fdi attacks in power system state estimation. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):763–776, 8 2018. ISSN 1941-0484. doi: 10.1109/JSTSP.2018.2846542. event: IEEE Journal of Selected Topics in Signal Processing.

[33] Fei Miao, Quanyan Zhu, Miroslav Pajic, and George J Pappas. Coding sensor outputs for injection attacks detection. In *53rd IEEE Conference on Decision and Control*, pages 5776–5781. IEEE, 2014.

[34] Jinping Hao, Robert J Piechocki, Dritan Kaleshi, Woon Hau Chin, and Zhong Fan. Sparse malicious false data injection attacks and defense mechanisms in smart grids. *IEEE Transactions on Industrial Informatics*, 11(5):1–12, 2015.

[35] Tung T Kim and H Vincent Poor. Strategic protection against data injection attacks on power grids. *IEEE Transactions on Smart Grid*, 2(2):326–333, 2011.

[36] PRICE CODE. Communication networks and systems in substations–part 5: Communication requirements for functions and device models. 2003.

[37] Abdulrahman Hadbah, Akhtar Kalam, and Aladin Zayegh. Powerful ieds, ethernet networks and their effects on iec 61850-based electric power utilities security. In *2017 Australasian Universities Power Engineering Conference (AUPEC)*, pages 1–5. IEEE, 2017.

[38] International Electrotechnical Commission et al. Power systems management and associated information exchange–data and communications security. part 1: Communication network and system security–introduction to security issues. *IEC Technical Specification*, pages 62351–1, 2007.

[39] Jue Tian, Rui Tan, Xiaohong Guan, and Ting Liu. Enhanced hidden moving target defense in smart grids. *IEEE transactions on smart grid*, 10(2):2208–2223, 2018.

[40] Bo Liu and Hongyu Wu. Optimal planning and operation of hidden moving target defense for maximal detection effectiveness. *IEEE Transactions on Smart Grid*, 12(5): 4447–4459, 2021.

[41] Bo Liu, Hongyu Wu, Anil Pahwa, Fei Ding, Erfan Ibrahim, and Ting Liu. Hidden moving target defense against false data injection in distribution network reconfiguration. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2018.

[42] Hang Zhang, Bo Liu, Xuebo Liu, Anil Pahwa, and Hongyu Wu. Voltage stability constrained moving target defense against net load redistribution attacks. *IEEE Transactions on Smart Grid*, 13(5):3748–3759, 2022.

[43] Bo Liu and Hongyu Wu. Systematic planning of moving target defence for maximising detection effectiveness against false data injection attacks in smart grid. *IET Cyber-Physical Systems: Theory & Applications*, 6(3):151–163, 2021.

[44] Yilin Mo, Sean Weerakkody, and Bruno Sinopoli. Physical authentication of control

systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems Magazine*, 35(1):93–109, 2 2015. ISSN 1941-000X. doi: 10.1109/MCS.2014.2364724. event: IEEE Control Systems Magazine.

[45] Sean Weerakkody, Yilin Mo, and Bruno Sinopoli. Detecting integrity attacks on control systems using robust physical watermarking. In *53rd IEEE Conference on Decision and Control*, pages 3757–3764. IEEE, 2014.

[46] Fei Miao, Quanyan Zhu, Miroslav Pajic, and George J Pappas. Coding schemes for securing cyber-physical systems against stealthy data injection attacks. *IEEE Transactions on Control of Network Systems*, 4(1):106–117, 2016.

[47] Zhong-Hua Pang, Lan-Zhi Fan, Jian Sun, Kun Liu, and Guo-Ping Liu. Detection of stealthy false data injection attacks against networked control systems via active data modification. *Information Sciences*, 546:192–205, 2021.

[48] Chensheng Liu, Ruilong Deng, Wangli He, Hao Liang, and Wenli Du. Optimal coding schemes for detecting false data injection attacks in power system state estimation. *IEEE Transactions on Smart Grid*, 13(1):738–749, 2021.

[49] Rodrigo D Trevizan and Matthew Reno. Detection of false data injection attacks in power system state estimation using sensor encoding. In *2022 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–6. IEEE, 2022.

[50] U.S. National electric grid security and resilience action plan. [Online]. Available: https://www.whitehouse.gov/sites/whitehouse.gov/fles/images/National_Electric_Grid_Action_Plan_06Dec2016.pdf. Accessed: 2022-09-1.

[51] Xuan Liu and Zuyi Li. False data attacks against AC state estimation with incomplete network information. *IEEE Transactions on smart grid*, 8(5):2239–2248, 2016.

[52] Bo Liu and Hongyu Wu. Optimal D-FACTS placement in moving target defense against false data injection attacks. *IEEE Transactions on Smart Grid*, 11(5):4345–4357, 2020. doi: 10.1109/TSG.2020.2977207.

[53] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal. Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Transactions on Power Systems*, 19(3):1387–1401, 2004. doi: 10.1109/TPWRS.2004.825981.

[54] Mingjian Cui and Jianhui Wang. Deeply hidden moving-target-defense for cybersecure unbalanced distribution systems considering voltage stability. *IEEE Transactions on Power Systems*, 36(3):1961–1972, 2021. doi: 10.1109/TPWRS.2020.3031256.

[55] Subhash Lakshminarayana, E. Veronica Belmega, and H. Vincent Poor. Moving-target defense for detecting coordinated cyber-physical attacks in power grids. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–7, 2019. doi: 10.1109/SmartGridComm.2019.8909767.

[56] Zhenyong Zhang, Ruilong Deng, David K. Y. Yau, Peng Cheng, and Jiming Chen. Analysis of moving target defense against false data injection attacks on power grid. *IEEE Transactions on Information Forensics and Security*, 15:2320–2335, 2020. doi: 10.1109/TIFS.2019.2928624.

[57] Martin Higgins, Keith Mayes, and Fei Teng. Enhanced cyber-physical security using attack-resistant cyber nodes and event-triggered moving target defence. *arXiv preprint arXiv:2010.14173*, 2020.

[58] Hang Zhang, Bo Liu, and Hongyu Wu. Smart grid cyber-physical attack and defense: A review. *IEEE Access*, 9:29641–29659, 2021.

[59] Salsabeel Shapsough, Fatma Qatan, Raafat Aburukba, Fadi Aloul, and AR Al Ali. Smart grid cyber security: Challenges and solutions. In *2015 International conference on smart grid and clean energy technologies (ICSGCE)*, pages 170–175. IEEE, 2015.

[60] A Procopiou and N Komninos. Current and future threats framework in smart grid

domain. In *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, pages 1852–1857. IEEE, 2015.

[61] Rajendra Kumar Pandey and Mohit Misra. Cyber security threats—smart grid infrastructure. In *2016 National Power Systems Conference (NPSC)*, pages 1–6. IEEE, 2016.

[62] Anibal Sanjab, Walid Saad, Ismail Guvenc, Arif Sarwat, and Saroj Biswas. Smart grid security: Threats, challenges, and solutions. *arXiv preprint arXiv:1606.06992*, 2016.

[63] Ilhami Colak, Seref Sagiroglu, Gianluca Fulli, Mehmet Yesilbudak, and Catalin-Felix Covrig. A survey on the critical issues in smart grid technologies. *Renewable and Sustainable Energy Reviews*, 54:396–405, 2016.

[64] Peter Eder-Neuhauser, Tanja Zseby, Joachim Fabini, and Gernot Vormayr. Cyber attack models for smart grid environments. *Sustainable Energy, Grids and Networks*, 12:10–29, 2017.

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

[66] Guneet Bedi, Ganesh Kumar Venayagamoorthy, Rajendra Singh, Richard R Brooks, and Kuang-Ching Wang. Review of internet of things (iot) in electric power and energy systems. *IEEE Internet of Things Journal*, 5(2):847–870, 2018.

[67] Ting Liu, Jue Tian, JZ Wang, H Wu, L Sun, Y Zhou, and X Guan. Integrated security threats and defense of cyber-physical systems. *Acta Automatica Sinica*, 45(1):5–24, 2019.

[68] Muhammed Zekeriya Gunduz and Resul Das. Cyber-security on smart grid: Threats and potential solutions. *Computer Networks*, 169:107094, 2020.

[69] V. Ajjarapu and C. Christy. The continuation power flow: a tool for steady state

voltage stability analysis. *IEEE Transactions on Power Systems*, 7(1):416–423, 1992. doi: 10.1109/59.141737.

[70] Hang Zhang, Noah Fulk, Bo Liu, Lawryn Edmonds, Xuebo Liu, and Hongyu Wu. Load margin constrained moving target defense against false data injection attacks. In *2022 IEEE Green Technologies Conference (GreenTech)*, pages 51–56. IEEE, 2022.

[71] Jun Yan, Yihai Zhu, Haibo He, and Yan Sun. Multi-contingency cascading analysis of smart grid based on self-organizing map. *IEEE Transactions on Information Forensics and Security*, 8(4):646–656, 4 2013. ISSN 1556-6021. doi: 10.1109/TIFS.2013.2249065. event: IEEE Transactions on Information Forensics and Security.

[72] Yousif Dafalla, Bo Liu, Dalton A. Hahn, Hongyu Wu, Reza Ahmadi, and Alexandru G. Bardas. Prosumer nanogrids: A cybersecurity assessment. *IEEE Access*, 8:131150–131164, 2020. ISSN 2169-3536. doi: 10.1109/ACCESS.2020.3009611. event: IEEE Access.

[73] Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3):49–51, 2011.

[74] Thomas M Chen. Stuxnet, the real start of cyber warfare?[editor's note]. *IEEE Network*, 24(6):2–3, 2010.

[75] Nicolas Falliere, Liam O Murchu, and Eric Chien. W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 5(6):29, 2011.

[76] Linyuan Zhang, Guoru Ding, Qihui Wu, Yulong Zou, Zhu Han, and Jinlong Wang. Byzantine attack and defense in cognitive radio networks: A survey. *IEEE Communications Surveys & Tutorials*, 17(3):1342–1363, 2015.

[77] A Geetha and N Sreenath. Byzantine attacks and its security measures in mobile adhoc networks. *Int'l Journal of Computing, Communications and Instrumentation Engineering (IJCCIE 2016)*, 3(1):42–47, 2016.

[78] Guoru Ding, Jinlong Wang, Qihui Wu, Linyuan Zhang, Yulong Zou, Yu-Dong Yao,

and Yingying Chen. Robust spectrum sensing with crowd sensors. *IEEE Transactions on Communications*, 62(9):3129–3143, 2014.

[79] Jiahu Qin, Menglin Li, Ling Shi, and Xinghuo Yu. Optimal denial-of-service attack scheduling with energy constraint over packet-dropping networks. *IEEE Transactions on Automatic Control*, 63(6):1648–1663, 2017.

[80] Heng Zhang and Wei Xing Zheng. Denial-of-service power dispatch against linear quadratic control via a fading channel. *IEEE Transactions on Automatic Control*, 63 (9):3032–3039, 2018.

[81] Mark Zeller. Common questions and answers addressing the aurora vulnerability. *Schweitzer Engineering Laboratories Report*, 2011.

[82] M. Zeller. Myth or reality — does the aurora vulnerability pose a risk to my generator? In *2011 64th Annual Conference for Protective Relay Engineers*, pages 130–136, 2011. doi: 10.1109/CPRE.2011.6035612.

[83] Ieee standard for salient-pole 50 hz and 60 hz synchronous generators and generator/motors for hydraulic turbine applications rated 5 mva and above. *IEEE Std C50.12-2005*, pages 1–45, 2006. doi: 10.1109/IEEESTD.2006.99082.

[84] Anurag Srivastava, Thomas Morris, Timothy Ernster, Ceeman Vellaithurai, Shengyi Pan, and Uttam Adhikari. Modeling cyber-physical vulnerability of the smart grid with incomplete information. *IEEE Transactions on Smart Grid*, 4(1):235–244, 3 2013. ISSN 1949-3061. doi: 10.1109/TSG.2012.2232318.

[85] Mohammadreza FM Arani, Amir Abiri Jahromi, Deepa Kundur, and Marthe Kassouf. Modeling and simulation of the aurora attack on microgrid point of common coupling. In *2019 7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, pages 1–6. IEEE, 2019.

[86] Ieee approved draft standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces - amendment 1 to

ieee std 1547-2018 to provide more flexibility for adoption of abnormal operating performance category iii. *IEEE P1547a/D1.4, January 2020*, pages 1–17, 3 2020.

[87] Rui Tan, Varun Badrinath Krishna, David KY Yau, and Zbigniew Kalbarczyk. Impact of integrity attacks on real-time pricing in smart grids. page 439–450, 2013.

[88] Y. Zhang, V. V. G. Krishnan, J. Pi, K. Kaur, A. Srivastava, A. Hahn, and S. Suresh. Cyber physical security analytics for transactive energy systems. *IEEE Transactions on Smart Grid*, 11(2):931–941, 3 2020. ISSN 1949-3061. doi: 10.1109/TSG.2019.2928168.

[89] Sabita Maharjan, Quanyan Zhu, Yan Zhang, Stein Gjessing, and Tamer Basar. Dependable demand response management in the smart grid: A stackelberg game approach. *IEEE Transactions on Smart Grid*, 4(1):120–132, 3 2013. ISSN 1949-3061. doi: 10.1109/TSG.2012.2223766. event: IEEE Transactions on Smart Grid.

[90] Rui Tan, Hoang Hai Nguyen, Eddy. Y. S. Foo, David K. Y. Yau, Zbigniew Kalbarczyk, Ravishankar K. Iyer, and Hoay Beng Gooi. Modeling and mitigating impact of false data injection attacks on automatic generation control. *IEEE Transactions on Information Forensics and Security*, 12(7):1609–1624, 7 2017. ISSN 1556-6021. doi: 10.1109/TIFS.2017.2676721. event: IEEE Transactions on Information Forensics and Security.

[91] Jairo Giraldo, Alvaro Cárdenas, and Nicanor Quijano. Integrity attacks on real-time pricing in smart grids: Impact and countermeasures. *IEEE Transactions on Smart Grid*, 8(5):2249–2257, 9 2017. ISSN 1949-3061. doi: 10.1109/TSG.2016.2521339.

[92] Siddharth Sridhar and Manimaran Govindarasu. Model-based attack detection and mitigation for automatic generation control. *IEEE Transactions on Smart Grid*, 5(2): 580–591, 3 2014. ISSN 1949-3061. doi: 10.1109/TSG.2014.2298195.

[93] Chunyu Chen, Kaifeng Zhang, Kun Yuan, Lingzhi Zhu, and Minhui Qian. Novel detection scheme design considering cyber attacks on load frequency control. *IEEE*

Transactions on Industrial Informatics, 14(5):1932–1941, 5 2018. ISSN 1941-0050. doi: 10.1109/TII.2017.2765313. event: IEEE Transactions on Industrial Informatics.

[94] Yao Liu, Ning Peng, and Michael K. Reiter. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, page 21–32, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605588940. doi: 10.1145/1653662.1653666. URL https://doi.org/10.1145/1653662.1653666.

[95] Barry M. Horowitz and Katherine M. Pierce. The integration of diversely redundant designs, dynamic system models, and state estimation technology to the cyber security of physical systems. *Systems Engineering*, 16(4):401–412, 2013. ISSN 1520-6858. doi: 10.1002/sys.21239. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sys.21239.

[96] Rakesh B Bobba, Katherine M Rogers, Qiyan Wang, Himanshu Khurana, Klara Nahrstedt, and Thomas J Overbye. Detecting false data injection attacks on dc state estimation. volume 2010, 2010.

[97] Zhisheng Wang, Ying Chen, Feng Liu, Yue Xia, and Xuemin Zhang. Power system security under false data injection attacks with exploitation and exploration based on reinforcement learning. *IEEE Access*, 6:48785–48796, 2018. ISSN 2169-3536. doi: 10.1109/ACCESS.2018.2856520. event: IEEE Access.

[98] Jingwen Liang, Lalitha Sankar, and Oliver Kosut. Vulnerability analysis and consequences of false data injection attack on power system state estimation. *IEEE Transactions on Power Systems*, 31(5):3864–3872, 2016.

[99] Gabriela Hug and Joseph Andrew Giampapa. Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid*, 3(3):1362–1370, 9 2012. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2012.2195338.

[100] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven R. Weller, and Zhao Yang Dong. A

review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 7 2017. ISSN 1949-3061. doi: 10.1109/TSG.2015.2495133.

[101] Kin Cheong Sou, Henrik Sandberg, and Karl Henrik Johansson. Electric power network security analysis via minimum cut relaxation. pages 4054–4059. 2011 50th IEEE Conference on Decision and Control and European Control Conference, 12 2011. doi: 10.1109/CDC.2011.6160456. ISSN: 0743-1546.

[102] Huaizhi Wang, Jiaqi Ruan, Guibin Wang, Bin Zhou, Yitao Liu, Xueqian Fu, and Jianchun Peng. Deep learning-based interval state estimation of ac smart grids against sparse cyber attacks. *IEEE Transactions on Industrial Informatics*, 14(11):4766–4778, 11 2018. ISSN 1941-0050. doi: 10.1109/TII.2018.2804669. event: IEEE Transactions on Industrial Informatics.

[103] J. Liang, L. Sankar, and O. Kosut. Vulnerability analysis and consequences of false data injection attack on power system state estimation. *IEEE Transactions on Power Systems*, 31(5):3864–3872, 2016.

[104] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. Vincent Poor. Distributed models for sparse attack construction and state vector estimation in the smart grid. In *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pages 306–311, 2012.

[105] Junbo Zhao, Lamine Mili, and Meng Wang. A generalized false data injection attacks against power system nonlinear state estimator and countermeasures. *IEEE Transactions on Power Systems*, 33(5):4868–4877, 2018.

[106] Jinsub Kim, Lang Tong, and Robert J. Thomas. Subspace methods for data attack on state estimation: A data driven approach. *IEEE Transactions on Signal Processing*, 63 (5):1102–1114, 3 2015. ISSN 1941-0476. doi: 10.1109/TSP.2014.2385670. event: IEEE Transactions on Signal Processing.

[107] Zong-Han Yu and Wen-Long Chin. Blind false data injection attack using pca approximation method in smart grid. *IEEE Transactions on Smart Grid*, 6(3):1219–1226, 5 2015. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2014.2382714.

[108] Principle component analysis: Springer series in statistics. In I. T. Jolliffe, editor, *Principal Component Analysis*, pages 1–9. Springer New York, New York, NY, 2002. ISBN 978-0-387-22440-4. doi: 10.1007/0-387-22440-8_1. URL https://doi.org/10.1007/0-387-22440-8_1.

[109] Md. Ashfaqur Rahman and Hamed Mohsenian-Rad. False data injection attacks with incomplete information against smart power grids. pages 3153–3158. 2012 IEEE Global Communications Conference (GLOBECOM), 12 2012. doi: 10.1109/GLOCOM.2012.6503599. ISSN: 1930-529X.

[110] Vassilis Kekatos, Georgios B. Giannakis, and Ross Baldick. Grid topology identification using electricity prices. pages 1–5. 2014 IEEE PES General Meeting | Conference Exposition, 7 2014. doi: 10.1109/PESGM.2014.6939474. ISSN: 1932-5517.

[111] Mohammad Esmalifalak, Huy Nguyen, Rong Zheng, and Zhu Han. Stealth false data injection using independent component analysis in smart grid. pages 244–248. 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), 10 2011. doi: 10.1109/SmartGridComm.2011.6102326.

[112] Mr Martin Higgins, Dr Fei Teng, and Professor Thomas Parisini. Stealthy mtd against unsupervised learning-based blind fdi attacks in power systems. *arXiv preprint arXiv:2004.07004*, 2020.

[113] R. Deng and H. Liang. False data injection attacks with limited susceptance information and new countermeasures in smart grid. *IEEE Transactions on Industrial Informatics*, 15(3):1619–1628, 2019.

[114] Ying Chen, Shaowei Huang, Feng Liu, Zhisheng Wang, and Xinwei Sun. Evaluation of reinforcement learning-based false data injection attack to automatic voltage control.

IEEE Transactions on Smart Grid, 10(2):2158–2169, 3 2019. ISSN 1949-3061. doi: 10.1109/TSG.2018.2790704. event: IEEE Transactions on Smart Grid.

[115] Ian Markwood, Yao Liu, Kevin Kwiat, and Charles Kamhoua. Electric grid power flow model camouflage against topology leaking attacks. pages 1–9. IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, 5 2017. doi: 10.1109/INFOCOM.2017.8057060.

[116] Xuan Liu and Zuyi Li. Local load redistribution attacks in power systems with incomplete network information. *IEEE Transactions on Smart Grid*, 5(4):1665–1676, 2014.

[117] Liang Che, Xuan Liu, Zuyi Li, and Yunfeng Wen. False data injection attacks induced sequential outages in power systems. *IEEE Transactions on Power Systems*, 34(2): 1513–1523, 3 2019. ISSN 1558-0679. doi: 10.1109/TPWRS.2018.2871345. event: IEEE Transactions on Power Systems.

[118] D. Choeum and D. Choi. Vulnerability assessment of conservation voltage reduction to load redistribution attack in unbalanced active distribution networks. *IEEE Transactions on Industrial Informatics*, pages 1–1, 2020. doi: 10.1109/TII.2020.2980590.

[119] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang. Power system reliability evaluation considering load redistribution attacks. *IEEE Transactions on Smart Grid*, 8(2):889–901, 2017. doi: 10.1109/TSG.2016.2569589.

[120] J. Fu, L. Wang, B. Hu, K. Xie, H. Chao, and P. Zhou. A sequential coordinated attack model for cyber-physical system considering cascading failure and load redistribution. In *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pages 1–6, 2018. doi: 10.1109/EI2.2018.8582135.

[121] Jinsub Kim and Lang Tong. On topology attack of a smart grid: Undetectable attacks and countermeasures. *IEEE Journal on Selected Areas in Communications*, 31(7): 1294–1305, 7 2013. ISSN 0733-8716. doi: 10.1109/JSAC.2013.130712.

[122] Desmond Schmidt, Kenneth Radke, Seyit Camtepe, Ernest Foo, and Michał Ren. A survey and analysis of the gnss spoofing threat and countermeasures. *ACM Computing Surveys*, 48(4):1–31, 5 2016. ISSN 0360-0300, 1557-7341. doi: 10.1145/2897166.

[123] Yi Cui, Feifei Bai, Yilu Liu, Peter L. Fuhr, and Marissa E. Morales-Rodríguez. Spatio-temporal characterization of synchrophasor data against spoofing attacks in smart grids. *IEEE Transactions on Smart Grid*, 10(5):5807–5818, 9 2019. ISSN 1949-3061. doi: 10.1109/TSG.2019.2891852. event: IEEE Transactions on Smart Grid.

[124] Paresh Risbud, Nikolaos Gatsis, and Ahmad Taha. Vulnerability analysis of smart grids to gps spoofing. *IEEE Transactions on Smart Grid*, 10(4):3535–3548, 7 2019. ISSN 1949-3061. doi: 10.1109/TSG.2018.2830118. event: IEEE Transactions on Smart Grid.

[125] Liang Che, Xuan Liu, and Zuyi Li. Fast screening of high-risk lines under false data injection attacks. *IEEE Transactions on Smart Grid*, 10(4):4003–4014, 2018.

[126] Xuan Liu and Zuyi Li. Local topology attacks in smart grids. *IEEE Transactions on Smart Grid*, 8(6):2617–2626, 11 2017. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2016.2532347.

[127] Saleh Soltan, Mihalis Yannakakis, and Gil Zussman. React to cyber attacks on power grids. *IEEE Transactions on Network Science and Engineering*, 6(3):459–473, 7 2019. ISSN 2334-329X. doi: 10.1109/TNSE.2018.2837894.

[128] Saleh Soltan and Gil Zussman. Expose the line failures following a cyber-physical attack on the power grid. *IEEE Transactions on Control of Network Systems*, 6(1): 451–461, 3 2019. ISSN 2372-2533. doi: 10.1109/TCNS.2018.2844244.

[129] Saleh Soltan, Prateek Mittal, and H. Vincent Poor. Line failure detection after a cyber-physical attack on the grid using bayesian regression. *IEEE Transactions on Power Systems*, 34(5):3758–3768, 9 2019. ISSN 1558-0679. doi: 10.1109/TPWRS.2019.2910396.

[130] Hwei-Ming Chung, Wen-Tai Li, Chau Yuen, Wei-Ho Chung, Yan Zhang, and Chao-Kai Wen. Local cyber-physical attack for masking line outage and topology attack in smart grid. *IEEE Transactions on Smart Grid*, 10(4):4577–4588, 7 2019. ISSN 1949-3061. doi: 10.1109/TSG.2018.2865316.

[131] Ruilong Deng, Peng Zhuang, and Hao Liang. Ccpa: Coordinated cyber-physical attacks and countermeasures in smart grid. *IEEE Transactions on Smart Grid*, 8(5): 2420–2430, 9 2017. ISSN 1949-3061. doi: 10.1109/TSG.2017.2702125.

[132] Stamatis Karnouskos. Stuxnet worm impact on industrial cyber-physical system security. In *IECON 2011-37th Annual Conference of the IEEE Industrial Electronics Society*, pages 4490–4494. IEEE, 2011.

[133] Jue Tian, Rui Tan, Xiaohong Guan, Zhanbo Xu, and Ting Liu. Moving target defense approach to detecting stuxnet-like attacks. *IEEE Transactions on Smart Grid*, pages 1–1, 2019. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2019.2921245.

[134] Defense Use Case. Analysis of the cyber attack on the ukrainian power grid. *Electricity Information Sharing and Analysis Center (E-ISAC)*, 388, 2016.

[135] Zhiyi Li, Mohammad Shahidehpour, Ahmed Alabdulwahab, and Abdullah Abusorrah. Bilevel model for analyzing coordinated cyber-physical attacks on power systems. *IEEE Transactions on Smart Grid*, 7(5):2260–2272, 9 2016. ISSN 1949-3061. doi: 10.1109/TSG.2015.2456107. event: IEEE Transactions on Smart Grid.

[136] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah. Analyzing locally coordinated cyber-physical attacks for undetectable line outages. *IEEE Transactions on Smart Grid*, 9(1):35–47, 2018. doi: 10.1109/TSG.2016.2542925.

[137] Yilin Mo and Bruno Sinopoli. Secure control against replay attacks. In *2009 47th annual Allerton conference on communication, control, and computing (Allerton)*, pages 911–918. IEEE, 2009.

[138] Dan Simon. Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8):1303–1318, 2010.

[139] Raman K Mehra and J Peschon. An innovations approach to fault detection and diagnosis in dynamic systems. *Automatica*, 7(5):637–640, 1971.

[140] Yacine Chakhchoukh, Vijay Vittal, and Gerald T Heydt. Pmu based state estimation by integrating correlation. *IEEE Transactions on Power Systems*, 29(2):617–626, 2013.

[141] Annarita Giani, Eilyan Bitar, Manuel Garcia, Miles McQueen, Pramod Khargonekar, and Kameshwar Poolla. Smart grid data integrity attacks: characterizations and countermeasures. pages 232–237. 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), 10 2011. doi: 10.1109/SmartGridComm.2011.6102324.

[142] Junjian Qi, Kai Sun, and Wei Kang. Optimal pmu placement for power system dynamic state estimation by using empirical observability gramian. *IEEE Transactions on Power Systems*, 30(4):2041–2054, 7 2015. ISSN 1558-0679. doi: 10.1109/TPWRS.2014.2356797. event: IEEE Transactions on Power Systems.

[143] Anamitra Pal, Anil Kumar S. Vullikanti, and S. S. Ravi. A pmu placement scheme considering realistic costs and modern trends in relaying. *IEEE Transactions on Power Systems*, 32(1):552–561, 1 2017. ISSN 1558-0679. doi: 10.1109/TPWRS.2016.2551320. event: IEEE Transactions on Power Systems.

[144] Morteza Sarailoo and N. Eva Wu. Cost-effective upgrade of pmu networks for fault-tolerant sensing. *IEEE Transactions on Power Systems*, 33(3):3052–3063, 5 2018. ISSN 1558-0679. doi: 10.1109/TPWRS.2017.2756030. event: IEEE Transactions on Power Systems.

[145] Qingyu Yang, Dou An, and Wei Yu. On time desynchronization attack against ieee 1588 protocol in power grid systems. pages 1–5. 2013 IEEE Energytech, 5 2013. doi: 10.1109/EnergyTech.2013.6645332.

[146] Zhenghao Zhang, Shuping Gong, Aleksandar D. Dimitrovski, and Husheng Li. Time synchronization attack in smart grid: Impact and analysis. *IEEE Transactions on Smart Grid*, 4(1):87–98, 3 2013. ISSN 1949-3061. doi: 10.1109/TSG.2012.2227342. event: IEEE Transactions on Smart Grid.

[147] Xichen Jiang, Jiangmeng Zhang, Brian J. Harding, Jonathan J. Makela, and Alejandro D. Domı´nguez-Garcı´a. Spoofing gps receiver clock offset of phasor measurement units. *IEEE Transactions on Power Systems*, 28(3):3253–3262, 8 2013. ISSN 1558-0679. doi: 10.1109/TPWRS.2013.2240706. event: IEEE Transactions on Power Systems.

[148] Yawen Fan, Zhenghao Zhang, Matthew Trinkle, Aleksandar D. Dimitrovski, Ju Bin Song, and Husheng Li. A cross-layer defense mechanism against gps spoofing attacks on pmus in smart grids. *IEEE Transactions on Smart Grid*, 6(6):2659–2668, 11 2015. ISSN 1949-3061. doi: 10.1109/TSG.2014.2346088. event: IEEE Transactions on Smart Grid.

[149] Yi Huang, Husheng Li, Kristy A. Campbell, and Zhu Han. Defending false data injection attack on smart grid network using adaptive cusum test. pages 1–6. 2011 45th Annual Conference on Information Sciences and Systems, 3 2011. doi: 10.1109/CISS.2011.5766111.

[150] Lanchao Liu, Mohammad Esmalifalak, Qifeng Ding, Valentine A. Emesih, and Zhu Han. Detecting false data injection attacks on power grid by sparse optimization. *IEEE Transactions on Smart Grid*, 5(2):612–621, 3 2014. ISSN 1949-3061. doi: 10.1109/TSG.2013.2284438. event: IEEE Transactions on Smart Grid.

[151] Gu Chaojun, Panida Jirutitijaroen, and Mehul Motani. Detecting false data injection attacks in ac state estimation. *IEEE Transactions on Smart Grid*, 6(5):2476–2483, 9 2015. ISSN 1949-3061. doi: 10.1109/TSG.2015.2388545. event: IEEE Transactions on Smart Grid.

[152] Junbo Zhao, Gexiang Zhang, Massimo La Scala, Zhao Yang Dong, Chen Chen, and

Jianhui Wang. Short-term state forecasting-aided method for detection of smart grid general false data injection attacks. *IEEE Transactions on Smart Grid*, 8(4):1580–1590, July 2017. ISSN 1949-3061. doi: 10.1109/TSG.2015.2492827. event: IEEE Transactions on Smart Grid.

[153] Aditya Ashok, Manimaran Govindarasu, and Venkataramana Ajjarapu. Online detection of stealthy false data injection attacks in power system state estimation. *IEEE Transactions on Smart Grid*, 9(3):1636–1646, May 2018. ISSN 1949-3061. doi: 10.1109/TSG.2016.2596298. event: IEEE Transactions on Smart Grid.

[154] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor. Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1773–1786, 2016. doi: 10.1109/TNNLS.2015.2404803.

[155] J. Yan, B. Tang, and H. He. Detection of false data attacks in smart grid with supervised learning. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1395–1402, 2016. doi: 10.1109/IJCNN.2016.7727361.

[156] J. Sakhnini, H. Karimipour, and A. Dehghantanha. Smart grid cyber attacks detection using supervised learning and heuristic feature selection. In *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*, pages 108–112, 2019. doi: 10.1109/SEGE.2019.8859946.

[157] X. Niu, J. Li, J. Sun, and K. Tomsovic. Dynamic detection of false data injection attack in smart grid using deep learning. In *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6, 2019. doi: 10.1109/ISGT.2019.8791598.

[158] Y. Li, Y. Wang, and S. Hu. Online generative adversary network based measurement recovery in false data injection attacks: A cyber-physical approach. *IEEE Transactions on Industrial Informatics*, 16(3):2031–2043, 2020. doi: 10.1109/TII.2019.2921106.

[159] Guoru Ding, Qihui Wu, Yu-Dong Yao, Jinlong Wang, and Yingying Chen. Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions. *IEEE Signal Processing Magazine*, 30(4):126–136, 2013.

[160] Tianpei Xie, Nasser M Nasrabadi, and Alfred O Hero. Learning to classify with possible sensor failures. *IEEE Transactions on Signal Processing*, 65(4):836–849, 2016.

[161] Zhijin Qin, Yue Gao, and Mark D Plumbley. Malicious user detection based on low-rank matrix completion in wideband spectrum sensing. *IEEE Transactions on Signal Processing*, 66(1):5–17, 2017.

[162] Ehab Al-Shaer, Qi Duan, and Jafar Haadi Jafarian. Random host mutation for moving target defense. pages 310–327, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. ISBN 978-3-642-36883-7.

[163] Kate L. Morrow, Erich Heine, Katherine M. Rogers, Rakesh B. Bobba, and Thomas J. Overbye. Topology perturbation for detecting malicious data injection. pages 2104–2113. 2012 45th Hawaii International Conference on System Sciences, January 2012. doi: 10.1109/HICSS.2012.594. ISSN: 1530-1605.

[164] Katherine R. Davis, Kate L. Morrow, Rakesh Bobba, and Erich Heine. Power flow cyber attacks and perturbation-based defense. pages 342–347. 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm), November 2012. doi: 10.1109/SmartGridComm.2012.6486007.

[165] Mohammad Ashiqur Rahman, Ehab Al-Shaer, and Rakesh B. Bobba. Moving target defense for hardening the security of the power system state estimation. pages 59–68, Scottsdale, Arizona, USA, 2014. the First ACM Workshop, ACM Press. ISBN 978-1-4503-3150-0. doi: 10.1145/2663474.2663482. URL http://dl.acm.org/citation.cfm?doid=2663474.2663482. [Online; accessed 2020-05-16].

[166] Zhenyong Zhang, Ruilong Deng, David KY Yau, Peng Cheng, and Jiming Chen. Anal-

ysis of moving target defense against false data injection attacks on power grid. *IEEE Transactions on Information Forensics and Security*, 15:2320–2335, 2019.

[167] Beibei Li, Gaoxi Xiao, Rongxing Lu, Ruilong Deng, and Haiyong Bao. On feasibility and limitations of detecting false data injection attacks on power grid state estimation using d-facts devices. *IEEE Transactions on Industrial Informatics*, 16(2):854–864, 2 2020. ISSN 1941-0050. doi: 10.1109/TII.2019.2922215.

[168] Subhash Lakshminarayana and David KY Yau. Cost-benefit analysis of moving-target defense in power grids. *IEEE Transactions on Power Systems*, 2020.

[169] B. Liu, L. Edmonds, H. Zhang, and H. Wu. An interior-point solver for optimal power flow problem considering distributed facts devices. In *2020 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–5, 2020. doi: 10.1109/KPEC47870.2020.9167620.

[170] Jue Tian, Rui Tan, Xiaohong Guan, and Ting Liu. Enhanced hidden moving target defense in smart grids. *IEEE Transactions on Smart Grid*, 10(2):2208–2223, 3 2019. ISSN 1949-3053, 1949-3061. doi: 10.1109/TSG.2018.2791512.

[171] Zhenyong Zhang, Ruilong Deng, David KY Yau, Peng Cheng, and Jiming Chen. On hiddenness of moving target defense against false data injection attacks on power grid. *ACM Transactions on Cyber-Physical Systems*, 4(3):1–29, 2020.

[172] Fei Miao, Miroslav Pajic, and George J Pappas. Stochastic game approach for replay attack detection. In *52nd IEEE conference on decision and control*, pages 1854–1859. IEEE, 2013.

[173] Bharadwaj Satchidanandan and Panganamala R Kumar. Dynamic watermarking: Active defense of networked cyber–physical systems. *Proceedings of the IEEE*, 105(2): 219–240, 2016.

[174] Aidin Ferdowsi and Walid Saad. Deep learning-based dynamic watermarking for secure signal authentication in the internet of things. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2018.

[175] Ying Zhang, Jianhui Wang, and Jianzhe Liu. Attack identification and correction for pmu gps spoofing in unbalanced distribution systems. *IEEE Transactions on Smart Grid*, 11(1):762–773, 1 2020. ISSN 1949-3061. doi: 10.1109/TSG.2019.2937554. event: IEEE Transactions on Smart Grid.

[176] Jean Bélanger, P Venne, and Jean-Nicolas Paquin. The what, where and why of real-time simulation. *Planet Rt*, 1(1):25–29, 2010.

[177] CM Davis, JE Tate, H Okhravi, C Grier, Thomas J Overbye, and D Nicol. Scada cyber security testbed development. In *2006 38th North American Power Symposium*, pages 483–488. IEEE, 2006.

[178] Malaz Mallouhi, Youssif Al-Nashif, Don Cox, Tejaswini Chadaga, and Salim Hariri. A testbed for analyzing security of scada control systems (tasscs). In *ISGT 2011*, pages 1–7. IEEE, 2011.

[179] Thomas Strasser, Matthias Stifter, Filip Andrén, and Peter Palensky. Co-simulation training platform for smart grids. *IEEE Transactions on Power Systems*, 29(4):1989–1997, 2014.

[180] Bo Chen, Karen L Butler-Purry, Ana Goulart, and Deepa Kundur. Implementing a real-time cyber-physical system test bed in rtds and opnet. In *2014 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2014.

[181] Mir Hadi Athari and Zhifang Wang. Impacts of wind power uncertainty on grid vulnerability to cascading overload failures. *IEEE Transactions on Sustainable Energy*, 9 (1):128–137, 2017.

[182] A Muir and J Lopatto. Final report on the august 14, 2003 blackout in the united states and canada: causes and recommendations. *US–Canada Power System Outage Task Force, Canada*, 2004.

[183] Federal Energy Regulatory Commission et al. Arizona-southern california outages on september 8, 2011: Causes and recommendations. *FERC and NERC Staff, Apr*, 2012.

[184] Lei Wang, Zhaoyang Qu, Yang Li, Kewei Hu, Jian Sun, Kai Xue, and Mingshi Cui. Method for extracting patterns of coordinated network attacks on electric power cps based on temporal–topological correlation. *IEEE Access*, 8:57260–57272, 2020.

[185] Fabio Pasqualetti, Florian Dörfler, and Francesco Bullo. A divide-and-conquer approach to distributed attack identification. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5801–5807. IEEE, 2015.

[186] Francesco Fusco. General bad data identification and estimation in the presence of critical measurement sets. In *2014 IEEE PES General Meeting— Conference & Exposition*, pages 1–5. IEEE, 2014.

[187] Anggoro Primadianto and Chan-Nan Lu. A review on distribution system state estimation. *IEEE Transactions on Power Systems*, 32(5):3875–3883, 2016.

[188] Oliver Kosut, Liyan Jia, Robert J Thomas, and Lang Tong. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In *2010 first IEEE international conference on smart grid communications*, pages 220–225. IEEE, 2010.

[189] Yanling Yuan, Zuyi Li, and Kui Ren. Quantitative analysis of load redistribution attacks in power systems. *IEEE Transactions on Parallel and Distributed Systems*, 23 (9):1731–1738, 2012.

[190] Bo-Siang Fang, Chia-Chu Lai, Ying-Wei Lu, Kuan-Ta Chen, Mike Tasi, and Don-Son Jiang. A methodology to correct in-fixture measurement of impedance by a machine learning model. In *2019 IEEE 69th Electronic Components and Technology Conference (ECTC)*, pages 1704–1709. IEEE, 2019.

[191] Tadatoshi Sekine. An estimation method for the capacitance matrix of bundle of wires based on machine learning. In *2018 International Symposium on Electromagnetic Compatibility (EMC EUROPE)*, pages 1004–1007. IEEE, 2018.

[192] Snehal V Unde and Sanjay S Dambhare. Double circuit transmission line parameter

estimation using pmu. In *2016 IEEE 6th International Conference on Power Systems (ICPS)*, pages 1–4. IEEE, 2016.

[193] Dong Liang, Huashan Guo, and Tao Zheng. Real-time impedance estimation for power line communication. *IEEE Access*, 7:88107–88115, 2019.

[194] Lei Wang, Qun Zhou, and Shuangshuang Jin. Physics-guided deep learning for power system state estimation. *Journal of Modern Power Systems and Clean Energy*, 8(4): 607–615, 2020.

[195] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural networks*, 4(2):251–257, 1991.

[196] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, 2011. doi: 10.1109/TPWRS.2010.2051168.

[197] Beibei Li, Gaoxi Xiao, Rongxing Lu, Ruilong Deng, and Haiyong Bao. On feasibility and limitations of detecting false data injection attacks on power grid state estimation using d-facts devices. *IEEE Transactions on Industrial Informatics*, 16(2):854–864, 2019.

[198] Lei Wang and Hsiao-Dong Chiang. Toward online line switching for increasing load margins to static stability limit. *IEEE Transactions on Power Systems*, 31(3):1744–1751, 2015.

[199] B. Cui and X. A. Sun. A new voltage stability-constrained optimal power-flow model: Sufficient condition, SOCP representation, and relaxation. *IEEE Transactions on Power Systems*, 33(5):5092–5102, 2018. doi: 10.1109/TPWRS.2018.2801286.

[200] Chong Wang, Bai Cui, Zhaoyu Wang, and Chenghong Gu. SDP-based optimal power flow with steady-state voltage stability constraints. *IEEE Transactions on Smart Grid*, 10(4):4637–4647, 2018.

[201] Haoyu Yuan, Reetam Sen Biswas, Jin Tan, and Yingchen Zhang. Developing a reduced 240-bus wecc dynamic model for frequency response study of high renewable integration. In *2020 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–5. IEEE, 2020.

[202] GH Golub and CF Van Loan. Matrix computations, 476 pp. *Johns Hopkins University*, 1983.

[203] P-A Lof, G Andersson, and DJ Hill. Voltage stability indices for stressed power systems. *IEEE transactions on power systems*, 8(1):326–335, 1993.

[204] Y Wang, LCP Da Silva, W Xu, and Y Zhang. Analysis of ill-conditioned power-flow problems using voltage stability methodology. *IEE Proceedings-Generation, Transmission and Distribution*, 148(5):384–390, 2001.

[205] Jue Tian, Rui Tan, Xiaohong Guan, and Ting Liu. Hidden moving target defense in smart grids. In *Proceedings of the 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids*, pages 21–26, 2017.

[206] K. M. Rogers and T. J. Overbye. Some applications of distributed flexible AC transmission system (D-FACTS) devices in power systems. In *2008 40th North American Power Symposium*, pages 1–8, 2008. doi: 10.1109/NAPS.2008.5307314.

[207] Z. Wang, B. Cui, and J. Wang. A necessary condition for power flow insolvability in power distribution systems with distributed generators. *IEEE Transactions on Power Systems*, 32(2):1440–1450, 2017. doi: 10.1109/TPWRS.2016.2588341.

[208] Roberto Faranda, Antonio Pievatolo, and Enrico Tironi. Load shedding: A new proposal. *IEEE Transactions on Power Systems*, 22(4):2086–2093, 2007. doi: 10.1109/TPWRS.2007.907390.

[209] Charles Mozina. Undervoltage load shedding. In *2007 Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources*, pages 39–54, 2007. doi: 10.1109/PSAMP.2007.4740897.

[210] Bo Liu, Qihui Yang, Hang Zhang, and Hongyu Wu. An interior-point solver for AC optimal power flow considering variable impedance-based facts devices. *IEEE Access*, 9:154460–154470, 2021.

[211] Bo Liu and Hongyu Wu. Optimal planning and operation of hidden moving target defense for maximal detection effectiveness. *IEEE Transactions on Smart Grid*, pages 1–1, 2021. doi: 10.1109/TSG.2021.3076824.

[212] Subham Sahoo, Tomislav Dragičević, and Frede Blaabjerg. Cyber security in control of grid-tied power electronic converters—challenges and vulnerabilities. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(5):5326–5340, 2019.

[213] Mehmetcan Gursoy and Behrooz Mirafzal. On self-security of grid-interactive smart inverters. In *2021 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–6. IEEE, 2021.

[214] Tareq Hossen, Dushyant Sharma, and Behrooz Mirafzal. Smart inverter twin model for anomaly detection. In *2021 IEEE 22nd Workshop on Control and Modelling of Power Electronics (COMPEL)*, pages 1–6. IEEE, 2021.

[215] Jingwen Liang, Oliver Kosut, and Lalitha Sankar. Cyber attacks on ac state estimation: Unobservability and physical consequences. In *2014 IEEE PES General Meeting—Conference & Exposition*, pages 1–5. IEEE, 2014.

[216] Jingwen Liang, Lalitha Sankar, and Oliver Kosut. Vulnerability analysis and consequences of false data injection attack on power system state estimation. *IEEE Transactions on Power Systems*, 31(5):3864–3872, 2015.

[217] Alcir Monticelli. Electric power system state estimation. *Proceedings of the IEEE*, 88 (2):262–282, 2000.

[218] Ravindra Singh, Bikash C Pal, and Richard B Vinter. Measurement placement in distribution system state estimation. *IEEE Transactions on Power Systems*, 24(2): 668–675, 2009.

[219] Siddharth Bhela, Vassilis Kekatos, and Sriharsha Veeramachaneni. Enhancing observability in distribution grids using smart meter data. *IEEE Transactions on Smart Grid*, 9(6):5953–5961, 2017.

[220] Efthymios Manitsas, Ravindra Singh, Bikash C Pal, and Goran Strbac. Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling. *IEEE Transactions on power systems*, 27(4):1888–1896, 2012.

[221] Jianzhong Wu, Yan He, and Nick Jenkins. A robust state estimator for medium voltage distribution networks. *IEEE Transactions on Power Systems*, 28(2):1008–1016, 2012.

[222] Priya L Donti, Yajing Liu, Andreas J Schmitt, Andrey Bernstein, Rui Yang, and Yingchen Zhang. Matrix completion for low-observability voltage estimation. *IEEE Transactions on Smart Grid*, 11(3):2520–2530, 2019.

[223] Ruilong Deng, Peng Zhuang, and Hao Liang. False data injection attacks against state estimation in power distribution systems. *IEEE Transactions on Smart Grid*, 10(3): 2871–2881, 2018.

[224] Junjian Qi, Adam Hahn, Xiaonan Lu, Jianhui Wang, and Chen-Ching Liu. Cybersecurity for distributed energy resources and smart inverters. *IET Cyber-Physical Systems: Theory & Applications*, 1(1):28–39, 2016.

[225] Shu Tezuka. Linear congruential generators. In *Uniform Random Numbers*, pages 57–82. Springer, 1995.

[226] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[227] Hourly Load Data Archives. [Online]. Available: https://www.ercot.com/gridinfo/load/load_hist. Accessed: 2022-06-16.

[228] eGauge360. [Online]. Available: http://egauge33748.egaug.es/5AA89/. Accessed: 2022-06-16.

Appendix A

Acronyms

Acronyms

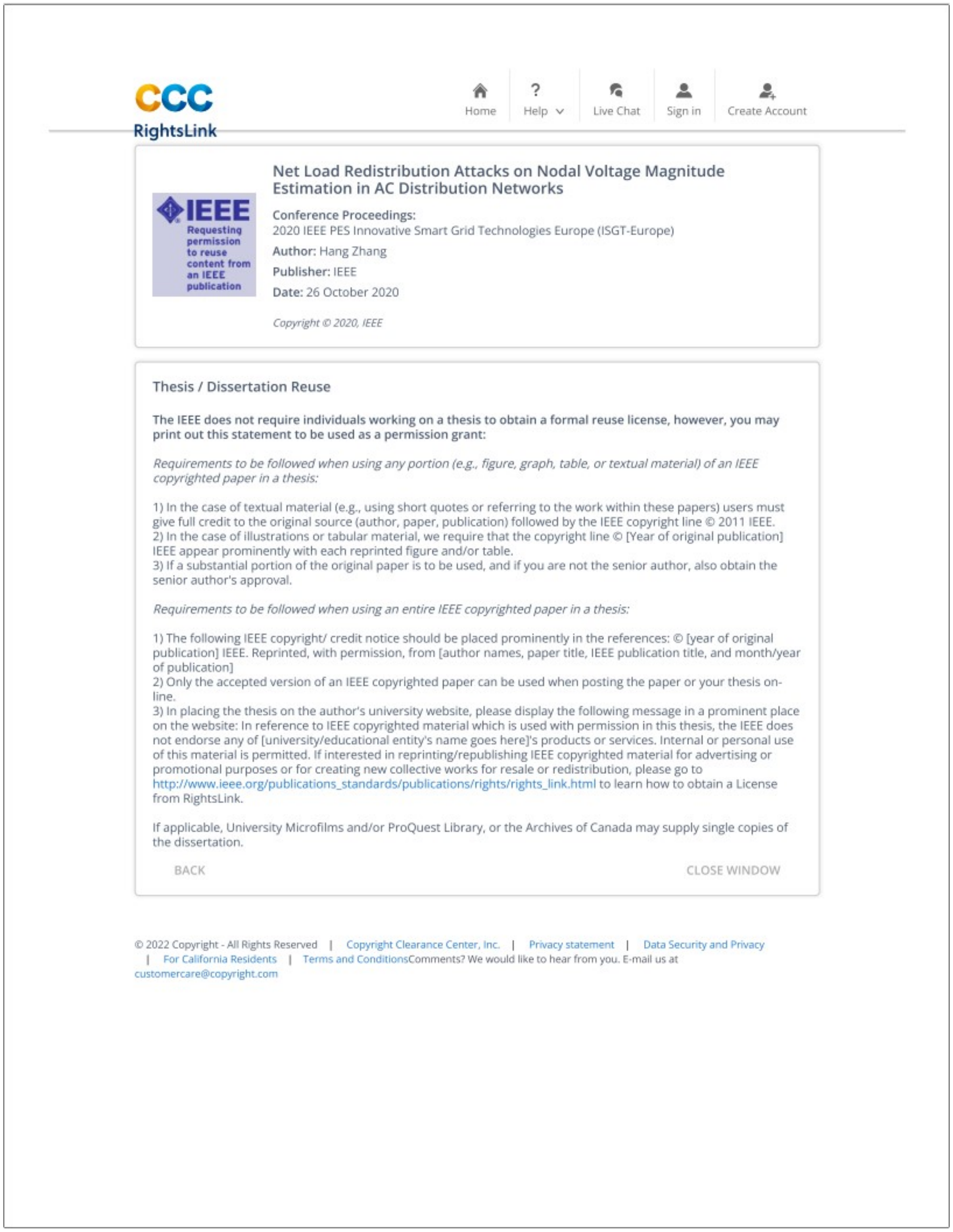
CPS	cyber-physical system
IoT	internet of things
SCADA	supervisory control and data acquisition
DoS	denial of service
ICT	information and communication technologies
CPSG	cyber-physical smart grid
NLRA	net load redistribution attack
MTD	moving target defense
EMS	energy management system
PMU	phasor measurement unit
SE	state estimation
BDD	bad data detection

NESCOR	national electric sector cybersecurity organization resource
WAMPAC	wide-area monitoring, protection, and control
PDC	phasor data concentrator
IT	information technologies
OT	operational technologies
OPF	optimal power flow
LR	load redistribution
DER	distributed energy resource
WLS	weighted least square
D-FACTS	distributed flexible AC transmission system
SVC	static var compensator
TCSC	thyristor controlled series capacitor
SSSC	static synchronous series compensator
CPF	contibuation power flow
PFEDL	power flow enhanced deep learning
AGC	automatic generation control
RTU	remote terminal unit
DSO	distributed system operator
RF	redundant factor
LSTM	long short-term memory
DNN	deep neural network
MAE	mean absolute error

ASP	attack stealthy probability
RMTD	random moving target defense
OMTD	optimized moving target defense
HMTD	hidden moving target defense
SNB	saddle node bifurcation
WECC	western electricity coordinating council
PII	power injection to impedance
PFI	power flow to impedance
SI	state to impedance
TII	*t*-index to impedance
ROC	receiver operating characteristic
TPR	true positive rate
FPR	false positive rate
AUC	area under curve

Appendix B

Reuse permissions from publishers


CCC
RightsLink
Home
Help
Live Chat
Sign in
Create Account
IEEE
Requesting permission to reuse content from an IEEE publication

Net Load Redistribution Attacks on Nodal Voltage Magnitude Estimation in AC Distribution Networks

Conference Proceedings:
2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)

Author: Hang Zhang

Publisher: IEEE

Date: 26 October 2020

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

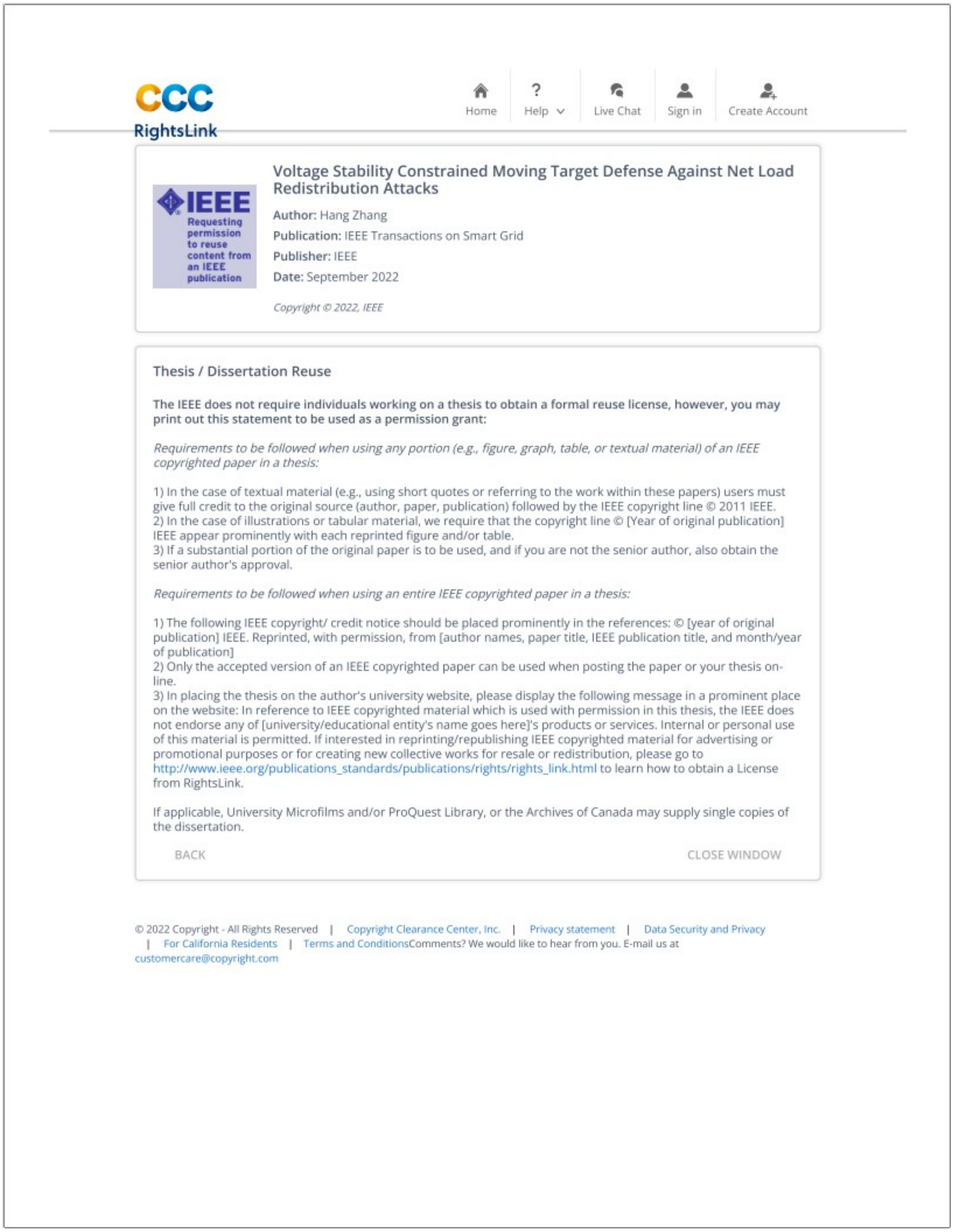
Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.


CCC
RightsLink
Home
Help
Live Chat
Sign in
Create Account
IEEE
Requesting permission to reuse content from an IEEE publication

Voltage Stability Constrained Moving Target Defense Against Net Load Redistribution Attacks

Author: Hang Zhang

Publication: IEEE Transactions on Smart Grid

Publisher: IEEE

Date: September 2022

Copyright © 2022, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

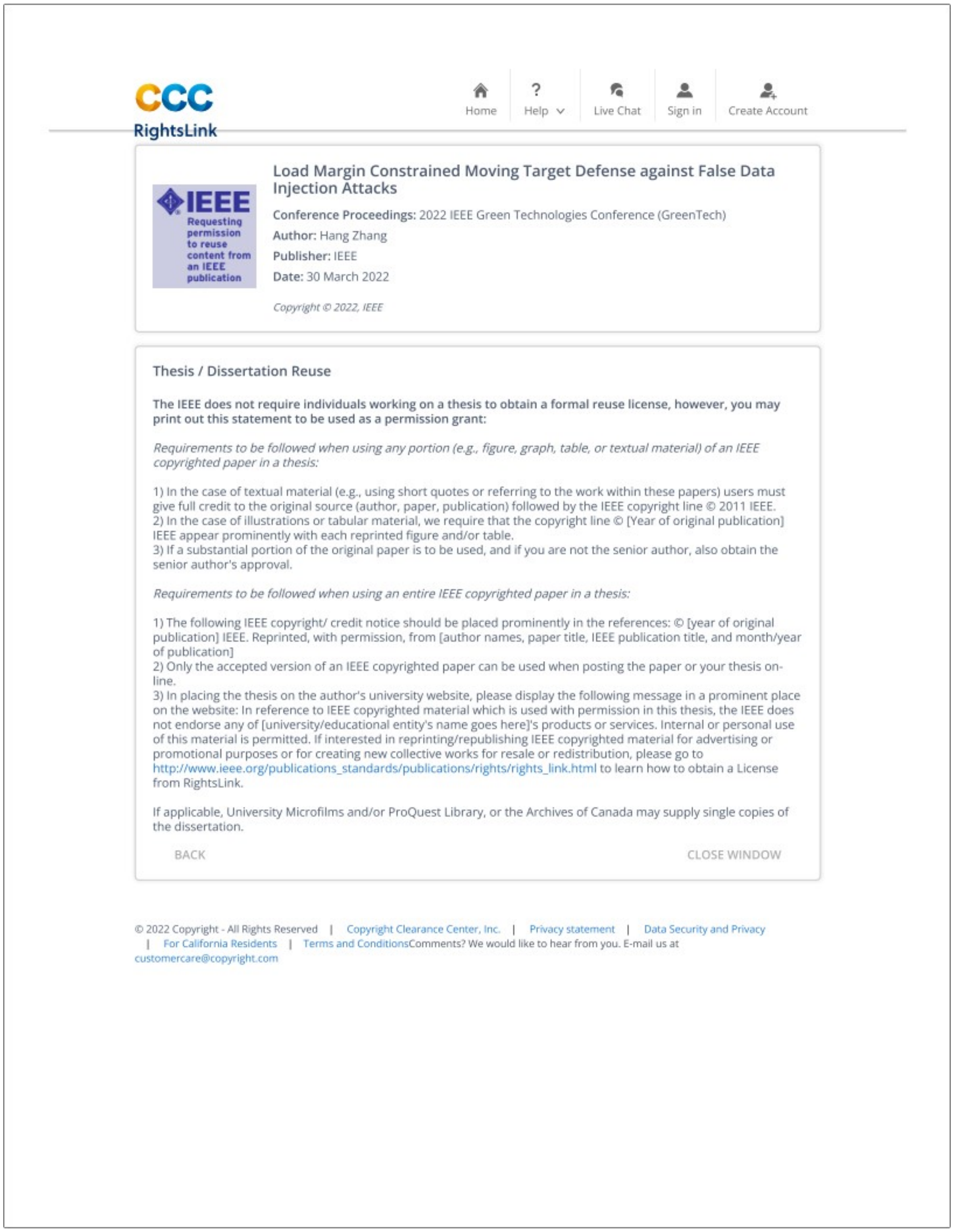
Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.


CCC
RightsLink
Home
Help
Live Chat
Sign in
Create Account
IEEE
Requesting permission to reuse content from an IEEE publication

Load Margin Constrained Moving Target Defense against False Data Injection Attacks

Conference Proceedings: 2022 IEEE Green Technologies Conference (GreenTech)
Author: Hang Zhang
Publisher: IEEE
Date: 30 March 2022

Copyright © 2022, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.